

UNIVERSIDAD NACIONAL DEL LITORAL



Desarrollo de un sensor inteligente para el monitoreo continuo de animales en pastoreo para ganadería de precisión

Luciano Sebastián Martínez Rau

FICH

FACULTAD DE INGENIERÍA Y CIENCIAS HÍDRICAS

INTEC

INSTITUTO DE DESARROLLO TECNOLÓGICO PARA LA INDUSTRIA
QUÍMICA

CIMEC

CENTRO DE INVESTIGACIÓN DE MÉTODOS COMPUTACIONALES

sinc(*i*)

INSTITUTO DE INVESTIGACIÓN EN SEÑALES, SISTEMAS E
INTELIGENCIA COMPUTACIONAL

Tesis de Doctorado **2022**



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas
Instituto de Desarrollo Tecnológico para la Industria Química

**DESARROLLO DE UN SENSOR INTELIGENTE
PARA EL MONITOREO CONTINUO DE
ANIMALES EN PASTOREO PARA
GANADERÍA DE PRECISIÓN**

Luciano Sebastián Martínez Rau

Tesis remitida al Comité Académico del Doctorado
como parte de los requisitos para la obtención
del grado de
DOCTOR EN INGENIERÍA
Mención Inteligencia Computacional, Señales y Sistemas
de la
UNIVERSIDAD NACIONAL DEL LITORAL

2022

Comisión de Posgrado, Facultad de Ingeniería y Ciencias Hídricas, Ciudad Universitaria, Paraje
“El Pozo”, S3000, Santa Fe, Argentina.



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas
Instituto de Desarrollo Tecnológico para la Industria Química

DESARROLLO DE UN SENSOR INTELIGENTE PARA EL MONITOREO CONTINUO DE ANIMALES EN PASTOREO PARA GANADERÍA DE PRECISIÓN

Luciano Sebastián Martínez Rau

Lugar de Trabajo:

Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional –
sinc(i), FICH-UNL/CONICET.

Director:

Dr. Leonardo Luis Giovanini *sinc(i)*-CONICET-UNL

Co-director:

Dr. Pablo Alejandro Kler CIMEC-CONICET-UNL

Jurado Evaluador:

Dr. Marcelo Risk UBA-CONICET
Dr. Omar Chiotti INGAR-CONICET-UTN
Dra. Gabriela Peretti UTN-FRVM

2022



UNIVERSIDAD NACIONAL DEL LITORAL
Facultad de Ingeniería y Ciencias Hídricas

Santa Fe, 28 de Marzo de 2022.

Como miembros del Jurado Evaluador de la Tesis de Doctorado en Ingeniería titulada **“Desarrollo de un sensor inteligente para el monitoreo continuo de animales en pastoreo para ganadería de precisión”**, desarrollada por el Ing. Luciano Sebastián MARTÍNEZ RAU, en el marco de la Mención “Inteligencia computacional, señales y sistemas”, certificamos que hemos evaluado la Tesis y recomendamos que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

La aprobación final de esta disertación estará condicionada a la presentación de dos copias encuadernadas de la versión final de la Tesis ante el Comité Académico del Doctorado en Ingeniería.

Dr. Marcelo Risk

Dr. Omar Chiotti

Dra. Gabriela Peretti

Santa Fe, 28 de Marzo de 2022.

Certifico haber leído la Tesis, preparada bajo mi dirección en el marco de la Mención “Inteligencia computacional, señales y sistemas” y recomiendo que sea aceptada como parte de los requisitos para la obtención del título de Doctor en Ingeniería.

Dr. Pablo Kler
Codirector de Tesis

Dr. Leonardo Giovanini
Director de Tesis



Universidad Nacional del Litoral
Facultad de Ingeniería y
Ciencias Hídricas

Secretaría de Posgrado

Ciudad Universitaria
C.C. 217
Ruta Nacional N° 168 - Km. 472,4
(3000) Santa Fe
Tel: (54) (0342) 4575 229
Fax: (54) (0342) 4575 224
E-mail: posgrado@fich.unl.edu.ar

DECLARACIÓN DEL AUTOR

Esta tesis ha sido remitida como parte de los requisitos para la obtención del grado académico de Doctor en Ingeniería - Mención Inteligencia Computacional, Señales y Sistemas ante la Universidad Nacional del Litoral y ha sido depositada en la Biblioteca de la Facultad de Ingeniería y Ciencias Hídricas para que esté a disposición de sus lectores bajo las condiciones estipuladas por el Reglamento de la mencionada Biblioteca.

Citaciones breves de esta tesis son permitidas sin la necesidad de un permiso especial, en la suposición de que la fuente sea correctamente citada. Solicitudes de permiso para la citación extendida o para la reproducción parcial o total de este manuscrito serán concebidos por el portador legal del derecho de propiedad intelectual de la obra.

Aclaración:

La presente tesis se encuentra organizada bajo el formato de Tesis por Compilación, aprobado en la resolución N° 255/17 (Expte. N° 888317-17) por el Comité Académico de la Carrera Doctorado en Ingeniería, Facultad de Ingeniería y Ciencias Hídricas, Universidad Nacional del Litoral (UNL). De dicha resolución:

“En el caso de optar por la Tesis por Compilación, ésta consistirá en una descripción técnica de al menos 30 páginas, redactada en español e incluyendo todas las investigaciones abordadas en la tesis. Se deberán incluir las secciones habituales indicadas a continuación en la Sección Contenidos de la Tesis. Los artículos científicos publicados por el autor, en el idioma original de las publicaciones, deberán incluirse en un Anexo con el formato unificado al estilo general de la Tesis indicado en la Sección Formato. El Anexo deberá estar encabezado por una sección donde el tesista detalle para cada una de las publicaciones cuál ha sido su contribución. Esta sección deberá estar avalada por su director de Tesis. El documento central de la Tesis debe incluir referencias explícitas a todas las publicaciones anexadas y presentar una conclusión que muestre la coherencia de dichos trabajos con el hilo conceptual y metodológico de la tesis. Los artículos presentados en los anexos podrán ser artículos publicados, aceptados para publicación (en prensa) o en revisión.”

Agradecimientos

En primer lugar quiero agradecer a mi familia por el apoyo incondicional.

A mis amigos y amigas que siempre se interesaron por mi investigación.

A mis directores, Dr. Leonardo Giovanini y Dr. Pablo Kler, por darme la confianza y la oportunidad de formar parte de un grupo de investigación. Su inconmensurable predisposición y ayuda en guiarme han permitido que alcance este objetivo.

A todos los miembros del grupo *chewbite*, un excelente grupo de trabajo y grupo humano, por haberme dejado formar parte del mismo y enseñado a valorar y explotar todas las ventajas del trabajo interdisciplinario.

A mis compañeros y compañeras del Instituto sinc(*i*) por las excelentes relaciones humanas y las charlas en los almuerzos.

Al Prof. Payá Vayá y demás miembros del grupo EIS por abrirme las puertas de su instituto y dejarme trabajar con ellos, permitiéndome vivir una experiencia única tanto en lo personal como en lo profesional.

Finalmente, quiero agradecer a las siguientes instituciones:

- sinc(*i*): Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional.
- FICH: Facultad de Ciencias Hídricas.
- UNL: Universidad Nacional del Litoral.
- CONICET: Consejo Nacional de Investigaciones Científicas y Técnicas.
- EIS: Abteilung Entwurf Integrierter Systeme (Departamento de Diseño de Sistemas Integrados) de la Universidad Técnica de Braunschweig, Alemania.
- TU Braunschweig: Universidad Técnica de Braunschweig, Alemania.
- DAAD: Servicio Alemán de Intercambio Académico.

Luciano S. Martínez Rau

Santa Fe, Abril de 2022.

Índice general

Resumen	ix
1. Introducción	1
1.1. Descripción del problema	1
1.2. Comportamiento alimentario del rumiante	2
1.2.1. Anatomía y fisiología digestiva	2
1.2.2. La rumia y el pastoreo	4
1.2.3. Eventos masticatorios	5
1.3. Objetivos	7
1.4. Antecedentes	8
1.5. Organización de la tesis	12
2. Metodología	14
2.1. Descripción general del sistema	14
2.2. Sistema de reconocimiento de patrones	14
2.3. Descripción del sistema propuesto	16
2.4. Clasificadores basados en aprendizaje maquina	18
2.5. Bases de datos	19
2.6. Medidas de desempeño	21
2.6.1. Reconocimiento de eventos	21
2.6.2. Reconocimiento de actividades alimentarias	22
3. Reconocimiento de eventos masticatorios	24
3.1. Análisis de características robustas	25
3.2. Descripción del algoritmo	28
3.3. Resultados y discusión	30
4. Reconocimiento de actividades alimentarias	33
4.1. Métodos propuestos	34
4.2. Resultados y discusión	39

5. Sensor acústico	43
5.1. Sistema Embebido	44
5.2. Sistema Híbrido	49
6. Conclusiones y Trabajo Futuro	56
6.1. Conclusiones	56
6.2. Artículos y patente	58
6.3. Trabajo futuro	59
Anexos	61
A. A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals	64
B. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle	106
C. Using jaw movement-agnostic features to recognize foraging activities in grazing cattle	138
D. Acoustic Real-time Sensor for Ingestive Behaviour of Grazing Cattle	186
E. Developments on real-time monitoring of grazing cattle feeding behavior using sound	199
F. Sistema para el seguimiento y evaluación de actividades alimentarias en producción animal basado en la energía del sonido	211

Índice de figuras

1.1.	Porciones del aparato digestivo de un rumiante vacuno.	4
1.2.	Registro acústico de actividades de rumia (a) y pastoreo (b).	5
1.3.	Zoom sobre registro acústico de un período de pastoreo donde se pueden observar los eventos masticatorios.	6
1.4.	Registro acústico de los cuatro tipos de eventos masticatorios asociados a los movimientos mandibulares.	6
1.5.	Zonas donde se generan los eventos masticatorios de interés.	7
2.1.	Diagrama de bloques general de un sistema de reconocimiento de patrones.	15
2.2.	Diagrama de flujo del sistema abordado para monitoreo alimentario.	17
3.1.	Señales de eventos masticatorios producidos por movimientos mandibulares y sus respectivas características.	26
3.2.	Análisis t-SNE del nuevo conjunto de características propuesto (a-b) y del conjunto de características usado en el CBIA (c-d) en términos de clases (GC: <i>grazing-chew</i> - RC: <i>ruminantion-chew</i> - C: <i>chew</i> - B: <i>bite</i> - CB: <i>chew-bite</i>).	27
3.3.	Diagrama en bloque de las etapas del CBEBA.	28
3.4.	Matrices de confusiones de eventos de CBEBA: <i>grazing-chew</i> (GC), <i>bite</i> (B), <i>chew-bite</i> (CB) y <i>ruminantion-chew</i> (RC) en las cuatro combinaciones de datasets.	31
4.1.	Actividades de pastoreo (a) y rumia (b). Las flechas en la rumia indican las pausas de deglución-regurgitación del alimento.	33
4.2.	Diagrama general del reconocedor de actividades.	35
4.3.	Diagrama general del BUFAR.	36
4.4.	Diagrama general del JMFAR.	37
4.5.	Matrices de confusión de los distintos métodos.	40
4.6.	Tasa de reconocimiento para distintos niveles de ruido.	42
5.1.	Diagrama de bloques del sistema embebido.	45

5.2. Diseño de la PCB del sistema integrado.	46
5.3. Ritmo circadiano del comportamiento de alimentación del ganado. . .	47
5.4. Vista macro (a) y zoom (b) sobre los principales eventos de consumo de corriente.	48
5.5. Consumo de corriente (línea negra) y voltaje (línea roja) durante 5,5 días.	49
5.6. Diagrama de bloques del sistema híbrido.	50
5.7. Diagrama de bloques del microprocesador conformado por procesa- dores (bloques azules) y memorias (bloques naranjas).	50
5.8. Partes de un procesador TTA.	53
5.9. Arquitectura del procesador TTA.	54

Índice de tablas

1.1. Escalas temporal y espacial utilizadas para describir y evaluar el comportamiento de rumiantes en pastoreo libre [13].	3
2.1. Ejemplo de alineamiento de secuencias de referencia y reconocida.	21
2.2. Métricas utilizadas para evaluar la detección de eventos.	22
2.3. Métricas utilizadas para evaluar la clasificación de eventos.	22
3.1. Comparación de resultados (media \pm desvió estándar; %) entre CBIA y CBEBA en las cuatro combinaciones de datasets.	31
4.1. Características utilizadas por los algoritmos propuestos.	38
4.2. Rendimiento general y costo computacional de las distintas variantes.	41
5.1. Consumo de corriente estimado y porcentaje de tiempo ocupado en cada modo de operación en cada módulo.	47
5.2. Comparación de resultados (media \pm desvió estándar; %) entre CBEBA utilizando representación en punto flotante y punto fijo en dos datasets.	55



A mi familia y amigos.

Resumen

La ganadería de precisión se basa en el uso y desarrollo de tecnologías de la información y las comunicaciones con el objetivo de lograr un monitoreo automático, preciso y en tiempo real del comportamiento animal, lo cual ayuda al productor a mejorar la gestión del ganado cambiando la escala de observación de rebaños a animales individuales. Con respecto al comportamiento animal, el monitoreo individualizado proporciona información importante y única sobre actividades colectivas relacionadas con el estado de salud, nutrición y bienestar de los animales. Así, estas tecnologías permiten optimizar y controlar los recursos de los sistemas de producción ganadera permitiéndoles a los productores mantener la competitividad en un mercado cada vez más desafiante.

Esta tesis se ocupa del desarrollo de un sistema para el monitoreo del comportamiento alimentario de rumiantes utilizando señales acústicas, lo que es imprescindible para conocer el estado nutricional del animal, estimar la cantidad de materia seca ingerida aumentando la eficiencia de los recursos pastoriles y predecir celo del animal en forma temprana, entre otros. Para ello fue necesario diseñar, desarrollar e implementar algoritmos de bajo costo computacional que puedan operar en sensores acústicos con capacidad de computo reducida como partes de un sistema de monitoreo alimentario. Los algoritmos desarrollados se encargan de (i) reconocer los eventos masticatorios (masticación durante la rumia, masticación durante el pastoreo, arranque y arranque-masticación) producidos a una escala temporal a nivel de segundos, y (ii) reconocer y estimar los tiempos de las actividades alimentarias (rumia y pastoreo) producidos a una escala temporal a nivel de minutos u horas. Se utilizaron diversas técnicas de procesamiento de señales y aprendizaje maquina durante el diseño y desarrollo de los distintos algoritmos permitiéndoles alcanzar buena robustez ante condiciones operativas cambiantes. El desempeño logrado por los distintos algoritmos fue superior a los métodos del estado del arte y, al mismo tiempo, con un menor costo computacional en la mayoría de los casos. También se diseñaron dos sensores acústicos capaces de reconocer eventos masticatorios de forma autónoma.

Abstract

Precision livestock farming is based on the use and development of information and communication technologies to achieve automatic, precise and real-time monitoring of animal behaviour, in order to help farmers to improve livestock management by changing the observation scale from herds to individual animals. Individualised behaviour monitoring provides important and unique information about collective activities related to the health, nutrition and welfare of animals. Thus, these technologies make it possible to optimise and control the resources of livestock production systems, allowing cattle farmers to remain competitive in an increasingly challenging market.

This thesis develops a livestock feeding behaviour monitoring system using acoustic signals, which is essential to know the nutritional status of the animal, to estimate the forage dry matter intake, to increase the efficiency of pastoral resources and to predict rut periods early, among others. It was necessary to design, develop and implement low-cost computational algorithms capable of operating on acoustic sensors with reduced computing capability as parts of a feeding monitoring system. The algorithms developed are capable of (i) recognising jaw movements (grazing-chew, rumination-chew, bite and chew-bite), which occurs on a short-time scale of seconds, and (ii) recognising and estimating the duration of feeding activities (rumination and grazing), which occurs on a long-time scale from minutes to hours. Various signal processing and machine learning techniques were used during the design and development of the algorithms, allowing them to achieve good robustness to changing operating conditions. The different algorithms achieve a superior performance to the methods of the state of the art, and with lower computational cost in most of the cases. Two autonomous acoustic sensors capable of recognising jaw movements were also designed.

1 Introducción

1.1. Descripción del problema

Se estima que en las próximas décadas la demanda mundial de carne y lácteos aumentará en más de un 40 %. Surge así el interrogante de como poder satisfacer esa demanda y, al mismo tiempo, lograr una producción de buena calidad, sostenible, segura y amigable con el medio ambiente. Además, los productores ganaderos, tanto de leche como de carne vacuna (cría e invernada), se enfrentan a diferentes dificultades cuando tratan de mejorar la producción y mantener la rentabilidad de sus establecimientos. Estas dificultades están en parte asociadas a la falta de información precisa sobre el manejo correcto de la alimentación y el estado de salud de los animales, lo cual tiene como consecuencia una baja eficiencia de productividad (conversión del alimento del animal en producto (leche y/o carne) debido a una gran pérdida de recursos en el proceso de alimentación.

Desde hace aproximadamente dos décadas se han desarrollado varios dispositivos y métodos que ayudan a los productores a mejorar la eficiencia de la alimentación, en forma particular, y del sistema de producción, en forma general. Estos dispositivos permiten la supervisión automatizada de la salud animal y brindan información sobre su estado nutricional. Una herramienta que ofrece posibilidades reales es la ganadería de precisión. Estos sistemas han cobrado relevancia en los últimos años y apuntan a ofrecer un sistema de seguimiento y gestión en tiempo real con el objetivo básico de mejorar la vida de los animales, alertando de eventuales problemas de salud para que el ganadero pueda tomar medidas inmediatas. Esto debería hacer posible un seguimiento continuo y completamente automatizado de la salud y el bienestar de los animales, su rendimiento y su impacto ambiental [1].

La producción láctea y cárnica en bovinos está altamente relacionada con la ingestión de forraje [2], y por eso su cuantificación y optimización es uno de los principales objetivos en una explotación comercial [3]. En este sentido, existen numerosos indicadores relacionados al estado nutricional del rumiante. Por ejemplo, el número total de masticaciones por unidad de alimento (producidos principalmente durante la rumia) está asociado a la reducción de las partículas y a la cantidad de saliva producida por el animal. Cuanto menor es el tamaño de las partículas y mayor la cantidad de saliva, mejor es el aprovechamiento de los nutrientes y se contribuye a mantener un adecuado ambiente ruminal [4], lo cual influye indirectamente sobre la producción y la salud de los animales. Además, existen otros factores como la

medición del consumo de materia seca ingerida [5, 6] y los tiempos de pastoreo y rumia que permiten determinar con antelación si existen problemas asociados con diversos estados de estrés [7], ansiedad [8] o enfermedad [9] que estén afectando a la salud o el rendimiento de los animales. Estos factores también permiten reflejar las condiciones fisiológicas del animal y las condiciones de las pasturas [10, 11], como así también, predecir cuando el animal se encuentra en celo [12]. Por otro lado, la producción de leche está estrechamente ligada a la ingestión de agua, la cual está directamente relacionada con las condiciones climáticas y el contenido de materia seca del alimento ingerido [4]. En consecuencia, los cambios en el patrón diario de estas actividades pueden explicar en gran parte el resultado productivo y/o poner de manifiesto condiciones limitantes en los sistemas de producción.

1.2. Comportamiento alimentario del rumiante

En condiciones de pastoreo libre, los animales rumiantes exhiben diferentes patrones de comportamiento alimentario que varían tanto en la superficie abarcada por el animal como en el tiempo empleado para realizar dicho desplazamiento [13]. De esta manera su comportamiento puede ser estudiado a corta escala espacio-temporal, donde se evalúan comportamientos como la selección de la dieta, o a gran escala donde se dan comportamientos de migración de superficies y trashumancia, en la que los animales se desplazan en el espacio adaptándose a zonas de alimentación cambiante. En la Tabla 1.1 se pueden observar las distintas escalas espacio-temporales relacionadas a las decisiones de pastoreo. Las escalas varían desde simples movimientos mandibulares (1-2 segundos) hasta áreas que cambian con mucho menos frecuencia (1 mes a 2 años).

1.2.1. Anatomía y fisiología digestiva

Los rumiantes son mamíferos herbívoros que digieren los alimentos (pasto o forraje) en dos etapas: primero el alimento es ingerido y preprocesado durante el pastoreo, para luego ser reprocesado durante la rumia. La rumia consiste en la regurgitación del material semi-digerido, re-masticación (desmenuzado) y agregación de saliva. Dentro del grupo de los rumiantes se encuentran los bovinos, caprinos, ovinos y cérvidos.

Para que la digestión de un rumiante sea posible es necesario un proceso de fermentación, el cual es llevado a cabo gracias a la presencia de microorganismos en distintas porciones de su aparato digestivo. Esta característica les permite degradar los hidratos de carbono estructurales del forraje, como la celulosa, la hemicelulosa y la pectina, las cuales son muy poco digeribles para las especies de estómago simple

Nivel espacial	Resolución espacial	Escala temporal	Características comportamiento	Información obtenida	Vegetación
Movimiento mandibular	0,0001 – 0,01m ²	1 – 2s	Movimientos de mandíbula, lengua y cuello	Característica del movimiento	Porción de planta
Estación de alimentación	0,1 – 1m ²	2s – 2min	Alimentación en línea recta	Frecuencia de movimientos	Planta o arbusto
Parcela de alimentación	1m ² – 1ha	1 – 30min	Pausa en el pastoreo para cambiar de orientación	Duración del bloque de alimentación	Grupos de plantas
Sitio de alimentación	1 – 10ha	1 – 4h	Turno de pastoreo	Movimientos de pastoreo	Asociación de especies vegetales
Rango diario	10 – 100ha	12 – 24h	Área de alimentación, descanso y bebida	Asignación del tiempo diario	Unidad de paisaje
Rango estacional	100 – 1000ha	3 – 12 meses	Migración	Asignación metabólica	Tipo de paisaje
Rango de por vida	> 1000ha	Varios años	Dispersión o migración	Registro histórico de vida	Región geográfica

Tabla 1.1: Escalas temporal y espacial utilizadas para describir y evaluar el comportamiento de rumiantes en pastoreo libre [13].

o no rumiantes. Basada en esta diferencia fundamental, la fisiología digestiva del rumiante adquiere características particulares.

Una característica distintiva de los rumiantes es su aparato digestivo poligástrico compuesto por cuatro divisiones (Figura 1.1): el retículo (red o redecilla), el rumen (panza), el omaso (libro o librilla) y el abomaso (cuañar). Sólo este último es glandular y funcionalmente análogo al estómago de los no rumiantes, mientras que las otras tres divisiones están cubiertas por un epitelio queratinizado y carecen de glándulas.

El retículo toma su nombre de la disposición en forma de red de los pliegues de su mucosa y está situado en contacto con el diafragma, comunicándose con el rumen a través del pliegue retículo-ruminal que los convierte en una sola unidad funcional (retículo-rumen). La función del retículo es movilizar el alimento digerido hacia el rumen o hacia el omaso en la regurgitación del bolo alimenticio después de la rumia. El rumen es el compartimiento más voluminoso y está en contacto con la pared abdominal izquierda. Allí se encuentran microorganismos anaeróbicos que en simbiosis con el animal le permiten digerir la fibra contenida en los forrajes (especialmente celulosa y hemicelulosa). Estas sustancias fibrosas se degradan hasta glucosa. Las bacterias utilizan la mayor parte de la glucosa y generan ácidos carboxílicos volátiles que constituyen la principal fuente de energía de los rumiantes, ya que dentro de su organismo se absorben y se metabolizan.

El omaso se ubica a la derecha del retículo y posee forma esférica. Su función principal es la absorción de agua, sales minerales y ácidos grasos contenidos en la ingesta, y se comunica con el retículo por el esfínter retículo-omasal y con el abomaso por el esfínter omaso-abomasal.

Por último, el abomaso se ubica a la derecha y ventralmente en la cavidad abdominal, tiene forma de saco alargado con un extremo ciego denominado fundus y un extremo pilórico que desemboca en el duodeno. La mucosa es de tipo glandular y en el fundus presenta pliegues que aumentan su superficie. En su interior a los microorganismos ruminales, que en buena parte están constituidos por proteína de alta calidad, se les somete a digestión hasta aminoácidos y péptidos, y posteriormente a absorción en el intestino delgado.

1.2.2. La rumia y el pastoreo

Los animales rumiantes dedican más de 12 horas por día a realizar dos actividades alimentarias principales: rumia y pastoreo (Figura 1.2) [14-17]. El resto del tiempo los animales están descansando, bebiendo, en ordeño o simplemente está en búsqueda de nuevo alimento.

El acto de pastoreo involucra la selección del forraje, su aprehensión, masticación, insalivación y el tragado de cada uno de los bocados. El animal varía el número y peso de los bocados que toma en función de la estructura de la vegetación (cantidad de materia seca disponible, altura de las plantas, densidad del forraje y cobertura) determinando una tasa de consumo. Esta tasa de consumo, junto con el tiempo de pastoreo, definen la cantidad de forraje consumido.

La rumia consiste en la regurgitación del alimento ingerido para continuar con

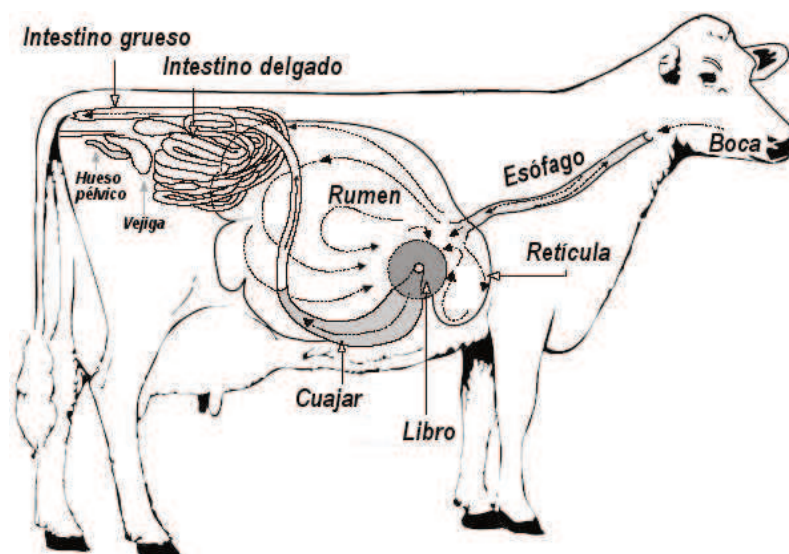


Figura 1.1: Porciones del aparato digestivo de un rumiante vacuno.

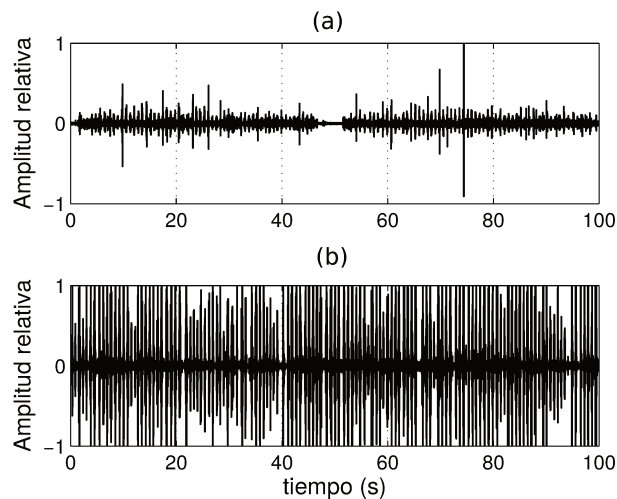


Figura 1.2: Registro acústico de actividades de rumia (a) y pastoreo (b).

la masticación del mismo. Durante este proceso se consigue una adecuada reducción de las partículas fibrosas para facilitar su paso al resto del tracto gastrointestinal y al exterior, ya que una porción de la fibra es de consistencia leñosa y aún para los microorganismos ruminales resulta indigerible. La rumia es un proceso repetitivo de 40-60 s de duración compuesto solo por masticaciones y con un patrón bien marcado de pausas de deglución y regurgitación de a 3-7 s de duración (Figura 1.2) [18-20].

La masticación durante el pastoreo y la rumia es esencial para reducir el tamaño de las partículas del alimento y aumentar la relación superficie/volumen expuesta a la acción de los microorganismos. Además se estimula la secreción de saliva, la cual contiene altos niveles de bicarbonatos y fosfatos que afectan la concentración de los iones de hidrógeno y colaboran en mantener un nivel de pH ruminal adecuado para la actividad microbiana.

1.2.3. Eventos masticatorios

Durante su alimentación (pastoreo y rumia), un bovino mueve la mandíbula de forma semi continua para realizar un total de 40.000 movimientos mandibulares al día, aproximadamente unos 10.000 durante el pastoreo y 30.000 durante la rumia (Figura 1.3). En un rumiante se pueden diferenciar tres tipos de eventos masticatorio asociados a los movimientos mandibulares producidos durante el pastoreo (Figura 1.4): *chew* (masticación), *bite* (arranque) y el movimiento combinado *chew-bite* (arranque-masticación). Durante la rumia el único tipo de evento masticatorio asociado a los movimientos mandibulares es el *chew* (masticación).

El evento *chew* se caracteriza por la no ingestión de alimento nuevo, sino que el animal mastica triturando y reprocesando lo que ya se encuentra en su boca con la ayuda de sus piezas dentales posteriores (molares y premolares) (Figura 1.5). Si

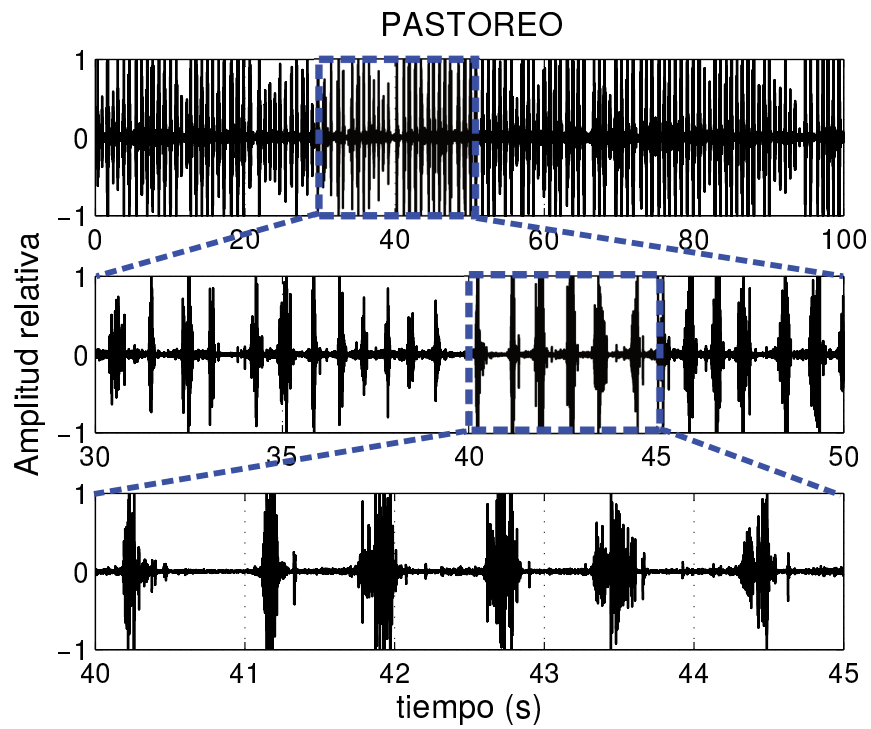


Figura 1.3: Zoom sobre registro acústico de un período de pastoreo donde se pueden observar los eventos masticatorios.

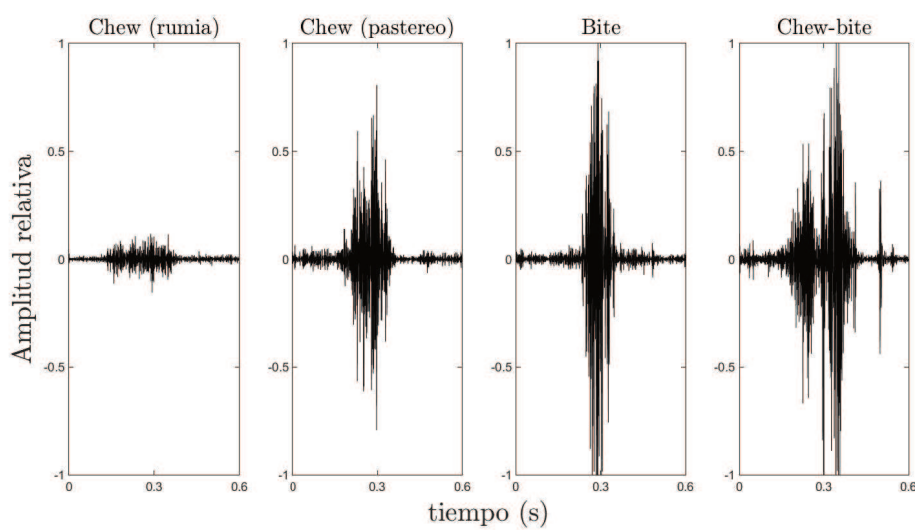


Figura 1.4: Registro acústico de los cuatro tipos de eventos masticatorios asociados a los movimientos mandibulares.

bien el movimiento mandibular producido durante ambas actividades alimentarias es el mismo, es posible distinguir en forma sonora los eventos *chews* producidos durante el pastoreo o la rumia debido al contenido de agua del material masticado, denominándolos *grazing-chew* (masticación en pastoreo) y *ruminantion-chew* (masticación en rumia) (Figura 1.4). Durante el pastoreo, el forraje masticado contiene su humedad interna y el sonido producido se debe en parte a la ruptura de las paredes celulares y a la extrusión del agua, mientras que durante la rumia, se mastica el bolo regurgitado, el cual ya ha sido parcialmente procesado y degradado en el rumen, donde se satura con los fluidos ruminales conteniendo gran cantidad de líquido [21]. En tanto, un evento *bite* se puede definir como la aprehensión de la vegetación y corte entre los dientes incisivos en la parte inferior y el rodete dentario en la parte superior [22] (Figura 1.5). Finalmente, el evento *chew-bite* es un evento compuesto que resulta de la superposición de los dos eventos simples mencionados en ese orden.

La importancia de considerar los *chew-bites* fue detectada al introducir al audio como método de obtención de los datos [22] y ha sido analizada en varios trabajos [23, 24], siendo uno de los principales predictores utilizados para estimar el consumo de materia seca ingerida por los animales [5]. En parte, la importancia de este tipo de evento radica en el ahorro de tiempo que se produce al masticar y arrancar en forma casi simultánea, afectando así la tasa de consumo.

1.3. Objetivos

Objetivo general

A lo largo de esta tesis se plantea el desarrollo, implementación y evaluación de un sistema robusto de adquisición, procesamiento y análisis de información para ganadería de precisión, utilizando técnicas de procesamiento de señales y aprendizaje

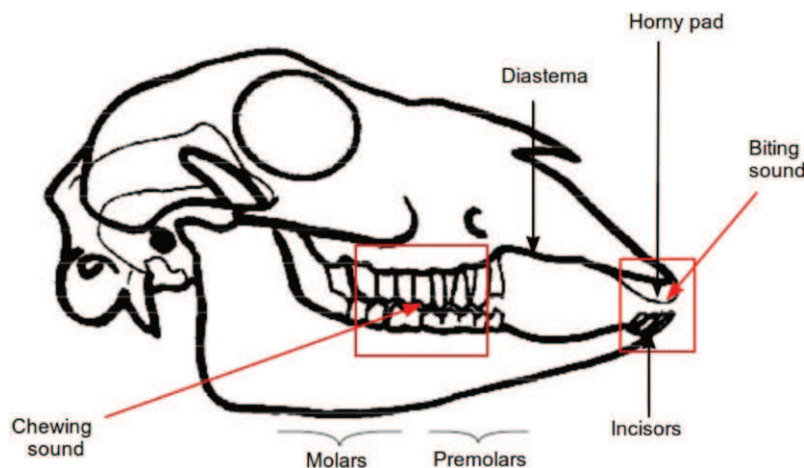


Figura 1.5: Zonas donde se generan los eventos masticatorios de interés.

maquinal para analizar señales acústicas derivadas de la alimentación del rumiante que puedan ser ejecutados en tiempo real en un sensor acústico con capacidad de cómputo reducida.

Objetivos particulares

Los objetivos específicos para lograr el objetivo general propuesto son:

- Diseñar un algoritmo robusto capaz de detectar y clasificar eventos masticatorios a partir de los sonidos producidos por los animales rumiantes durante las actividades alimentarias utilizando herramientas de procesamiento de señales y aprendizaje maquinal;
- Implementar y evaluar el algoritmo de detección y clasificación de eventos desarrollado en el punto anterior con señales registradas en distintas condiciones operativas de pastoreo libre;
- Diseñar un algoritmo robusto para la clasificación de actividades alimentaria basado en la clasificación de eventos y otras fuentes de información, utilizando herramientas de aprendizaje maquinal;
- Implementar y evaluar el algoritmo de clasificación de actividad alimentaria desarrollado en el punto anterior con señales registradas en distintas condiciones operativas de pastoreo libre;
- Diseñar un sensor acústico capaz de ejecutar los algoritmos en tiempo real y poseer autonomía energética.

1.4. Antecedentes

Uno de las primeras formas utilizadas para caracterizar la actividad alimentaria de rumiantes era a través de la medición de la digestibilidad en forma indirecta [25]. Otro antiguo método utilizaba grabaciones de movimientos mandibulares obtenido mediante transductores mecánicos [26]. Sin embargo, ninguno de estos métodos tuvo éxito y durante mucho tiempo la caracterización del comportamiento ingestivo de rumiantes era realizada por una persona mediante la observación directa del animal, antes de que se comiencen a desarrollar métodos que automaticen esta tarea. Por supuesto la observación directa presenta ciertas limitaciones asociadas a la falta de objetividad, la imposibilidad de un análisis minucioso del comportamiento y la falta de concentración del observador durante períodos largos. Estos factores hacen que esta metodología sea tediosa, ineficiente y difícil de implementar a gran escala [27].

Los primeros sistemas automáticos de monitoreo alimentario estaban basados en la utilización de diferentes dispositivos colocados en la cabeza del animal, que detectaban el movimiento mandibular para saber si el animal estaba realizando algún tipo de actividad ingestiva. Uno de estos sistemas utilizaba un tubo de goma que cruzaba por debajo de la mandíbula inferior del animal, y mediante circuitos eléctricos se detectaba cuando el circuito era interrumpido entre dos puntos de contacto. Cada interrupción en el circuito se registraba como una masticación [28]. La principal desventaja de este método es que no permite determinar si el animal está pastoreando o rumiando, ya que sólo detecta movimientos mandibulares.

Gracias al avance de la industria y de la tecnología en el área de la ganadería de precisión, han surgido en las últimas décadas diversas técnicas de monitoreo del comportamiento alimentario utilizando distintos tipos de transductores portables, como por ejemplo, sensores de presión nasal [29-35], acelerómetros [36-39], unidades de medición inercial [40-42], señales de electromiografía [43, 44] y micrófonos o grabadores acústicos. Estudios comparativos entre los acelerómetros y los transductores acústicos demostraron las ventajas de utilizar estos últimos para monitorear el comportamiento alimentario de corta escala, es decir, reconocer eventos masticatorios [23]. A su vez, un cambio o movimiento en la posición de los transductores mecánicos podría traer datos inestables o incluso incorrectos impactando negativamente en la precisión de la identificación del comportamiento alimentario [45]. Además, se demostró que el análisis del sonido de las masticaciones en bovinos permite mediciones aceptables del consumo de materia ingerida en distintos tipos de forrajes y que la energía del sonido de las masticaciones está altamente relacionada con la cantidad de materia seca ingerida por el animal [5, 46, 47]. Con ello, se comprobó que las características acústicas poseen gran cantidad de información relacionada con el consumo. Por estos motivos, el método acústico es el más utilizado para reconocer eventos masticatorios en rumiantes.

A finales de la década de los 80', se comenzó a utilizar el sonido en los estudios de comportamiento animal. Uno de los primeros sistemas desarrollados utilizando esta metodología, fue capaz de detectar y clasificar correctamente las acciones de puercoespines (comer, beber, caminar) [48].

Los primeros trabajos utilizando señales acústicas para caracterizar la actividad alimentaria de rumiantes combinaban las grabaciones acústicas con grabaciones de vídeo [49, 50] o con un interruptor de mercurio situado debajo de la mandíbula inferior para determinar la inclinación de la cabeza durante las masticaciones, el cual resulta invasivo para el animal [51].

En trabajos posteriores se lograron avances en el reconocimiento de los sonidos masticatorios en ovinos [6, 27] y bovinos [52] a través de la adaptación de métodos del área del reconocimiento automático del habla, utilizando modelos ocultos de

Markov para segmentar y clasificar los eventos masticatorios durante el pastoreo de diferentes especies y alturas de forraje, con resultados satisfactorios en condiciones controladas con bajo niveles de ruido.

En los años siguientes, Clapham et al. [53] utilizaron un software para analizar el sonido del ganado producido durante el pastoreo. Este método solo es capaz de detectar las mordeduras (*bites*) y alcanza buenos resultados, pero exige una calibración cuidadosa y específica del sitio antes de ser empleado con diferentes animales, pasturas o condiciones experimentales. Además, este sistema presentó dificultades al tratar de escalar el monitoreo a varios días. Otro método basado en el análisis de las características del dominio temporal del sonido registrado para la detección de eventos masticatorios fue propuesto por Navon et al. [54], mostrando buenas tasas de desempeño. El algoritmo utiliza técnicas de aprendizaje automático para identificar los ruidos de fondo presentes en los sonidos grabados. Sin embargo, el algoritmo no realiza la clasificación de los eventos masticatorios en sus diferentes tipos.

Posteriormente, Chelotti et al. [55] desarrollaron un algoritmo capaz de detectar y clasificar eventos masticatorios en términos de *chew*, *bite* y *chew-bite* mediante el análisis de señales acústicas en el dominio temporal logrando un desempeño similar a [52] a un costo computacional mucho menor, lo cual permitió implementarlo como parte de un dispositivo embebido [56]. Sin embargo, este algoritmo presenta algunos inconvenientes relacionados con que sus parámetros deben ser ajustados manualmente, y que su desempeño se ve muy degradado ante cualquier tipo de ruido presente en las señales acústicas. Estos problemas fueron solucionados por Chelotti et al. [57] quienes incorporan al algoritmo anterior una etapa de preprocesamiento utilizando técnicas de procesamiento de señales y utilizan un clasificador basado en técnicas de aprendizaje maquina, sin aumentar considerablemente el costo computacional, logrando mejores resultados de detección y clasificación que [52, 55] en condiciones controladas con bajo ruido.

Mientras que los transductores acústicos son el método predilecto para abordar el reconocimiento de eventos masticatorios, los acelerómetros son el transductor más utilizado para el reconocimiento de actividades alimentarias de largo plazo. Esto se debe a que muchos métodos, además de reconocer actividades alimentarias, también reconocen otras actividades relacionadas a posición o movimiento. Martiskainen et al. [39] fueron capaces de identificar dentro de un galpón las actividades alimentarias de rumia e ingesta de silaje, y los movimientos acostarse, ponerse de pie, estar de pie, estar acostado y caminar normal y cojo. Diosdado et al. [58] utilizaron un acelerómetro y un podómetro para identificar si el animal localizado dentro de un establo se encuentra acostado, de pie o ingiriendo silaje. Rayas-Amor et al. [59] utilizaron un equipo comercial HOB0 (Onset Computer Corporation, Bourne, MA, Estados Unidos) que almacena los datos obtenidos de un sensor de aceleración y de

inclinación para cálculos tiempos de rumia y pastoreo. Sin embargo, la clasificación está basada en las componentes estáticas de los ejes del acelerómetro, lo cual no es práctico en situaciones reales donde un ligero movimiento de la vaca podría cambiar el valor de las componentes estáticas dentro del mismo comportamiento [60]. Benaissa et al. [61] desarrollaron un método utilizando un acelerómetro para reconocer las actividades de rumia y alimentación, estando esta última conformada por la actividad de pastoreo y bebida de agua. Finalmente, Shen et al. [62] obtuvieron mejores resultados que Benaissa et al. [61] a un menor costo computacional posibilitando implementar dicho método en un dispositivo embebido.

Algunos autores indican que el uso de acelerómetros presenta mayor inmunidad al ruido que el uso de micrófonos para reconocer actividades alimentarias. Sin embargo, ninguno de los métodos anteriores [39, 58, 59, 61, 62] evalúa su afectación ante movimientos en la posición de los acelerómetros. En este sentido, deberían realizarse experimentos de mayor duración para ver este efecto. Además, a mi leal saber y entender, aún no se ha desarrollado un método reconocedor de eventos masticatorios utilizando acelerómetros como fuente de información.

Un sistema comercial para el monitoreo de actividades alimentarias en rumiantes que incluye un registrador de señales acústicas con un micrófono integrado es el Hi-Tag (SCR Engineers Ltd., Netanya, Israel), el cual está basado en una patente [63]. Su diseño está enfocado en la autonomía, portabilidad y robustez del hardware. Sin embargo, dicho sistema solo informa la tasa de masticaciones y el tiempo de rumia en periodos de 2 horas, no brindando otra información que pudiese resultar de interés [64-67].

Un método reconocedor de actividades alimentarias utilizando señales acústicas, a partir de analizar sus regularidades fue propuesto por Vanrell et al. [68], obteniendo resultados superiores a los de Hi-Tag. Este método reconoce las actividades de rumia y pastoreo, y estima sus respectivas duraciones. Sin embargo, este método requiere tener toda la señal registrada previo a realizar el análisis, lo cual imposibilita ser ejecutado en tiempo real y ser implementado en un dispositivo electrónico de bajo consumo con capacidad limitada de almacenamiento.

Finalmente, en los últimos años ha habido un gran avance de las técnicas de aprendizaje maquina en el área de las redes neuronales profundas y en el área de la visión artificial. Distintos autores han empleado estas técnicas utilizando grabaciones obtenidas con cámaras de vídeo dentro de establos o galpones para reconocer, entre otras cosas, cuando el animal esta alimentándose [69-71] o rumiando [72], o para estimar la cantidad de materia ingerida [73]. Otros autores han desarrollado métodos basados en redes neuronales profundas utilizando señales acústicas para reconocer eventos masticatorios en bovinos [74] y en ovinos [75] en pastoreo libre, debido al gran área a cubrir y al libre movimiento de los animales. Si bien estos métodos obtienen

muy buenos resultados, las cámaras de vídeos no pueden ser utilizadas en condiciones de pastoreo libre. Además, los métodos basados en redes neuronales profundas requieren un mayor costo computacional que las técnicas clásicas de aprendizaje maquinal, haciendo más difícil su implementación en dispositivos embebidos.

1.5. Organización de la tesis

Esta tesis se encuentra organizada bajo el formato de Tesis por compilación de la siguiente forma:

- En el presente capítulo se describió la motivación y los problemas a abordar dentro del marco de esta tesis. Además, se introdujeron los conceptos necesarios para la comprensión del problema describiendo el comportamiento alimentario de rumiantes. Se establecieron los objetivos planteados, y se presentaron los antecedentes y las técnicas existentes en el estado del arte que han intentado dar solución al problema planteado.
- En el capítulo 2 se presenta brevemente el sistema propuesto siguiendo un modelo de reconocimiento de patrones, para luego llegar a la descripción de las metodologías necesarias para abordar el problema. Estas metodologías son básicamente para: (i) el reconocimiento de eventos masticatorios (*chew* durante el pastoreo, *chew* durante la rumia, *bite* y *chew-bite*), y (ii) el reconocimiento de actividades alimentarias (rumia y pastoreo). Además, se presentan de forma breve los métodos de aprendizaje maquinal utilizados durante las fases de diseño del sistema, como así también, las bases de datos donde serán probados los distintos reconocedores y las métricas utilizadas para cuantificar los resultados obtenidos.
- En el capítulo 3 se presenta en forma detallada un conjunto de características robustas ante ruidos típicos en condiciones de pastoreo libre. Además, se presenta un algoritmo de bajo costo computacional para el reconocimiento de eventos masticatorios, el cual realiza tres tareas: (i) detección de eventos, (ii) clasificación de eventos y (iii) ajuste de parámetros internos. La etapa de detección se basa en umbrales variables y adaptativos, en tanto la etapa de clasificación se basa en un perceptrón multicapa. El ajuste de parámetros permite al algoritmo adaptarse a condiciones operativas cambiantes o tendencias.
- El capítulo 4 presenta dos algoritmos, y algunas variantes de cada uno de ellos, para el reconocimiento de la rumia y el pastoreo. Ambos algoritmos siguen un esquema básico de reconocimiento de patrones. Uno de los algoritmos utiliza

los eventos reconocidos por el algoritmo presentado en 3 como información, mientras que el otro algoritmo utiliza otras fuentes de información.

- En el capítulo 5 se presentan dos sensores acústicos diseñados para la aplicación de los algoritmos propuestos. Ambos sensores trabajan en forma autónoma pudiendo comunicarse en forma inalámbrica con una estación base.
- Finalmente, en el capítulo 6 se presentan las conclusiones generales y específicas de la tesis, y los trabajos que aún quedan por abordar.

2 Metodología

2.1. Descripción general del sistema

Muchos de los estudios reportados hasta el momento presentan distintos métodos tanto para reconocer eventos masticatorios (*chew* (de rumia y/o pastoreo), *bite* y *chew-bite*) como para clasificar y estimar los tiempos de actividades alimentarias (rumia y pastoreo). Sin embargo, ninguno de estos métodos forman parte o proponen un sistema integral completo capaz de brindar una caracterización pormenorizada, tanto a nivel de eventos masticatorios como a nivel de actividades alimentarias en rumiantes. A su vez, la mayoría de estos métodos no han sido lo suficientemente examinados para asegurar su correcto funcionamiento en diversas condiciones operativas, y además, algunos requieren una gran capacidad de procesamiento y/o de almacenamiento imposibilitando su implementación en dispositivos electrónicos. Es fundamental contar con un sistema robusto capaz de analizar en tiempo real la actividad alimentaria en rumiantes en distintas condiciones operativas, sin la necesidad de almacenar los sonidos propiamente. En este sentido, el grupo de trabajo cuenta con una patente internacional en la que propone el desarrollo de un sistema integral completo como parte de un sensor acústico [76].

En esta tesis se utilizaron características temporales de los sonidos registrados para reconocer eventos masticatorios de corta resolución temporal. Esto evita la necesidad del calcular el espectro de frecuencia y, por lo tanto, disminuye el costo computacional de procesamiento. Sin embargo, se utilizaron características temporales, espectrales y estadísticas de los sonidos para reconocer actividades alimentarias de larga resolución temporal sin aumentar considerablemente el costo computacional de procesamiento, dado que las características son calculadas en escasas ocasiones. A partir de esta descripción es posible proponer y desarrollar un sistema integral completo que utilice información a diferentes escalas temporales con el fin de clasificar y cuantificar la actividad alimentaria (con duraciones en el orden de los minutos a horas) a partir del análisis de la secuencia de eventos masticatorios detectados o clasificados (con duraciones en el orden del segundo).

2.2. Sistema de reconocimiento de patrones

Los distintos algoritmos diseñados y desarrollados en esta tesis se basan en sistemas de reconocimiento de patrones. Un sistema de reconocimiento de patrones es

un sistema automático cuyo objetivo es clasificar datos de entrada en un conjunto de clases o categorías específicas usando propiedades y características distintivas de las clases. De esta forma, se busca diseñar modelos que sean capaces de reconocer y clasificar patrones dentro de ciertas categorías para su futuro uso. Este sistema puede ser descrito por una serie de etapas genéricas (Figura 2.1) que permiten: (i) la descripción y análisis de la señal de entrada, la cual facilita la extracción de características discriminativas y (ii) el reconocimiento y clasificación de los patrones correspondientes en termino de clases [77].

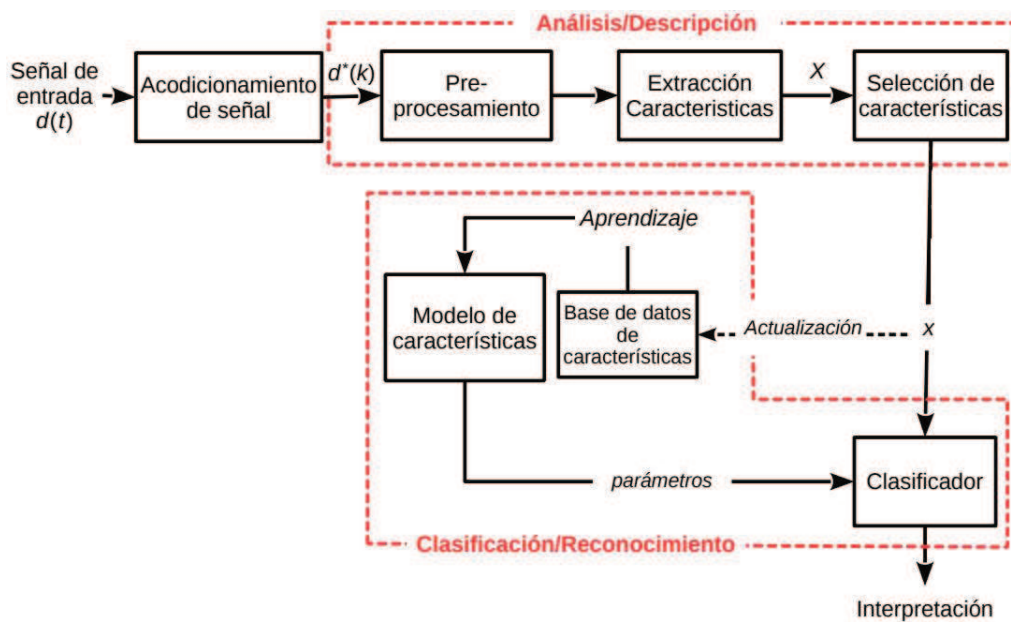


Figura 2.1: Diagrama de bloques general de un sistema de reconocimiento de patrones.

En un sistema de reconocimiento de patrones generalmente su entrada es la información de salida de algún tipo de transductor que transforma un tipo de energía en otra. Un ejemplo es la utilización de micrófonos para transformar las señales acústicas en señales eléctricas. En la etapa de acondicionamiento de señal se utilizan técnicas analógicas y digitales para adecuar la señal de entrada ($d(t)$) a los requerimientos de la etapa siguiente. Esta adecuación normalmente puede incluir la normalización, ajuste del rango dinámico, digitalización y filtrado de $d(t)$, entre otros. En tanto en la etapa de preprocesamiento, la señal digital acondicionada ($d^*(k)$) es procesada con el fin de facilitar la extracción de un conjunto de características (X). El preprocesamiento puede consistir en reducir la carga computacional por medio de la transformación de $d^*(k)$ en una señal intermedia, o en realizar una transformación de $d^*(k)$ al dominio de las frecuencias.

El objetivo de la etapa de extracción de características es encontrar una representación en la forma de un conjunto de medidas cuyos valores tengan, al mismo tiempo, una gran similitud para los elementos pertenecientes a una misma clase y una gran

diferencia respecto a los elementos de perteneciente a otras clases. Este concepto lleva a buscar un conjunto de características X que sea invariante a transformaciones irrelevantes de $d^*(k)$. Finalmente, la etapa de selección de características apunta a optimizar el conjunto de características X para abastecer al clasificador. Las características más discriminativas son retenidas en un nuevo conjunto $x \in X$, mientras que las restantes son descartadas. Eventualmente, se pueden proponer nuevas características y generarlas iterativamente con las etapas previas. Además, durante esta etapa se pueden obtener características de alto nivel por combinación y/o transformación de características de bajo nivel. Las características de alto nivel deben tener el mayor grado de discriminación posible para minimizar los errores de clasificación. En este punto del sistema, se ha completado la descripción de la señal, modificando su descripción original con el objetivo de mejorar su análisis. La entrada $d(t)$ ha sido transformada en un conjunto de características x que simplifica y mejora la tarea de clasificación. Esta transformación de $d(t)$ puede ser “continua” (generalmente basada en *frames*) o recaer en una etapa previa de segmentación.

Durante la tarea de clasificación, el sistema usa el conjunto de características x para evaluar el modelo de características construido durante la etapa de aprendizaje. El proceso de aprendizaje reconoce patrones y regularidades en las características y las clasifica en diferentes categorías. Como se ilustra en la Figura 2.1, el conjunto de características empleadas para construir el modelo son extraídas desde la base de datos durante el proceso de aprendizaje. Si la base de datos y el modelo son creados antes de que cada consulta ocurra, el proceso de aprendizaje es realizado de forma *offline*. En caso contrario, la base de datos y el modelo son actualizados en tiempo real y el proceso de aprendizaje es realizado *online*. Estas formas de aprendizaje tienen sus propias ventajas y desventajas, además de que su aplicabilidad depende del tipo de problema que se está intentando resolver: en el caso de que el conjunto de características cambie en el transcurso del tiempo, entonces la única opción es el aprendizaje *online*, el cual permite seguir tales cambios. En el caso de que el conjunto de características sea invariante en el tiempo, la mejor opción es el aprendizaje *offline*.

2.3. Descripción del sistema propuesto

Basado en un sistema de reconocimiento de patrones, el esquema del sistema propuesto se puede ver en la Figura 2.2. La señal de entrada al sistema es capturada por un micrófono electret unidireccional (transductor). La señal eléctrica analógica a la salida del micrófono debe ser acondicionada antes de ser digitalizada. Durante este acondicionamiento es posible que la señal sea amplificada o atenuada, en cuyo caso es importante conocer el nivel de ganancia aplicado. Una vez en formato digital,

la señal es preprocesada con el fin de mejorar su relación señal-ruido y utilizada en la siguiente etapa para calcular señales de interés, las cuales son submuestreadas para reducir el costo computacional de etapas posteriores. Como todo sistema de reconocimiento de patrones debe existir una etapa de almacenamiento o segmentación para identificar las unidades con las que se va a trabajar. Con la etapa de detección de eventos se logra identificar estas unidades, a las cuales se les extrae un conjunto de características que luego son pasadas a un clasificador de eventos. En la siguiente etapa se realiza un ajuste interno de parámetros asociado con la detección de eventos.

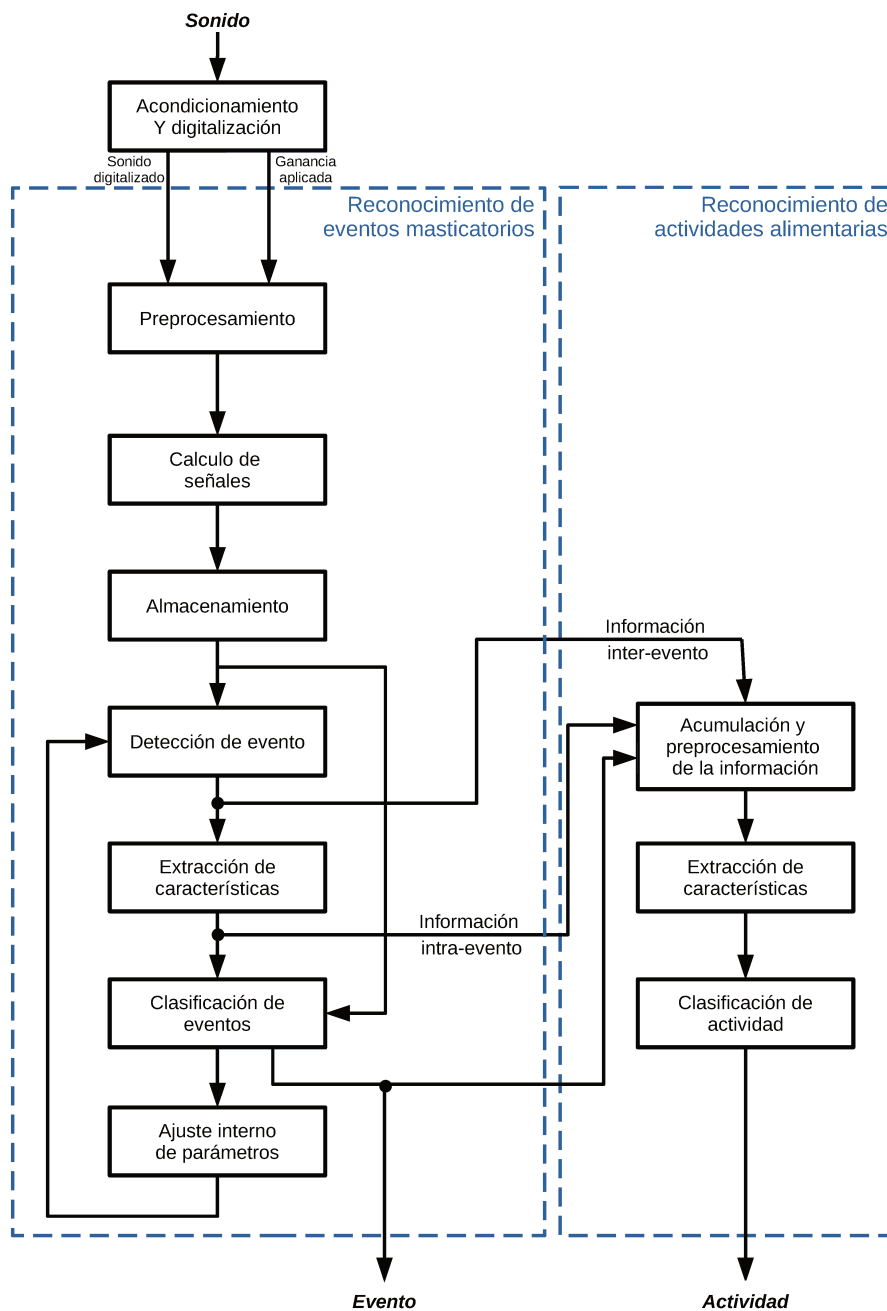


Figura 2.2: Diagrama de flujo del sistema abordado para monitoreo alimentario.

Una vez identificados los eventos masticatorios, es posible estimar los períodos de actividades de pastoreo y rumia, y sus parámetros indicativos (tiempo de rumia, tiempo de pastoreo, cantidad de bocados, tipo de bocado, energía de cada bocado, entre otros) utilizando información a nivel de eventos. Para ello, es necesario realizar la acumulación de información de un determinado bloque de eventos, preprocesar esa información y poder extraer características discriminativas que permitan diferenciar las actividades alimentarias de interés entre sí y de otras actividades a la vez. Una vez extraída esta información, la misma debería ser pasada a otro clasificador de manera similar a la etapa de reconocimiento de eventos masticatorios.

De esta forma el sistema completo contaría con dos niveles de clasificadores o métodos de aprendizaje maquina: (i) uno a nivel de eventos masticatorios, cuyas entradas serían las características temporales de los mismos, con el fin de identificar el tipo de evento; y (ii) uno a nivel de actividades alimentarias cuyas entradas serían las características intra-evento, inter-eventos y la secuencia de los mismos, con el fin de clasificar actividades de rumia y pastoreo, y así poder estimar sus duraciones.

2.4. Clasificadores basados en aprendizaje maquina

A continuación se describen brevemente las técnicas de aprendizaje maquina utilizadas por los clasificadores de los distintos algoritmos desarrollados. Dichos clasificadores fueron seleccionados debido a su buena capacidad de modelado y su bajo costo computacional. Esta última característica es imprescindible para lograr implementar los algoritmos en un sensor acústico, ya que se requiere de una ejecución en tiempo real, y además, de un bajo consumo energético debido a los períodos de operación necesarios.

Árbol de decisión

El árbol de decisión (DT, del inglés *Decision Tree*) ha probado ser exitoso en diferentes áreas y aplicaciones a pesar de su aparente simplicidad [78, 79]. Tienen la capacidad de aprender reglas simples de decisión y sistematizarlas con el objetivo de arribar a decisiones complejas. Para atributos numéricos, los árboles dividen el espacio de características en regiones rectangulares de ejes paralelos y etiquetan cada región con la clase correspondiente. Además, los DT proveen soluciones que son fáciles de implementar y entender.

Durante la fase de entrenamiento se evalúa una función de impureza (H), la cual define qué regla permanece en cada nodo. Por ejemplo, los árboles C4.5 usan la entropía cruzada como función de impureza:

$$H(X_m) = - \sum_k p_{mk} \log(p_{mk}), \quad (2.1)$$

donde p_{mk} es la fracción de ejemplos en el nodo m que corresponde a la clase k [80].

La complejidad computacional de clasificación solo depende de la profundidad del árbol, ya que ésta define el máximo número de comparaciones requeridas para alcanzar el nodo hoja.

Perceptrón multicapa

El perceptrón multicapa (MLP, del inglés *Multilayer Perceptron*) es un tipo de red neuronal artificial que puede operar con datos no separables linealmente [81]. Consiste en dos o más capas de nodos (perceptrones simples), donde cada capa está completamente conectada a la siguiente, pero no existen conexiones entre los nodos de una misma capa. La salida de un nodo (y) es la suma pesada (w_i) de las entradas (x_i) más un término de bias (θ), y está afectada por la función de activación (f) de la forma:

$$y = f \left(\sum_{i=1}^n w_i x_i + \theta \right). \quad (2.2)$$

Generalmente f es una función lineal para la capa de entrada, y no lineal (función sigmoidea o tangente hiperbólica) para las capas siguientes.

A diferencia de un árbol de decisión binario, una red neuronal tiene fronteras de decisión más suaves, llevando a mejores modelos de decisión. El costo computacional de la tarea de clasificación con una red neuronal entrenada es fijo y requiere de un número determinado de multiplicaciones y sumas que están establecidos por el número de capas, el número de neuronas en cada capa, el tamaño del patrón de entrada y el número de salidas.

2.5. Bases de datos

Para llevar a cabo los experimentos necesarios para la realización de esta tesis se utilizaron 2 bases de datos acústicas del comportamiento alimentario de ganado bovino, obtenidas por el grupo de trabajo en diferentes condiciones experimentales (dispositivos de grabación, pasturas y ambientes).

La primer base de datos se grabó en el *Kellogg Biological Station* (Universidad del estado de Michigan, Estados Unidos) en Agosto de 2014. En los experimentos se registró de manera diaria y continua el comportamiento alimentario de 5 vacas lecheras durante 6 días no consecutivos, consumiendo mezclas de trébol blanco y

raigrás, y trébol blanco y pasto ovillo. Para las grabaciones se utilizaron micrófonos electret convencionales y grabadores comerciales (SONY ICDPX312). Durante el ensayo, los 5 dispositivos de grabación fueron rotados en forma diaria entre las 5 vacas. Las señales fueron almacenadas en formato WAV utilizando una resolución de 16 bits y muestreada a 44.1 kHz para su posterior procesamiento.

Con esta base de datos se armó un dataset, llamado DS1, compuesto de 18 segmentos de pastoreo de 150 s de duración cada uno, y 18 segmentos de rumia de 75 s de duración cada uno, correspondiente a animales en condiciones de pastoreo libre. Este dataset está compuesto por segmentos libres de ruidos y es utilizado para estudiar el reconocimiento de eventos masticatorio en condiciones ideales.

Además, se formaron tres conjuntos de datos diferentes entre sí para estudiar el reconocimiento de las actividades alimentarias. Cabe remarcar que los datos de DS1 no pertenecen a ninguno de estos conjuntos. Estos conjuntos corresponden a animales en condición de pastoreo libre, y están compuestos por 24, 137 y 137 horas de señal acústica y son utilizados para entrenar los modelos y validar los modelos en ausencia de ruido y en presencia de ruido artificial, respectivamente.

La segunda base de datos fue obtenida en el tambo del campo experimental Villarino (Universidad Nacional de Rosario, Zavalla, Argentina) durante Octubre de 2014. Se grabaron los sonidos alimentarios producidos por 5 vacas lecheras Holstein durante 6 días no consecutivos, mientras estas pastaban una mezcla de alfalfa y festuca. Los dispositivos, micrófonos y procedimientos de grabación fueron los mismos utilizados para la obtención de la primera base de datos. Con esta base de datos se formó un dataset, llamado DS2, compuesto de 16 segmentos de pastoreo de 150 s de duración cada uno, y 16 segmentos de rumia de 75 s de duración cada uno, correspondiente a animales en condiciones de pastoreo libre. DS2 está compuesto por segmentos en los que la señal acústica está contaminada con ruidos externos típicos, y es utilizado para estudiar el reconocimiento de eventos masticatorio en condiciones adversas.

Todas las pruebas para el manejo de los animales fueron revisadas y aprobadas por los respectivos comités de ética y cuidado animal de cada universidad.

Las etiquetas de referencias a nivel de eventos masticatorios (*grazing-chew*, *ruminantion-chew*, *bite* y *chew-bite*) y a nivel de actividades alimentarias (*rumia*, *pastoreo* u *otras actividades*), contra las que comparar los resultados de los algoritmos propuestos en este trabajo, fueron generadas en forma auditiva por dos expertos en comportamiento alimentario de vacas con experiencia previa en esta tarea. Además para el etiquetado de las actividades alimentarias, los expertos contaban con información visual obtenida durante las jornadas diurnas del experimento.

2.6. Medidas de desempeño

Es necesario establecer ciertas medidas de desempeño tanto para el reconocimiento de eventos masticatorios como para el reconocimiento de actividades alimentarias, debido a que no hay medidas estandarizadas para esta aplicación. Algunos autores, como por ejemplo Navon et al. [54], han usado medidas o criterios propios para obtener sus resultados. Otros autores, como por ejemplo Zehner et al. [31], usan medidas de desempeño para el reconocimiento de actividades alimentarias que, si bien son muy utilizadas en la literatura (coeficientes de Pearson y de Spearman), no proveen una descripción completa del problema en cuestión.

2.6.1. Reconocimiento de eventos

Para evaluar el reconocimiento de eventos es muy importante la correcta sincronización entre la secuencia de eventos de referencia provista por los expertos y la secuencia de eventos reconocido por el algoritmo. En la Tabla 2.1 se puede ver que a partir de la correcta sincronización de ambas secuencias es posible establecer distintas métricas que permitan cuantificar el nivel de detección y clasificación de un reconocedor de eventos masticatorios. Analizando los eventos resaltados en negrita, podemos ver que la primera aparición (de izquierda a derecha) es una clasificación errónea (el algoritmo identifica un B cuando en realidad era un GC). La segunda aparición corresponde a un evento GC que no fue detectado. En tanto el tercer caso se debe a una falsa detección, y por consiguiente a una falsa clasificación como CB.

Referencia:	CB	GC	GC	B	GC		B	B
Reconocida:	CB	B		B	GC	CB	B	B

Tabla 2.1: Ejemplo de alineamiento de secuencias de referencia y reconocida.

La detección de eventos masticatorios solo considera la existencia o no de un evento, ignorando su correspondiente clase. La efectividad del detector está afectada por el número de eventos correctamente detectados (verdaderos positivos (VP)), el número de eventos no detectados (falsos negativos (FN)) y el número de detecciones que no corresponden a eventos (falsos positivos (FP)). Así, el desempeño del detector es cuantificado por el *precision*, el *recall* y el *F1-score*, definidos en la Tabla 2.2 [82].

La clasificación de eventos considera el tipo de eventos detectados, es decir, que el evento en la etiqueta de referencia y en la etiqueta reconocida sean iguales. El número de eventos correctamente detectados es utilizado para calcular, para cada clase i de eventos ($clase_i$), el número de eventos de la $clase_i$ correctamente clasificados (verdaderos positivos (vp_i)); el número de eventos de una clase distinta a la

Medida	Formula
<i>Precision</i>	$\frac{VP}{VP+FP}$
<i>Recall</i>	$\frac{VP}{VP+FN}$
<i>F1-score</i>	$\frac{2VP}{2VP+FP+FN}$

Tabla 2.2: Métricas utilizadas para evaluar la detección de eventos.

$clase_i$ correctamente clasificados (verdaderos negativos (vn_i)); el número de eventos clasificados como $clase_i$ que pertenecen a otra clase (falsos positivos (fp_i)); y el número de eventos reconocidos como una clase distinta a la $clase_i$ que pertenecen a la $clase_i$ (falsos negativos (fn_i)). El desempeño de clasificación obtenido para cada $clase_i$ es promediado aritméticamente y reportado por el $recall_M$, $precision_M$ y $F1 - score_M$ [82], y por los equivalentes promediado geométricamente $recall_G$, $precision_G$ y $F1 - score_G$ [83] (Tabla 2.3). La dispersión del desempeño de clasificación por clase tiende a afectar más negativamente los promedios geométricos que los promedios aritméticos, permitiendo observar si una o más clases son mejor clasificadas que el resto.

Medida	Formula
$Precision_M$	$\sum_{i=1}^l \frac{vp_i}{vp_i+fp_i}$
$Recall_M$	$\sum_{i=1}^l \frac{vp_i}{vp_i+fn_i}$
$F1 - score_M$	$\frac{2*Precision_M*Recall_M}{Precision_M+Recall_M}$
$Precision_G$	$\sqrt[l]{\prod_{i=1}^l \frac{vp_i}{vp_i+fp_i}}$
$Recall_G$	$\sqrt[l]{\prod_{i=1}^l \frac{vp_i}{vp_i+fn_i}}$
$F1 - score_G$	$\frac{2*Precision_G*Recall_G}{Precision_G+Recall_G}$

Tabla 2.3: Métricas utilizadas para evaluar la clasificación de eventos.

2.6.2. Reconocimiento de actividades alimentarias

Ha existido mucha controversia respecto a la forma de evaluación de métricas en el reconocimiento de actividades continuas en humanos [84]. Algunos autores sostienen que cierta información acerca de las características típicas de las actividades es ignorada al hacer que de alguna manera los resultados se adecuen a métricas estándares medidas en escala de corta duración (*frames*) o de larga duración (bloques).

Por ejemplo, las métricas estándares no revelan si una actividad ha sido fragmentada en varias actividades más pequeñas, o si varias actividades han sido unidas para formar una actividad de duración más larga, o si existen desfases temporales entre la secuencia de referencia y la reconocida. En este sentido, Vanrell et al. [68] realizaron un análisis para cuantificar la actividad alimentaria de rumiantes utilizando múltiples métricas a nivel de *frames* y bloque, permitiendo obtener resultados a escala temporal corta y larga, respectivamente, y representándolos en gráficos de araña. Sin embargo, en este análisis se asume que en cada señal evaluada la proporción y la duración de cada actividad son iguales. En esta tesis se tiene en cuenta la duración de cada actividad.

Así, tanto las etiquetas de referencia, como así también las etiquetas reconocidas están conformadas por una lista de bloques donde se especifican el inicio, el final y el tipo de actividad para cada señal de audio. Estos bloques son divididos en secuencias de *frames* no superpuestos de 1 s, permitiendo obtener una resolución muy detallado de las actividades. A partir de esta transformación de bloques a *frames*, se puede evaluar el reconocimiento utilizando métricas convencionales. El *Average Accuracy* es utilizado para cuantificar el nivel medio del reconocimiento de las actividades de interés (rumia y pastoreo). Además, la matriz de confusión permite observar el nivel de reconocimiento y confusión entre las distintas actividades.

De esta forma, al incorporar la duración de cada actividad y utilizando *frames* con 1 s de resolución, se pueden obtener muy buenos resultados en la estimación del tiempo de cada actividad.

3 Reconocimiento de eventos masticatorios

Uno de los objetivos de la ganadería de precisión es el correcto reconocimiento de eventos masticatorios de animales rumiantes, a partir de una correcta detección y clasificación de los mismos. Con esto se obtiene información útil para determinar actividades alimentarias de largo plazo (rumia y pastoreo) [68], estimar la cantidad de materia ingerida por el animal [5, 6], y determinar posibles problemas de salud del mismo [7-9], entre otros. El problema del reconocimiento de eventos masticatorios es bien conocido y ha sido abordado por diversos autores utilizando mayormente señales acústicas.

Milone et al. [52] propusieron un algoritmo basado en el análisis de señales acústicas capaz de detectar y clasificar los eventos masticatorios, logrando una tasa de reconocimiento del 85 % bajo condiciones controladas con bajo niveles de ruido. Como contrapartida, este algoritmo tiene un costo computacional muy elevado. En este sentido, Chelotti et al. [55] desarrollaron un algoritmo de bajo costo computacional mediante el análisis en el dominio temporal de las señales acústicas utilizando técnicas de reconocimiento de patrones logrando un desempeño similar a [52]. La detección de posibles eventos se realiza por medio de umbrales adaptativos. Una etapa de segmentación identifica los eventos candidatos sobre los cuales se extraen características, para finalmente ser clasificados siguiendo reglas heurísticas. Los inconvenientes que presenta este algoritmo son que los parámetros deben ser ajustados manualmente, y que su desempeño se ve muy degradado ante cualquier tipo de ruido presente en las señales acústicas. La solución a estos problemas fue propuesta por Chelotti et al. [57] quienes incorporaron al algoritmo una etapa de preprocesamiento para remover tendencias y ruidos usando técnicas de procesamiento de señales. Además, se mejoró el clasificador de eventos usando técnicas de aprendizaje maquina. Este algoritmo llamado *Chew Bite Intelligent Algorithm* (CBIA) tiene un bajo costo computacional y logra tasas de detección y clasificación cercana al 98 % y al 90 %, respectivamente, en condiciones controladas con bajo ruido. Sin embargo, su desempeño disminuye al ser utilizado en entornos agrícolas y ganaderos a cielo abierto, donde existen distintos y variados tipos de ruidos. Dichos ruidos no pueden ser completamente removidos por la etapa de preprocesamiento y afectan tanto la tarea de detección como la extracción de características y, consecuentemente, la clasificación de eventos.

Como solución a estos problemas, esta tesis propone un algoritmo robusto denominado *Chew-Bite Energy Based Algorithm* (CBEBA), el cual mejora la detección de eventos mediante el uso de la envolvente de la potencia instantánea de la señal

de entrada, y además, mejora la clasificación de eventos mediante la incorporación de un nuevo conjunto de características con gran inmunidad a condiciones adversas.

Por otro lado, los algoritmos previamente mencionados [52, 55, 57] son capaces de clasificar solo tres tipos de eventos masticatorios (*chew*, *bite* y *chew-bite*). Sin embargo, ninguno de ellos es capaz de distinguir si los eventos *chew* corresponden a una actividad alimentaria de rumia o pastoreo. Una propiedad distintiva de CBEBA radica en que es el primer algoritmo que logra realizar dicha distinción en animales bovinos, dando lugar a una clasificación en termino de cuatro tipo de eventos masticatorios (*rumination-chew*, *grazing-chew*, *bite* y *chew-bite*).

A continuación se mencionará brevemente el conjunto de características seleccionadas y se comparará su robustez frente a las utilizadas por el CBIA. También se describirá brevemente las etapas del CBEBA y se evaluará su tasa de reconocimiento en distintas bases de datos. Para mayor detalle se recomienda dirigirse al Anexo A del presente documento, donde se encuentra el artículo que presenta el CBEBA.

3.1. Análisis de características robustas

El novedoso conjunto de características temporales son extraídas a partir de la potencia instantánea de la señal de entrada de sonido procesada o de su envolvente (Figura 3.1, fila 2):

- **Signo de la pendiente de la envolvente:** representa la forma de un evento y es calculada como la cantidad de cruces por cero del valor de la pendiente de la envolvente, es decir, cuenta la cantidad de picos de la envolvente (Figura 3.1, fila 2). Esta característica permite diferenciar eventos *chew-bite*, que presentan dos o más cruces por cero, de los otros eventos presentes en el pastoreo (*grazing-chew* y *bite*) los cuales presentan un solo cruce por cero (Figura 3.1, fila 3).
- **Rapidez acumulada de la envolvente:** esta características está relacionada con la intensidad del sonido y sus variaciones (Figura 3.1, fila 1). Su valor se calcula como la suma acumulada del valor absoluto de la pendiente de la envolvente. Permite distinguir eventos de baja intensidad (*rumination-chew*) de cualquier otro tipo de evento de pastoreo. Además, diferencia entre los *grazing-chew* de los *bite* y *chew-bite*, los cuales presentan valores similares entre sí. La Figura 3.1, fila 4, muestra el incremento de rapidez acumulada en función del tiempo.
- **Duración:** es el tiempo que en que un evento tiene lugar. Un evento compuesto *chew-bite* tiene una duración mayor que otros eventos de pastoreo. Además, la duración de un *rumination-chew* es similar o mayor que un *grazing-chew*

debido al contenido de humedad del material ingerido por el animal (Figura 3.1, fila 5) [21].

- **Simetría:** está relacionado a la forma del evento y se mide como la simetría de la longitud temporal de un evento calculada a partir del pico máximo de la envolvente. Esta características permite diferenciar *bite* de *chew-bite*, al tener un mayor valor en los *bite* (Figura 3.1, fila 5).
- **Energía total:** es la suma acumulada de la señal de potencia instantánea y está relacionado con la intensidad y duración de un evento. En general, su valor es mayor en los eventos compuestos *chew-bite* que en los *bite*, y en los *bite* es mayor que en los *grazing-chew* y *ruminantion-chew*. De hecho, un *ruminantion-chew* tiene una energía mucho menor que cualquier evento de pastoreo debido a que el material alimentario ya ha sido masticada y triturada parcialmente. La Figura 3.1, fila 6, muestra el incremento de energía acumulada en función del tiempo.

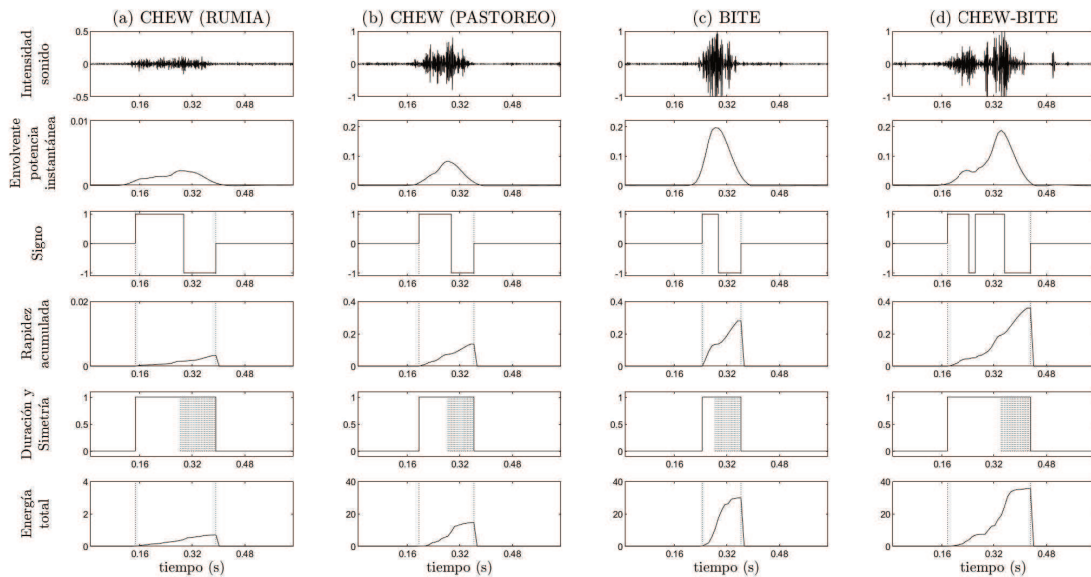


Figura 3.1: Señales de eventos masticatorios producidos por movimientos mandibulares y sus respectivas características.

Para poder determinar el grado de discriminación entre las distintas clases de eventos masticatorios, a partir del conjunto de características propuesto, se llevó a cabo un análisis cualitativo del método de reducción de dimensionalidad t-SNE [85]. Este método es útil para visualizar datos de alta dimensionalidad en dos o tres dimensiones, ya que es capaz de captar la estructura local de los datos sin perder de vista su estructura global.

La Figura 3.2 muestra la distribución de los eventos masticatorios luego de aplicar dicho método. Las figuras de la izquierda corresponden al nuevo conjunto de

características (Figura 3.2-a) y al conjunto de características utilizado previamente en CBIA (Figura 3.2-c) al ser extraídas de señales adquiridas con muy bajo niveles de ruido ambiente. Se observa que los grupos de datos correspondiente al nuevo conjunto de datos se encuentran menos solapados entre sí que los correspondientes al CBIA, dándole soporte a la idea de que el nuevo conjunto de características permite un mejor discernimiento entre las distintas clases de eventos masticatorios. En forma similar, las Figuras 3.2-c y 3.2-d corresponden al nuevo conjunto y al conjunto del CBIA al ser identificadas en señales que presentan una condición de ruido ambiente considerable. La forma y la separación de los grupos para el nuevo conjunto de datos es similar al caso libre de ruido con un mayor solapamiento entre ellos, mientras que los grupos son menos densos y están muy solapados entre sí para el conjunto de datos correspondiente al CBIA. Esto muestra que el nuevo conjunto de características propuestas presentan una gran robustez a las posibles variabilidades de la señal de audio adquirida.

Cabe remarcar que en todos los casos anteriores es posible identificar dos grupos bien identificados de *chew*, independientemente de si fueron explícitamente separados en *grazing-chew* y *ruminantion-chew* (nuevo conjunto de datos) o no (conjunto de datos del CBIA), lo cual refuerza la hipótesis de que es posible diferenciar entre distintos tipos de evento *chew*.

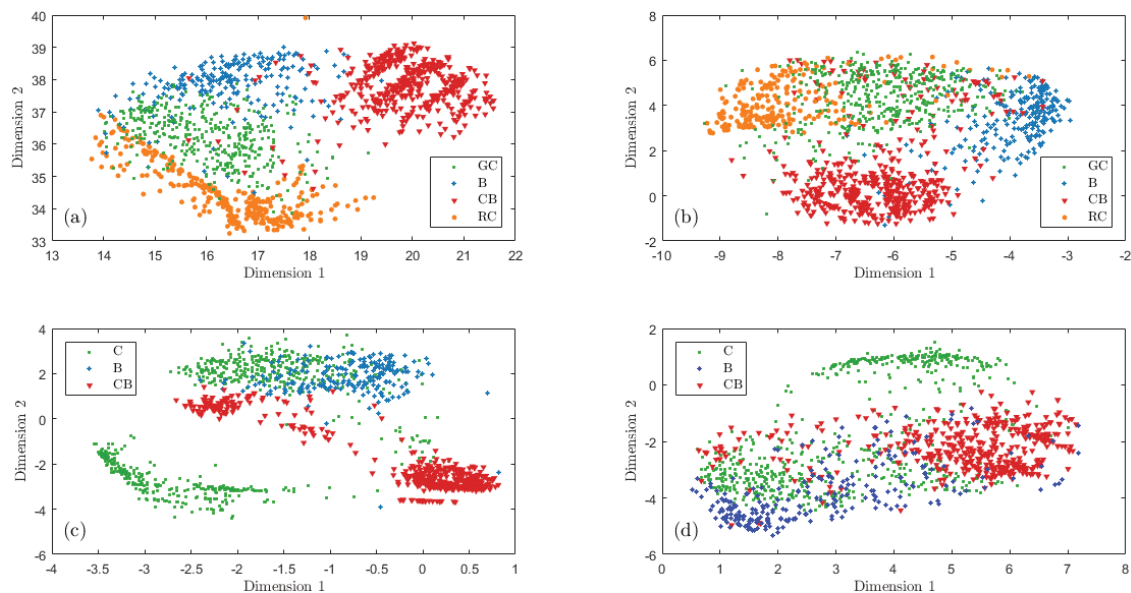


Figura 3.2: Análisis t-SNE del nuevo conjunto de características propuesto (a-b) y del conjunto de características usado en el CBIA (c-d) en términos de clases (GC: *grazing-chew* - RC: *ruminantion-chew* - C: *chew* - B: *bite* - CB: *chew-bite*).

3.2. Descripción del algoritmo

El algoritmo para clasificar eventos masticatorios en tiempo real CBEBA presenta el enfoque clásico de un sistema de reconocimiento de patrones. El mismo se divide en 6 etapas secuenciales con realimentación interna según se muestra en la Figura 3.3.

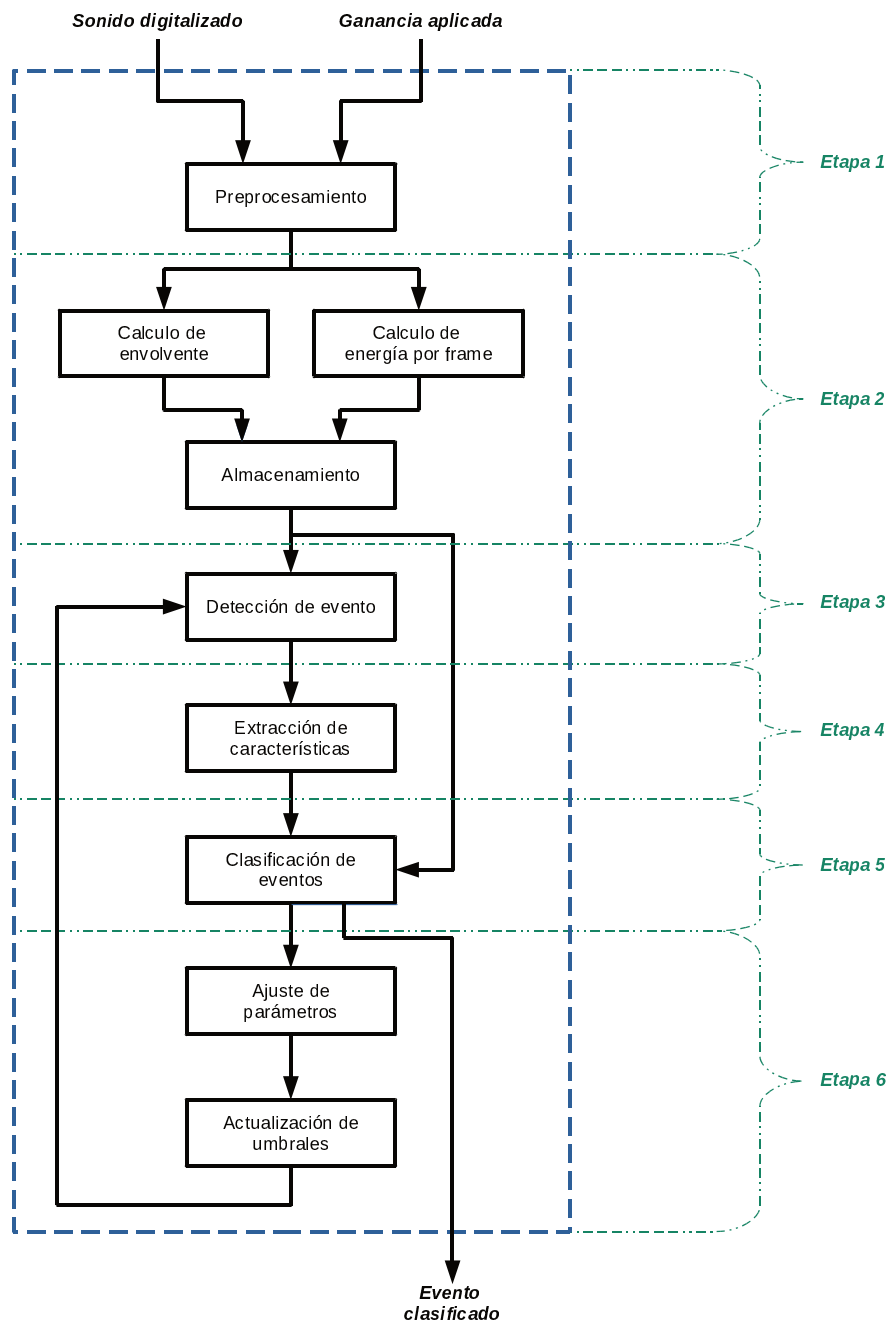


Figura 3.3: Diagrama en bloque de las etapas del CBEBA.

Preprocesamiento

La señal de entrada de audio digital puede provenir de un conversor analógico/digital (A/D) o de una memoria, dependiendo de la implementación del algoritmo. Esta etapa se encarga de:

- Normalizar: en caso de contar con información respecto a la ganancia aplicada por el sistema adquirente, esta puede ser utilizada para normalizar la señal de audio.
- Limitar el ancho de banda: se ajusta a la banda de frecuencia donde la densidad espectral de energía de los eventos masticatorios se encuentra localizado.
- Calcular la potencia instantánea de la señal preprocesada.

Calculo de señales y almacenamiento

La señal de potencia instantánea se usa para calcular una señal envolvente y una señal de energía por tramos. Ambas señales son submuestreadas y almacenadas. El submuestreo permite reducir la frecuencia de trabajo de las etapas posteriores del algoritmo, sin perder información relevante. A medida que se almacenan nuevas muestras, se descartan las muestras más antiguas, contando siempre con un segundo de información almacenada.

Detección

La posición media de la señal envolvente almacenada es utilizada para buscar un posible evento masticatorio. Para ello se buscan picos en la envolvente que sean superior a un umbral adaptativo. Cuando se produce una detección se delimitan los límites de dicho evento candidato al comparar la señal de energía por tramos almacenada con otro umbral adaptativo.

Extracción de características

Los límites del evento candidato en conjunto con las dos señales almacenadas son utilizados para extraer el conjunto de características mencionadas en la sección 3.1.

Clasificación

Las características extraídas son utilizadas para decidir si el evento candidato debe ser clasificado o descartado. En caso afirmativo, un clasificador de bajo costo computacional realiza dicha tarea. Los clasificadores evaluados utilizando técnicas

de aprendizaje maquina son los mencionados en el capítulo 2.4: DT y MLP. Los principales hiper-parámetros de cada clasificador fueron optimizados utilizando el método de búsqueda por grilla.

Ajuste de parámetros

Las señales almacenadas son utilizadas para actualizar los valores de ambos umbrales. Para ello se utilizan valores extraídos del propio evento y valores inter-evento. Esto permite adaptar los umbrales al rango de los eventos teniendo en cuenta el nivel de ruido de base de las señales, el cual puede ser variable en el tiempo. Un análisis detallado de cómo se calculan los umbrales se encuentra disponible en el material suplementario del artículo del CBEBA (Anexo A).

3.3. Resultados y discusión

El CBEBA fue evaluado en los datasets sin ruido (DS1) y con ruido (DS2) descritas en el capítulo 2.5, en distintas combinaciones: *i*) entrenamiento con DS1 y testeo con DS1 (denominado DS11); *ii*) entrenamiento con DS1 y testeo con DS2 (denominado DS12); *iii*) entrenamiento con DS2 y testeo con DS1 (denominado DS21); *iv*) entrenamiento con DS2 y testeo con DS2 (denominado DS22). Las tasas de reconocimientos obtenidas en cada caso se muestran en las matrices de confusión en la Figura 3.4.

Los resultados muestran que los *ruminacion-chew* tienen un muy alto nivel de reconocimiento permitiendo diferenciarlos de todos los otros eventos asociados al pastoreo, demostrando la hipótesis planteada. Además, la tasa de reconocimiento promedio es superior al 87.4% en todas las combinaciones.

Para comparar los resultados obtenidos por CBEBA contra el algoritmo previo CBIA, ambos algoritmos fueron configurados, evaluados y testados en las mismas condiciones, utilizando en ambos algoritmos un clasificador MLP. Además, en CBEBA se combinaron los eventos *ruminacion-chew* y *grazing-chew* en un simple evento *chew*, debido a que no es posible realizar una comparación directa entre ambos algoritmos, debido a que cada uno clasifica distintas clases de eventos masticatorios. La tabla 3.1 muestra los resultados obtenidos en cada caso. Se comprobó si la discrepancia de los resultados obtenidos eran estadísticamente significativa ($p < 0,05$) usando la prueba de los rangos con signo de Wilcoxon [86]. Los valores con fondo verde dentro de una misma fila y para una combinación de dataset determinada difieren significativamente, mientras que los valores con fondo rosa, en iguales condiciones, no muestran diferencias significativas. Se observa que al evaluar los algoritmos en presencia de ruido (DS12 y DS22), la detección de evento de CBEBA logra mejores

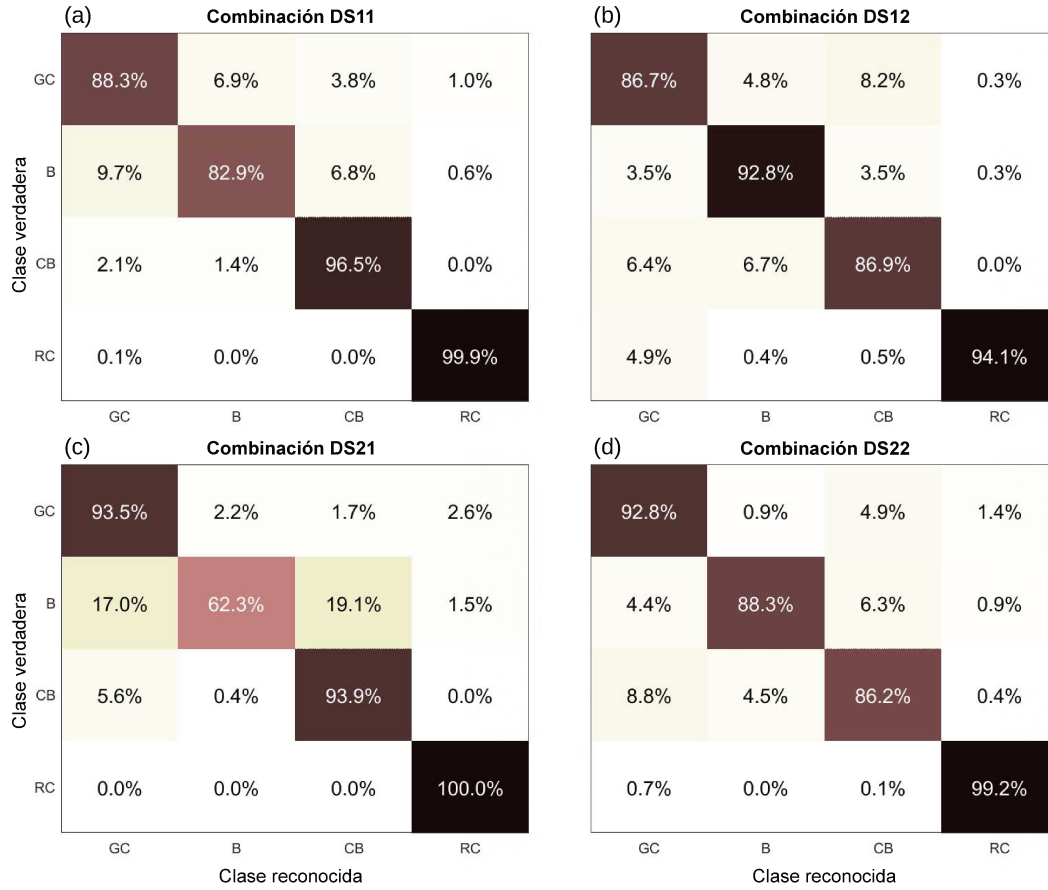


Figura 3.4: Matrices de confusiones de eventos de CBEBA: *grazing-chew* (GC), *bite* (B), *chew-bite* (CB) y *rumination-chew* (RC) en las cuatros combinaciones de datasets.

resultados que el utilizado en CBIA, gracias a una menor cantidad de falsos eventos detectados (falsos positivos), lo cual se cuantifica con el *Precision*. Esto indica que CBEBA detecta eventos de forma superior que CBIA en ambientes ruidosos.

	DS11		DS12		DS21		DS22		
	CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA	
Detección	<i>Recall</i>	98.8 ± 0.7	98.4 ± 1.0	98.5 ± 0.7	98.0 ± 0.9	98.0 ± 0.9	97.9 ± 1.3	98.7 ± 0.6	98.4 ± 1.2
	<i>Precision</i>	98.8 ± 0.8	99.2 ± 0.5	91.7 ± 2.8	95.8 ± 2.1	98.0 ± 0.7	98.7 ± 0.8	91.9 ± 2.6	96.1 ± 2.6
	<i>F1 - score</i>	98.8 ± 0.5	98.8 ± 0.6	95.0 ± 1.7	96.9 ± 1.4	97.8 ± 0.5	98.3 ± 0.9	95.1 ± 1.5	97.2 ± 1.8
Clasificación	<i>Recall_M</i>	91.7 ± 3.3	91.8 ± 3.4	72.9 ± 5.0	90.5 ± 4.6	82.0 ± 6.8	92.9 ± 2.6	86.7 ± 6.2	90.0 ± 4.2
	<i>Precision_M</i>	93.7 ± 2.4	92.1 ± 3.5	79.6 ± 6.7	85.9 ± 5.7	83.0 ± 4.9	85.1 ± 4.8	86.8 ± 5.5	90.2 ± 5.4
	<i>F1 - score_M</i>	92.5 ± 1.8	91.9 ± 2.4	77.5 ± 5.2	88.0 ± 4.4	82.9 ± 5.1	88.8 ± 3.1	87.3 ± 5.0	90.0 ± 4.3
	<i>Recall_G</i>	90.8 ± 3.9	91.5 ± 3.8	68.0 ± 7.6	90.2 ± 5.1	80.2 ± 7.8	92.7 ± 2.8	85.3 ± 8.2	89.5 ± 4.7
	<i>Precision_G</i>	93.5 ± 2.8	91.6 ± 4.0	78.5 ± 7.6	84.3 ± 7.3	81.6 ± 5.8	83.3 ± 5.7	85.7 ± 7.3	89.6 ± 6.0

Tabla 3.1: Comparación de resultados (media ± desvío estándar; %) entre CBIA y CBEBA en las cuatro combinaciones de datasets.

Al evaluar las tasas de reconocimiento, ambos algoritmos logran tasas similares de clasificación en ausencia de ruido (DS11), pero CBEBA logra mejores resultados que CBIA en presencia de ruido, tanto para entrenar, testear o ambas. Observando la disminución de las tasas de clasificación al comparar los resultados al testear con ruido

y sin ruido (DS12 contra DS11), se observa que la disminución en CBEBA es menos de la mitad producida en CBIA. Para los casos de condiciones de entrenamiento con ruido (DS21 y DS22), CBEBA logró mejores métricas generales de clasificación que CBIA, respectivamente, lo que indica una mayor sensibilidad de CBIA al ruido ambiental. Como contrapartida el costo computacional de CBEBA (43178 operaciones/s) aumenta en 11 % respecto al costo de CBIA (38960 operaciones/s). Sin embargo, este incremento no es significativo y aún permite su implementación en un dispositivo electrónico embebido.

Se comprobó en CBEBA que modificar la frecuencia de muestreo del audio de entrada (cumpliendo el teorema de muestreo de Nyquist-Shannon) o aumentar la longitud de palabra del audio de entrada digitalizado por encima de 8 bits tiene una variación insignificante en todas las métricas ($<0,5\%$), lo cual está en concordancia con [55].

El artículo donde se propone y evalúa el CBEBA se encuentra en el Anexo A del presente documento.

4 Reconocimiento de actividades alimentarias

Los animales rumiantes realizan distintas actividades de forma habitual, en las que puede o no existir la interacción de personas. Entre las actividades más comunes se destacan: rumiar, pastorear, hidratarse, descansar, desplazarse, ser ordeñados, ser pesados y aquellas relacionadas con cuestiones reproductivas y de salud, entre otras. Se pretende detectar, clasificar y estimar los tiempo de las dos actividades alimentarias: *rumia* y *pastoreo*. En este sentido, toda actividad que no corresponda a rumia o pastoreo, será identificada como *otro*, sin discriminar la actividad de la que se trate.

Como se ha mencionado anteriormente, las actividades alimentarias de largo plazo (rumia y pastoreo) se pueden pensar como una secuencia de eventos masticatorios de corto plazo (*ruminacion-chew* y *grazing-chew*, *bite* y *chew-bite*). En el caso de la rumia dicha secuencia consiste en una serie de masticaciones, asociadas a eventos *ruminacion-chew*, que cada cierto intervalo de tiempo (40-60 s) son separadas por un breve intervalo (3-7 s) de deglución-regurgitación del alimento. En el caso del pastoreo, la secuencia está compuesta por una serie de eventos masticatorios compuesto por *grazing-chew*, *bite* y *chew-bite*, sin poseer un determinado patrón u orden aparente. En este caso, la duración de la secuencia de eventos no tiene una duración definida en el tiempo (Figura 4.1).

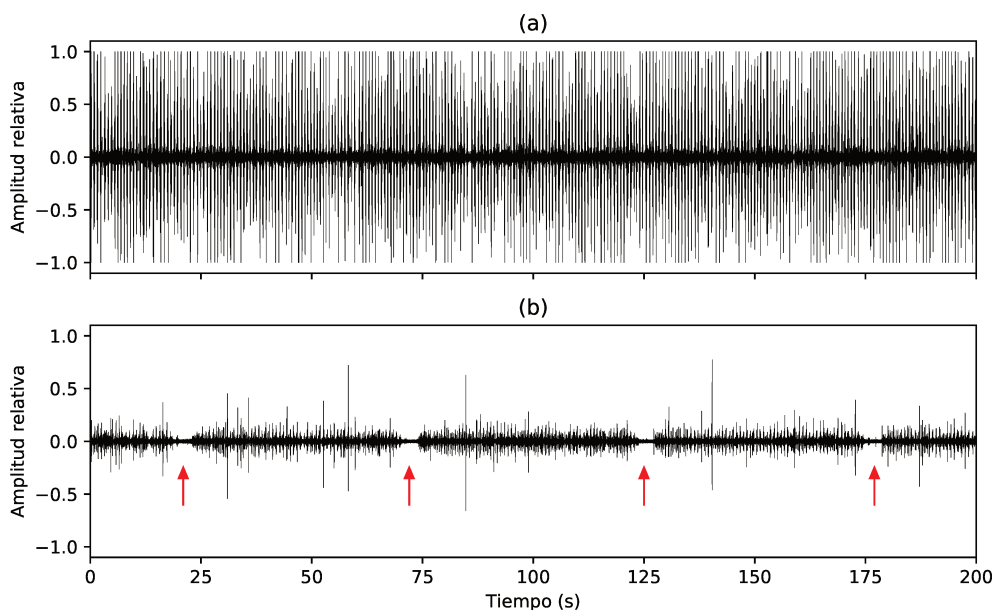


Figura 4.1: Actividades de pastoreo (a) y rumia (b). Las flechas en la rumia indican las pausas de deglución-regurgitación del alimento.

El reconocimiento de actividades alimentarias ha sido abordado previamente utilizando distintos tipos de transductores para adquirir las señales a procesar. La mayoría de los métodos desarrollados en el estado del arte utilizan acelerómetros, ya que además de reconocer actividades alimentarias también reconocen actividades de posición o movimiento de los rumiantes [39, 58, 59, 61, 62]. Todos estos algoritmos fueron desarrollados en ambientes de galpón o establo, donde los animales no tienen gran superficie para desplazarse. Sin embargo, en condiciones de pastoreo libre existe un mayor riesgo de que el collar, donde se ubican los acelerómetros, se muevan generando posiblemente datos inestables o incorrectos. Sin embargo, ninguno de estos algoritmos anteriores [39, 58, 59, 61, 62] evalúa su afectación ante movimientos en la posición de los acelerómetros. En este sentido, deberían realizarse experimentos de mayor duración para ver este efecto.

Un algoritmo capaz de reconocer la rumia y el pastoreo en condiciones de pastoreo libre fue propuesto por Vanrell et al. [68]. Este método *offline* está basado en la autocorrelación de la envolvente del sonido y obtuvo resultados superiores al equipo comercial Hi-Tag. Sin embargo, su principal desventaja es que se debe contar con toda la señal almacenada para poder ejecutarlo. Esto lo vuelve un algoritmo incapaz de reconocer actividades en tiempo real, aunque se lo puede utilizar como estándar de comparación. Como solución a esta limitante, esta tesis propone una serie de algoritmos *online* de bajo costo computacional capaces de ser implementado en dispositivos embebidos.

4.1. Métodos propuestos

La arquitectura general del sistema reconocedor de actividades alimentarias propuesto se muestra en la Figura 4.2, el cual se encuentra internamente dividido en dos etapas: reconocimiento de eventos masticatorios y reconocimiento de actividades alimentarias. El reconocedor de eventos masticatorios utiliza uno de los métodos introducidos en el capítulo 3 (CBIA [57] y CBEBA (Anexo A)) e internamente está separado en cuatro bloques: preprocesamiento del sonido, detección, extracción de características y clasificación del evento masticatorio. Parte de la información de salida de cada uno de estos bloques es transmitida a la etapa del reconocedor de actividades alimentarias (flechas anaranjadas en Figura 4.2). Esta etapa almacena internamente esta información para su posterior análisis en segmentos de longitud fija. Así, cada un determinado periodo de tiempo, a partir de los datos almacenados, se calcula un conjunto de características que es utilizado por un clasificador de actividades para determinar si el segmento considerado corresponde a *rumia*, *pastoreo* u *otro*. El conjunto de características de actividad está conformado por características dependiente e independiente de la clase de eventos reconocidos.

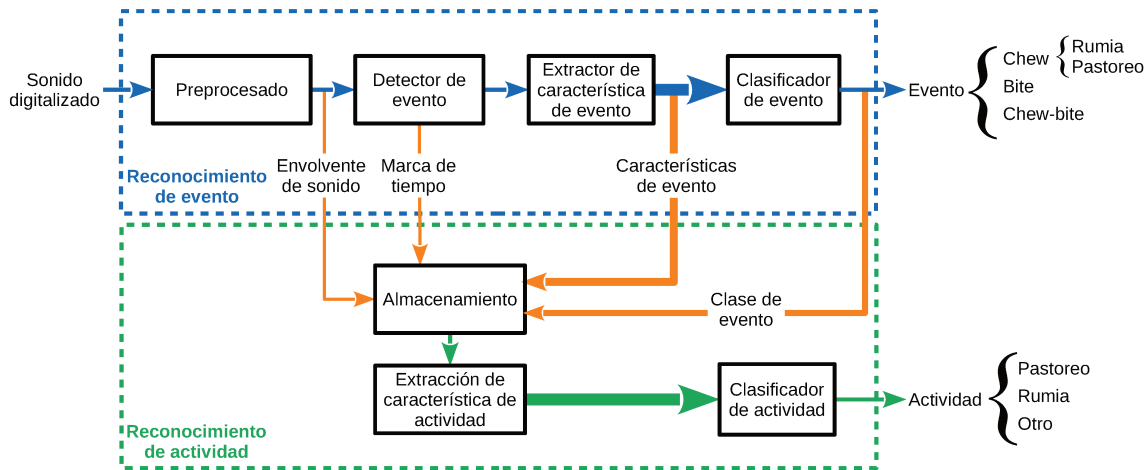


Figura 4.2: Diagrama general del reconocedor de actividades.

Métodos dependientes del reconocimiento de eventos

Uno de los algoritmos propuesto para estimar los tiempo de rumia y pastoreo se llama *Bottom-Up Foraging Activity Recognizer* (BUFAR). La etapa de reconocimiento de eventos masticatorios es realizada por el método CBIA [57], denominándolo BUFAR-CBIA.

El diagrama de bloques del BUFAR se muestra en la Figura 4.3. El bloque de almacenamiento guarda la marca de tiempo junto con la información de los eventos masticatorios reconocidos (*chew*, *bite* y *chew-bite*). El conjunto propuesto de cuatro características a nivel de actividad son estadísticos del segmento considerado: la tasa de eventos masticatorios (f_1)¹, la proporción de eventos *chew* (f_2), la proporción de eventos *bite* (f_3) y la proporción de eventos *chew-bite* (f_4). El clasificador de actividades de bajo costo computacional utilizado puede ser un árbol de decisión (DT) o una red neuronal perceptrón multicapa (MLP), que clasifica los segmentos como *rumia*, *pastoreo* u *otro*. Por último, un proceso de suavizado es aplicado sobre la secuencia de segmentos etiquetados para eliminar breves diferencias y así reducir la fragmentación de los periodos de actividad alimentaria.

Para mayor detalle sobre la descripción del algoritmo se recomienda dirigirse al Anexo B del presente documento, donde se encuentra el artículo que presenta el BUFAR.

La principal ventaja que presenta este algoritmo es que solo ocupa cuatro características a nivel de actividad, y el bajo costo computacional asociado al cálculo de cada una de ellas. Sin embargo, el cálculo de las proporciones de cada tipo de evento (características f_2 a f_4) se encuentra afectado por el error cometido por el clasificador de eventos masticatorios del CBIA.

¹El número entre paréntesis indica el número ordinal de la característica dentro del conjunto de características de actividad

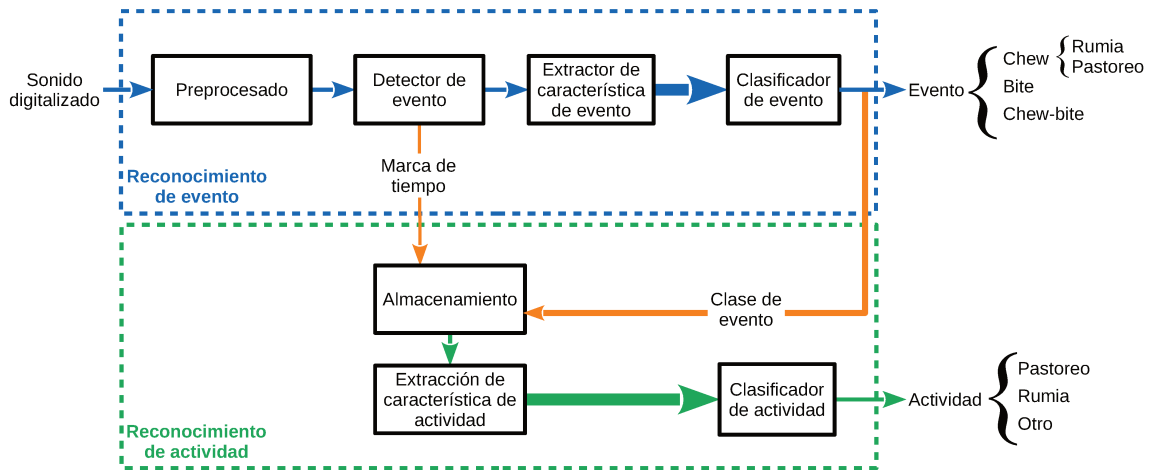


Figura 4.3: Diagrama general del BUFAR.

Una forma de disminuir el error inherente en la clasificación de eventos, es utilizando un reconocedor de eventos masticatorios que cometa menores errores de clasificación. A su vez, si el reconocedor de eventos masticatorios es capaz de distinguir eventos producidos únicamente durante la rumia y únicamente durante el pastoreo, entonces se disminuye el efecto del error cometido por el clasificador de eventos, ya que ahora solo interesa el error de clasificación de eventos asociados a rumia con los asociados a pastoreo, y viceversa.

Estas dos soluciones propuestas son incorporadas como una variante del BUFAR manteniendo el mismo principio de funcionamiento y la estructura general del sistema (ver Figura 4.3), utilizando el CBEBA como reconocedor de eventos masticatorios en lugar del CBIA, denominándolo BUFAR-CBEBA. Como consecuencia, el nuevo conjunto de características de actividad propuesto son: la tasa de eventos masticatorios (f_1), la proporción de eventos *bite* (f_3), la proporción de eventos *chew-bite* (f_4), la proporción de eventos *grazing-chew* (f_5) y la proporción de eventos *ruminantion-chew* (f_6).

Métodos independientes del reconocimiento de eventos

El algoritmo *Jaw Movement segment-based Foraging Activity Recognizer* (JMFAR) propone una forma diferente para estimar los tiempos de rumia y pastoreo que ambas variantes de BUFAR, utilizando un conjunto alternativo de características a nivel de actividad, las cuales no contienen información del evento masticatorio reconocido. Es decir, en este caso, solo es necesario detectar el evento masticatorio, pero no su correspondiente clasificación. Con esto se logra eliminar el error inherente de clasificación de los eventos masticatorios e independizar la clasificación de actividades alimentarias del reconocimiento de eventos masticatorios.

El diagrama general de bloques se muestra en la Figura 4.4. El JMFAR utiliza el

reconocedor de eventos masticatorios CBIA. El bloque de almacenamiento guarda la información de la envolvente del sonido, la marca de tiempo asociada a la detección de un evento y el conjunto de características extraídas por el CBIA. Esta información es utilizada en segmentos de longitud fija para calcular un conjunto de características a nivel de actividad asociadas a información temporal, estadística y espectral de los eventos masticatorios detectados.

En el artículo científico donde se presenta el JMFAR localizado en el Anexo C del presente documento, se encuentra un análisis heurístico, una descripción detallada y el cálculo de cada una de las características que conforman el conjunto, mencionadas a continuación:

- Tasa de eventos masticatorios (f1).
- Estadísticos de la amplitud: la amplitud máxima de cada evento detectado es utilizada para calcular su media (f7), su desviación estándar (f8), su asimetría estadística (f9) y su curtosis (f10).
- Estadísticos de la duración: la duración de cada evento detectado es utilizada para calcular su media (f11), su desviación estándar (f12), su asimetría estadística (f13) y su curtosis (f14).
- Espectro de la envolvente de sonido: esta característica representa la energía en la banda de frecuencia de 1.0-1.5 Hz (f15) relativa a la energía total del segmento.
- Variantes del jitter: el jitter se define como la variación del periodo entre dos eventos detectados consecutivos. Las características propuestas son: el jitter absoluto (f16), el jitter relativo (f17), el jitter PPQ5 (del inglés *5-Point Period Perturbation Quotient*) (f18) y la desviación estándar del jitter (f19).

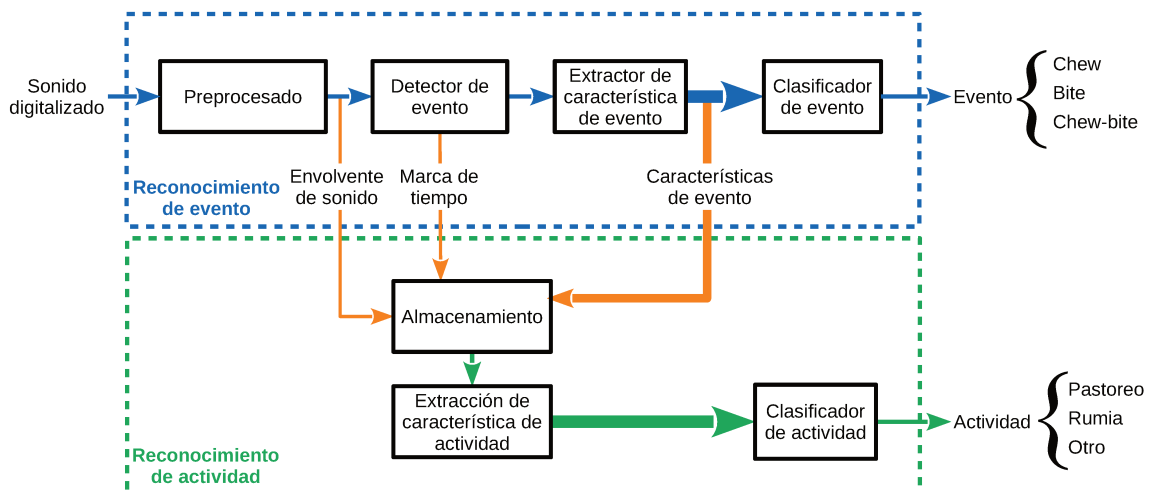


Figura 4.4: Diagrama general del JMFAR.

- Variantes del shimmer: el shimmer se define como la variación de la amplitud entre dos eventos detectados consecutivos. Las características propuestas son: el shimmer absoluto (f20), el shimmer relativo (f21), el shimmer APQ5 (del inglés *5-Point Amplitud Perturbation Quotient*) (f22) y la desviación estándar del shimmer (f23).
- Tacograma temporal: el tacograma considera los patrones de la regularidad de los eventos masticatorios. Esta característica indica la frecuencia con que se encuentra variabilidad en el periodo de eventos masticatorios consecutivos en el rango 3-10 s (f24).
- Tacograma espectral: estas características representan la energía relativa en la banda 0.017-0.020 Hz (f25) y en la banda 0.00-0.02 Hz (f26), respecto a la energía total del tacograma en el segmento considerado y provee información sobre las pausas de deglución y regurgitación producidas durante la rumia.

En la Tabla 4.1 se marcan las características de actividad utilizadas por las variantes del BUFAR y por dos variantes del JMFAR. Cada variante del JMFAR usa un subconjunto de las características buscando maximizar la relación de compromiso entre el reconocimiento de actividades alimentarias y el costo computacional:

- JMFAR-sel: Esta variante considera el subconjunto de características resultantes luego de aplicar un método de selección de características.
- JMFAR-ns: Esta variante reduce el costo computacional y de almacenamiento de datos al no considerar aquellas características que contienen información espectral (f15, f25 y f26), evitando realizar operaciones costosas como es la transformada rápida de Fourier.

Métodos	Características																									
	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	f11	f12	f13	f14	f15	f16	f17	f18	f19	f20	f21	f22	f23	f24	f25	f26
BUFAR (CBIA)	•	•	•	•																						
BUFAR (CBEBA)	•		•	•	•	•																				
JMFAR	•						•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
JMFAR-sel								•	•	•			•	•	•		•		•	•	•	•			•	
JMFAR-ns	•						•	•	•	•	•	•	•	•		•	•	•	•	•	•	•	•	•	•	•

Tabla 4.1: Características utilizadas por los algoritmos propuestos.

Para mayor detalle se recomienda dirigirse al Anexo C del presente documento, donde se encuentra el artículo que presenta el JMFAR.

4.2. Resultados y discusión

Las tasas de reconocimiento de actividades obtenidas para cada método y variante propuesto se muestran en la Figura 4.5. Además, la relación de compromiso entre el costo computacional asociado y el nivel medio de reconocimiento (*average accuracy*) de las clases de interés (*rumia* y *pastoreo*) de cada variante se muestra en la Tabla 4.2. Se incluyen los resultados obtenidos con el algoritmo RAFAR como referencia comparativa [68].

RAFAR es capaz de reconocer mejor el *pastoreo* que BUFAR-CBIA a costa de un peor reconocimiento de la *rumia* y *otro*, obteniendo en ambos casos tasas de reconocimiento promedio similares. Sin embargo, el costo computacional de BUFAR-CBIA es 47 veces menor que RAFAR, permitiendo implementarlo en un dispositivo embebido.

Por otro lado, JMFAR reconoce igual la *rumia* que BUFAR-CBIA, pero logra una mejora considerable del *pastoreo* a costa de una pequeña disminución de *otro*. Esto implica una mejora promedio del 8.5 % de las clases de interés, incrementando su costo computacional en un 33 %, lo cual no es significativo ya que permanece en el mismo orden de magnitud.

En general, el uso de menos características de actividad en JMFAR-sel y JMFAR-ns condujo a un deterioro del reconocimiento en comparación con JMFAR. Es notable la disminución del reconocimiento de la *rumia* para JMFAR-sel (25 %). La caída promedio del reconocimiento para JMFAR-sel fue del 13.5 % en comparación con JMFAR, y del 4.0 % en comparación con BUFAR, que se produjo sin incurrir en grandes mejoras de los costos computacionales. Particularmente la variante JMFAR-sel alcanzó los peores resultados entre todos los métodos evaluados, evidenciando que durante el proceso de selección de características se pierde información relevante para distinguir las actividades. El JMFAR-ns también muestra una disminución en el reconocimiento en comparación con JMFAR (6 % para el *pastoreo* y 2 % para la *rumia*) y una mejora en comparación con BUFAR-CBIA (13 % para el *pastoreo* y 2 % para la *rumia*). Sin embargo, el costo computacional es menor para JMFAR-ns que para JMFAR y que BUFAR-CBIA. Así, JMFAR-ns alcanza mejor tasa de reconocimiento y menor costo computacional que BUFAR-CBIA.

Finalmente, BUFAR-CBEBA obtuvo la mejor tasa de reconocimiento de *rumia* y *pastoreo* de todas las variantes propuestas, teniendo una mínima confusión entre ellas. La clase *otro* tiene una disminución del 7 %, 1 % y 9 % en comparación con BUFAR-CBIA, JMFAR y JMFAR-ns, respectivamente.

Por otro lado, se realizó un ensayo para comparar la robustez de los métodos ante distintos niveles de ruido. Se utilizó ruido blanco gaussiano sumado a cada una de las señales utilizada en las proporciones adecuadas para lograr la relación señal a ruido

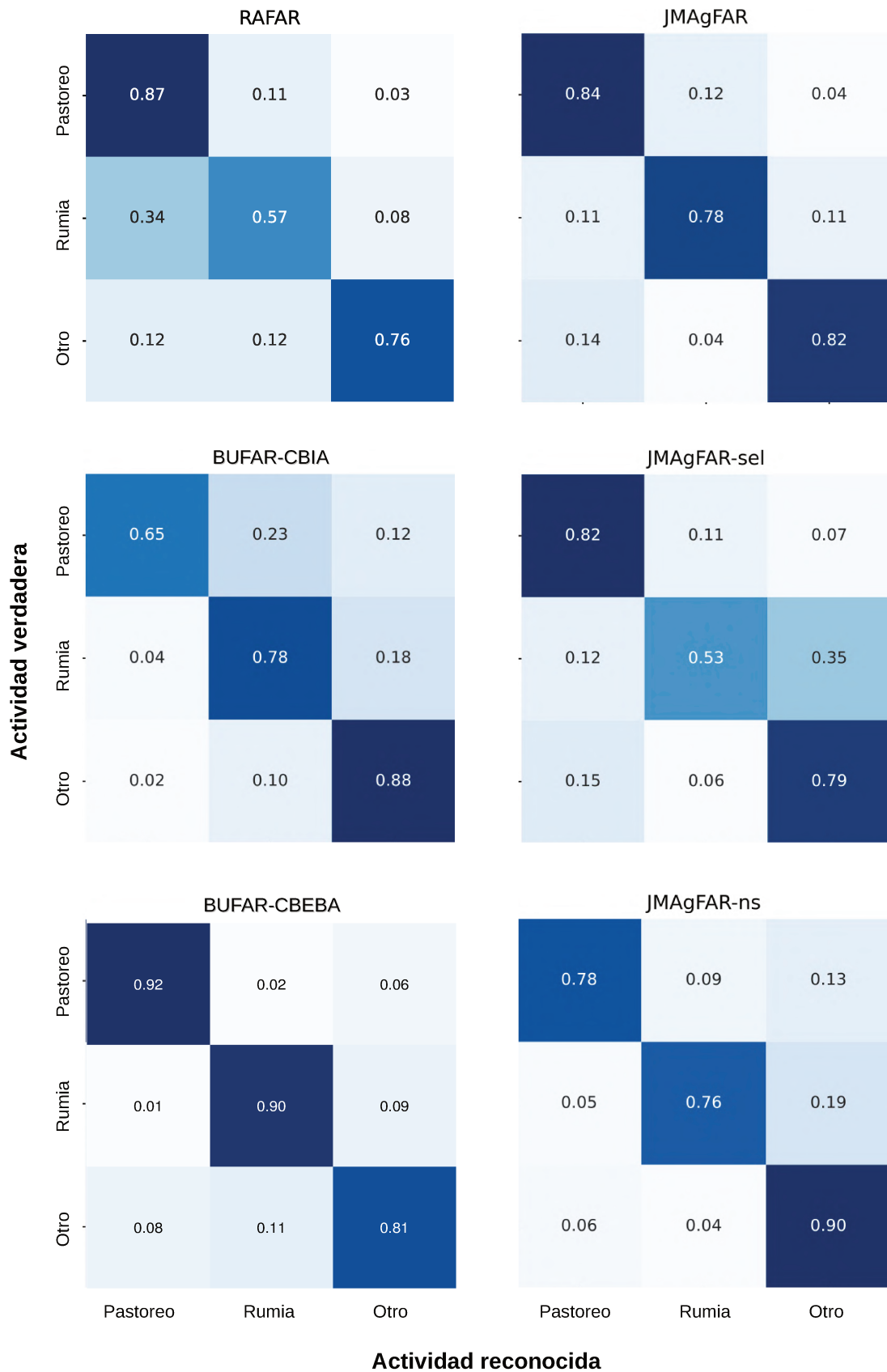


Figura 4.5: Matrices de confusión de los distintos métodos.

Método	Average accuracy (%)	Costo computacional (operaciones/s)
RAFAR	72,0	1.892.354
BUFAR-CBIA	71,5	37.966
BUFAR-CBEBA	91,0	43.185
JMFAR	81,0	50.445
JMFAR-sel	67,5	43.736
JMFAR-ns	77,0	37.645

Tabla 4.2: Rendimiento general y costo computacional de las distintas variantes.

(SNR) deseada. En la Figura 4.6 se muestran las tasas de reconocimiento promedio obtenido para distintos niveles SNR. Los resultados muestran que BUFAR-CBEBA siempre alcanza una tasa de reconocimiento superior a BUFAR-CBIA lo cual se debe, por un lado, a que las características de actividad fueron menos afectadas por los errores de clasificación de eventos masticatorios, y por el otro, a que CBEBA es un clasificador de eventos masticatorios más robusto en condiciones de ruido que CBIA, como se demostró en el capítulo 3. Por otro lado, la tasa de reconocimiento de BUFAR-CBIA disminuye considerablemente para SNR menores a 20 dB mostrando poca robustez al ruido aditivo, mientras que en BUFAR-CBEBA la tasa de reconocimiento se mantiene prácticamente constante hasta 0 dB, evidenciando una gran tolerancia frente al ruido. Finalmente, JMFAR logra una tasa de reconocimiento aproximadamente constante y superior a BUFAR-CBIA para condiciones de bajo ruido. Sin embargo, cuando la SNR disminuye a menos de 10 dB, su rendimiento disminuye notablemente, debido a que su conjunto de características de actividad se ve muy afectado por el ruido.

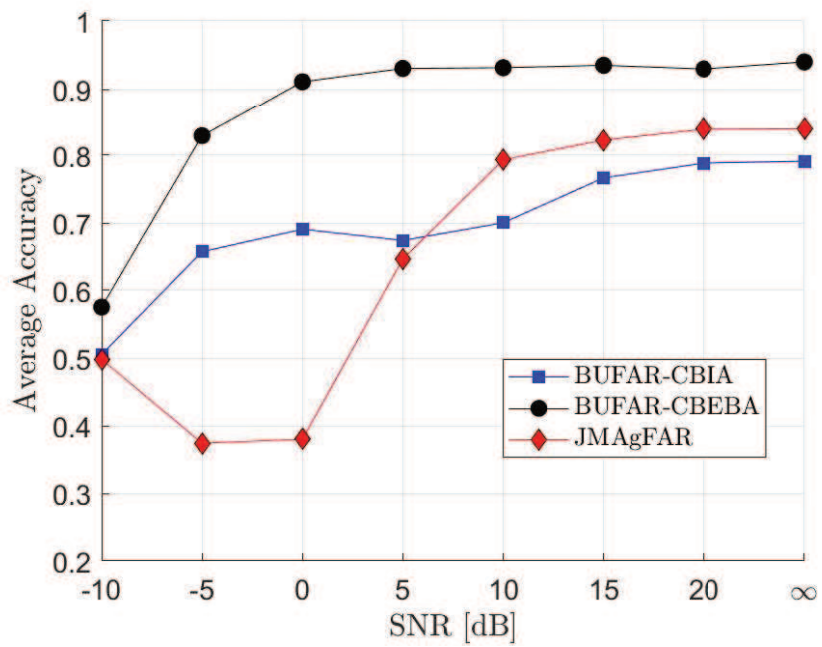


Figura 4.6: Tasa de reconocimiento para distintos niveles de ruido.

5 Sensor acústico

Uno de los factores tecnológicos claves de los sensores portátiles para el monitoreo del comportamiento animal es que éstos sean capaces de brindar información útil en forma precisa y eficaz, permitiéndole al usuario interpretarla rápidamente para poder tomar las medidas o decisiones necesarias. Para ello deberán ser lo suficientemente robustos para funcionar correctamente en distintas condiciones operativas de trabajo. También se requiere que sean de dimensiones reducidas, portables y no invasivos, para no interferir con el comportamiento natural del animal. Finalmente, se desea que trabajen en forma autónoma y continua por largos periodos de tiempo [87].

Bajo estas premisas, se busca diseñar un sensor acústico capaz de monitorear en tiempo real el comportamiento alimentario de rumiantes mediante el reconocimiento de eventos masticatorios.

Entre las estrategias generales de diseño podemos mencionar:

- Diseño completamente embebido: está basado en el uso de componentes electrónicos comerciales disponibles en el mercado. El consumo de energía del sensor no es óptimo debido a que los componentes electrónicos, utilizados para construir el sensor, muchas veces son capaces de realizar mayor cantidad de funciones que las requeridas, lo que implica un gasto innecesario de energía.
- Diseño completamente dedicado: está basado en el uso de componentes electrónicos dedicados, los cuales son diseñados y fabricados en forma específica para cumplir una determinada función. De esta forma, se logra optimizar al máximo el consumo de energía del sensor, como así también, reducir su tamaño. Sin embargo, el tiempo y complejidad requerida para diseñar cada uno de los componentes dedicados, muchas veces lo vuelve una limitante.
- Diseño híbrido: combina en distintas proporciones las dos estrategias anteriores, utilizando a su vez componentes electrónicos comerciales y dedicados.

A continuación se presenta el diseño de dos sensor acústico desarrollado siguiendo un diseño completamente embebido y un diseño híbrido, dejando el diseño completamente dedicado como trabajo futuro.

5.1. Sistema Embebido

Antecedentes

El primer dispositivo embebido utilizado para monitorear en tiempo real el comportamiento alimentario de rumiantes fue desarrollado por Deniz et al. [56]. La señal de entrada para el sistema es el sonido proveniente de las actividades alimentarias del rumiante, el cual es convertido a señal eléctrica por medio de un micrófono electret colocado sobre la frente del animal. El sistema registra tanto la información y características de la señal acústica analizada, como la información brindada por un receptor de GPS, constituyendo así paquetes de información que se van almacenando en memorias no volátiles, los cuales son transmitidos a una estación base en forma remota cada determinado periodos de tiempo.

El diagrama de bloques del sistema se puede ver en la Figura 5.1. El sistema esta internamente dividido en cuatro partes:

1. Fuente de energía: se compone por dos baterías de ion-litio de 3,7 V, la cual recarga su energía por medio de un panel solar o por un puerto USB. El manejo adecuado de los niveles de energía es administrado por convertidores de corriente continua y reguladores de voltaje.
2. Manejo de datos y comunicación: un módulo GPS es utilizado para obtener los datos de las coordenadas de geo-localización y horario, los cuales son empleados para constituir los paquetes de información que se van almacenando en memorias no volátiles del tipo flash. Además, el sistema cuenta con un módulo de comunicaciones XBee que permite la transferencia inalámbrica de los datos almacenados en las memorias.
3. Acondicionamiento analógico de la señal: se encarga de limitar el ancho de banda de la señal de adquirida por el micrófono, mejorar la relación señal a ruido mediante un amplificador con control automático de ganancia y adaptar el rango de la señal al nivel requerido por el conversor A/D del microcontrolador.
4. Procesamiento digital: El *software* embebido en el microcontrolador implementa el algoritmo CBRTA [55] e incorpora rutinas para el manejo de periféricos y la configuración del dispositivo. El modo de operación normal del *software* mantiene al dispositivo dormido (*sleep*) gran parte del tiempo. Sin embargo, el microcontrolador se despierta cada vez que se genera una interrupción. En dicho caso, se ejecutará una rutina según sea la fuente de interrupción que se genere:

- a) Interrupción del convertor A/D: cada $500 \mu\text{s}$ el microcontrolador toma una muestra de la señal acústica acondicionada. La muestra digitalizada es utilizada como entrada por el algoritmo CBRTA.
- b) Interrupción del contador: cada 5 minutos se genera un paquete de datos con la información acústica de la actividad alimentaria resultante del algoritmo CBRTA, el tiempo y la geo-localización.
- c) Interrupción del transceptor inalámbrico: se encarga de encender el módulo de transmisión y enviar el paquete de datos a la estación base.
- d) Interrupción del módulo USB: sirve para configurar el dispositivo, como así también, descargar a una PC los datos internos almacenados en las memorias flash.

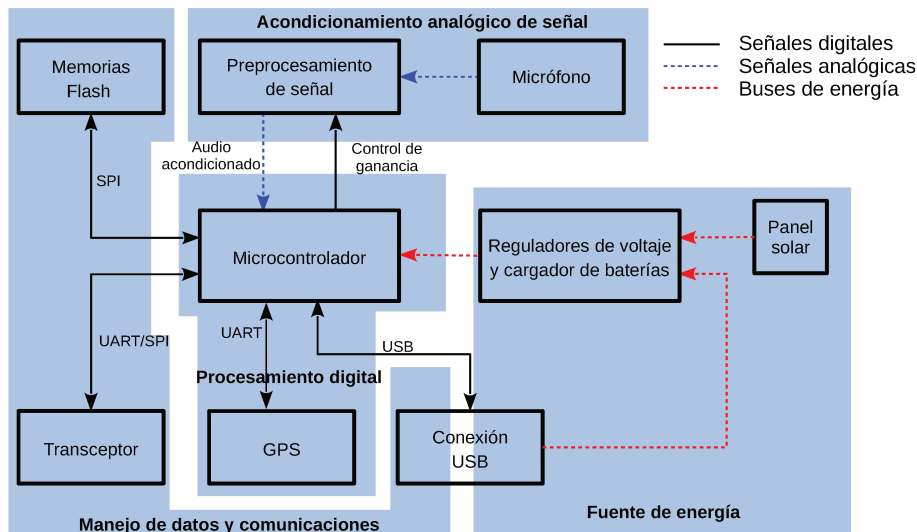


Figura 5.1: Diagrama de bloques del sistema embebido.

Para mayor detalle sobre el *hardware* y el *software* del dispositivo embebido se recomienda dirigirse a la publicación [56].

El dispositivo embebido logro una tasa de detección y de clasificación del 92% y 78%, respectivamente, bajo pruebas de laboratorio. Sin embargo, dicho sistema no fue evaluado en condiciones operativas de campo, donde el clima y el consumo energético juegan un papel importante para asegurar su robustez y su autonomía energética.

Desarrollo

Con el objetivo de reducir los posibles problemas operativos en campo asociados al hardware desarrollado, se realizó un nuevo diseño del dispositivo embebido. En la Figura 5.2 se observan los componentes físicos del sistema. La placa de circuito

impreso (PCB) del nuevo dispositivo es aproximadamente 20 % más pequeña que la anterior [56]. Además, se reemplazó el módulo de transmisión inalámbrica XBee de 500 m de alcance por un módulo LoRa permitiendo expandir el radio de cobertura hasta 6 km, brindándole al dispositivo mayor escalabilidad a grandes área de pastoreo. Para mayor detalle se recomienda dirigirse al Anexo D del presente documento, donde se encuentra el artículo que presenta dicho sistema embebido.

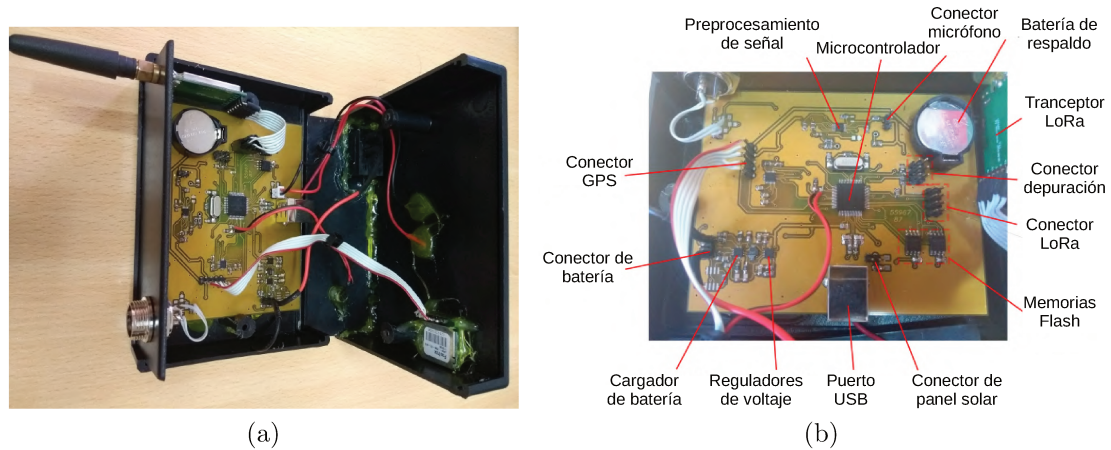


Figura 5.2: Diseño de la PCB del sistema integrado.

Resultados y discusión

Se realizó una prueba de laboratorio para verificar la correcta programación del dispositivo. Se validó el algoritmo reconocedor de eventos masticatorios obteniendo los mismos resultados reportados por Deniz et al. [56].

A fin de poder determinar el comportamiento eléctrico del dispositivo es necesario poder estimar el tiempo en que el sistema estará en modo activo, ya que es cuando se producirá el mayor consumo energético. Este tiempo está compuesto por la suma de las actividades de rumia y pastoreo, debido a que estos son los períodos en los que están presentes los eventos masticatorios.

En la Figura 5.3 se gráfica el ciclo circadiano del comportamiento alimenticio del ganado por hora del día [15]. Los valores observados reportan entre 50-70 % del tiempo diario tanto para actividades de rumia como de pastoreo, lo que concuerda con los resultados de trabajos relacionados [16, 17]. Esto indica que el dispositivo estará en modo activo alrededor del 60 % del tiempo diario.

El consumo de corriente, estimado a partir de las hojas de datos de los componentes, y el tiempo de trabajo de cada módulo en modo activo y dormido se resumen en la Tabla 5.1.

Debido al comportamiento de la fuente de alimentación conmutada, en lugar hacer de un análisis de consumo de energía, se considera más apropiado realizar un

análisis de balance de carga para determinar la autonomía del dispositivo embebido.

Módulo	Consumo de corriente [mA]	Tiempo
Microcontrolador (activo)	11,5	55 %
Microcontrolador (dormido)	5	45 %
Acondicionamiento de señal	2,5	100 %
GPS	11,5	0,18 %
Comunicación RX/TX	13/110	2,5 % / 0,12 %
Almacenamiento	4	0,18 %
Fuente de energía	20	100 %

Tabla 5.1: Consumo de corriente estimado y porcentaje de tiempo ocupado en cada modo de operación en cada módulo.

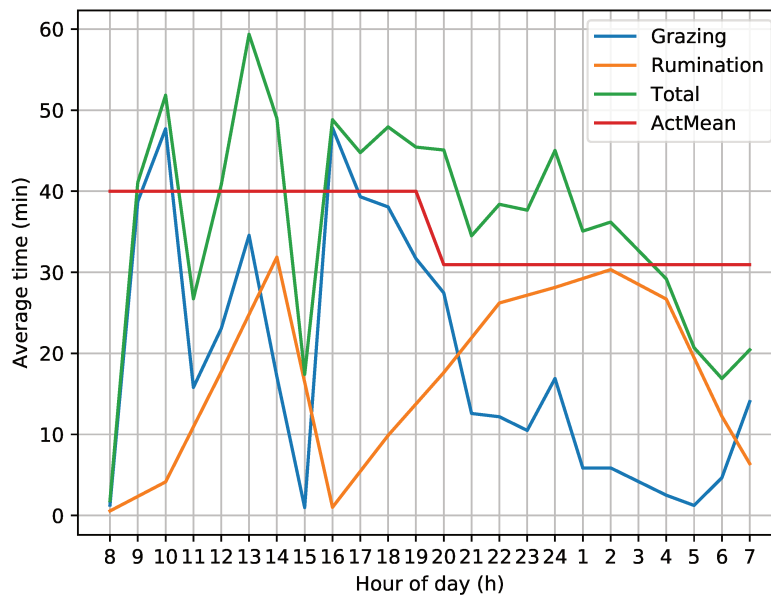


Figura 5.3: Ritmo circadiano del comportamiento de alimentación del ganado.

Según los valores establecido, el consumo de corriente promedio del dispositivo durante el modo activo y dormido es de 34,9 mA y 27,5 mA, respectivamente; por lo que su consumo de carga es de 34,9 mAh y 27,5 mAh. Por otro lado, la corriente instantánea medida proporcionada por el panel solar es de 150 mA durante las horas del día. Si suponemos que está corriente se mantiene durante 6 horas diarias, la carga media proporcionada por hora será $Q_{SP} = 150mA \cdot (6/24)h = 37,5mAh$. Según el ciclo circadiano del comportamiento de alimentación del ganado (Figura 5.3), el dispositivo estará funcionando en modo activo 40 min/hora durante 12 horas y 31 min/hora durante 12 horas. De esta forma, la carga promedio consumida por hora durante un día entero es $Q_{CR} = 31,9mAh$. Por lo tanto, el balance de carga neto es: $Q_{SP} - Q_{CR} = 5,6mAh$, demostrando que la energía recolectada es suficiente para energizar el sensor.

Por otro lado, se realizó un ensayo suponiendo que el dispositivo está en modo activo durante todo el ensayo, es decir, en las peores condiciones operativas posibles. En estas condiciones, teniendo en cuenta que la capacidad de las baterías es de 5000 mAh y sin considerar la carga proporcionada por el panel solar, el sensor tendría una autonomía de $T_A = 5000mAh/34,9mA = 143,3h \approx 6$ días. Se analizó el perfil de corriente y voltaje de las baterías durante 5,5 días con las baterías completamente cargadas inicialmente, como se observa en la Figura 5.5. Durante el día, la corriente proporcionada por las baterías disminuye a cero, cuando el sensor es energizado íntegramente por el panel solar y, al mismo tiempo, las baterías se recargan parcialmente. Sin embargo, esta energía no es suficiente para realizar una carga completa de las baterías. Durante la noche y los períodos en los que la luz solar no incide en el panel debido a la sombra de algún objeto, las baterías energizan al sistema embebido y su voltaje disminuye considerablemente. Se observa que a pesar de que las baterías son recargas, su voltaje tiende a disminuir a lo largo de los días. Al finalizar el ensayo, el sensor aún funcionaba correctamente y el voltaje tiende a estabilizarse cerca de 3,7 V, que es el voltaje nominal de las baterías.

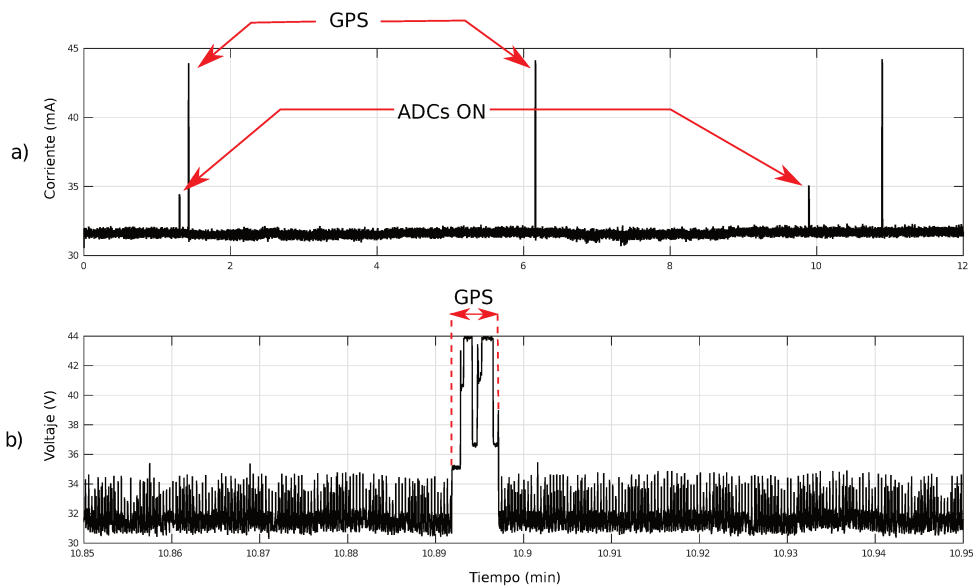


Figura 5.4: Vista macro (a) y zoom (b) sobre los principales eventos de consumo de corriente.

La Figura 5.4 muestra que el consumo de corriente medido es de 32 mA en modo activo. Cada 5 minutos, el GPS se activa y hay un consumo máximo de 44 mA, mientras que cada 9 minutos son encendidos en forma simultánea todos los canales del convertor A/D, vinculados a procesos de monitoreo y análisis. Los pequeños picos de 34 mA cada 500 μs corresponden a la ejecución del algoritmo CBRTA tras finalizar la conversión A/D, lo cual se encuentra en concordancia con el valor teórico calculado considerando los módulos que trabajan todo el tiempo (microcontrolador,

acondicionamiento de señal y fuente de alimentación) (Tabla 5.1).

Vale la pena resaltar que los componentes electrónicos utilizados para desarrollar el sensor fueron adquiridos durante el año 2016, con excepción de las baterías y el modulo LoRa. Si dicho dispositivo fuese desarrollado con componentes electrónicos actuales, se lograría disminuir considerablemente el consumo de energía. Sin embargo, como ya se mencionó anteriormente, utilizando componentes electrónicos comerciales difícilmente se logre un menor consumo energético que usando componentes electrónicos dedicados.

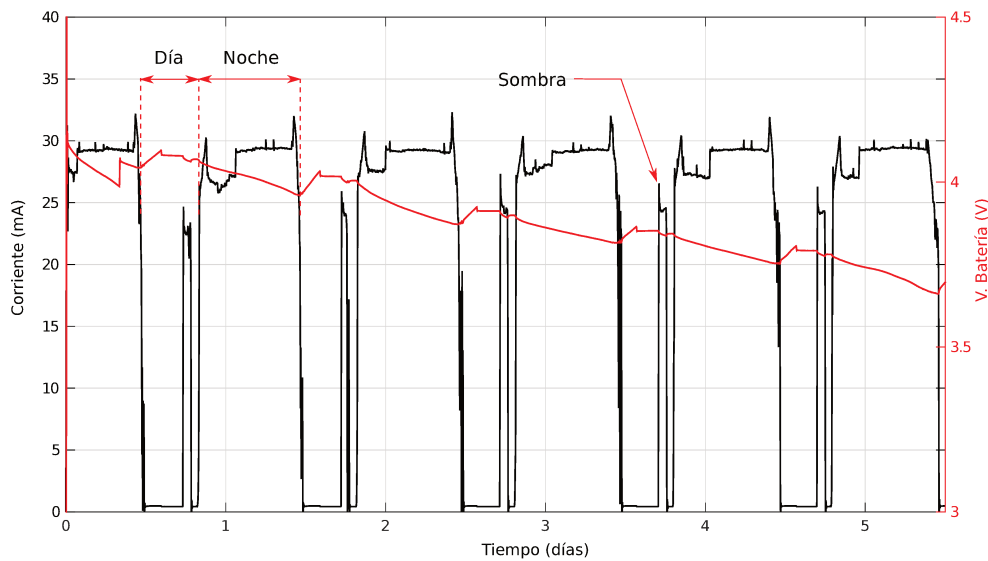


Figura 5.5: Consumo de corriente (línea negra) y voltaje (línea roja) durante 5,5 días.

5.2. Sistema Híbrido

Para diseñar un sensor acústico utilizando un sistema híbrido que cumpla los requerimientos, existe una gran variedad de optimizaciones posibles a realizar, como por ejemplo, implementar en un circuito integrado de aplicación específica (*i*) el acondicionamiento de la señal de audio, (*ii*) la fuente de alimentación, (*iii*) el microcontrolador o (*iv*) una combinación de ellos.

El sistema híbrido aquí propuesto realiza un nuevo diseño del procesamiento digital de los datos. El diagrama general del sistema desarrollado se muestra en la Figura 5.6, en el cual los bloques de color naranja son diseñados con componentes comerciales y el bloque de color verde es diseñado en forma de circuito dedicado. La señal acústica es adquirida y digitalizada por un micrófono del tipo MEMS (del inglés Micro-electro-mechanical systems), la cual pasa por un convertor de datos para ser entregada en el formato correcto a un microprocesador. El microprocesador

se encarga de reconocer los eventos masticatorios y almacenar los resultados en una memoria FIFO (del inglés First In, First Out). Por otro lado, el microcontrolador, el módulo GPS y el módulo LoRa se encuentran apagados. Sin embargo, estos módulos se encienden cuando la memoria FIFO se encuentra llena. Cuando esto ocurre, el microcontrolador envía a la estación base los datos de la memoria FIFO y los datos de fecha y geo-localización obtenidos del GPS. Una vez finalizada la transmisión, estos módulos vuelven a apagarse.

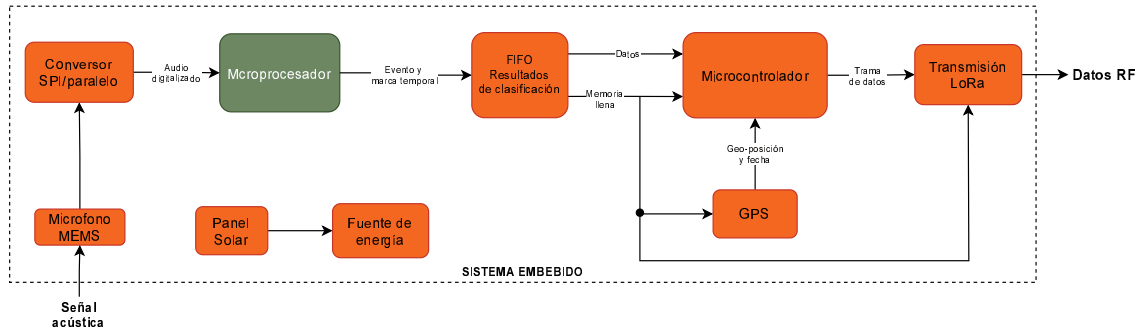


Figura 5.6: Diagrama de bloques del sistema híbrido.

Internamente, el microprocesador está compuesto por dos procesadores con un conjunto de instrucciones de uso específico (ASIP, del inglés Application-specific instruction set processor) llamados *NanoController* y *TTA* (del inglés Transport Triggered Architecture), y por distintas memorias para almacenar datos en forma interna. El *NanoController* se encuentra siempre encendido y se encarga de controlar la activación del *TTA* cuando es necesario. El diagrama general de la arquitectura del microprocesador implementado se muestra en la Figura 5.7.

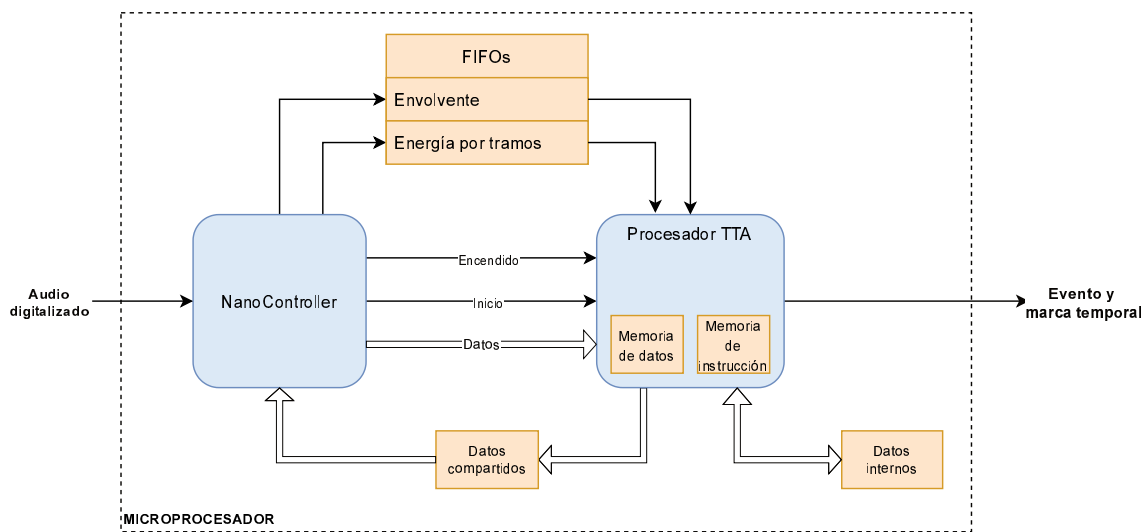


Figura 5.7: Diagrama de bloques del microprocesador conformado por procesadores (bloques azules) y memorias (bloques naranjas).

El algoritmo reconocedor de eventos masticatorios embebido en el microproce-

sador es el CBEBA, previamente descrito en el capítulo 3.2, el cual se encuentra internamente dividido en 6 etapas (ver Figura 3.3). Las primeras etapas del algoritmo asociadas a la detección de un posible evento fueron implementadas en el *NanoController*, mientras que el resto de las etapas, asociadas a la clasificación de evento, fueron implementadas en el *TTA*.

Inicialmente el *TTA* se encuentra apagado, y el *NanoController* se encarga de leer y preprocesar las muestras entregadas por el conversor SPI/paralelo en tiempo real (etapa 1). Con estas muestras se calcula la envolvente decimada y la energía por tramo, las cuales son almacenadas en dos memorias FIFOs (etapa 2). Además, el valor de la envolvente decimada es utilizada para encontrar picos en la envolvente y, a su vez, comparada con el umbral adaptativo para detectar un posible evento masticatorio (etapa 3). El valor del umbral adaptativo y otros parámetros internos son leídos de una memoria de datos compartidos por el *TTA* y el *NanoController*. Cabe remarcar que la búsqueda de un evento se produce sobre la muestra actual de la señal envolvente, y no en la posición media de la señal envolvente almacenada, tal como se describió en el capítulo 3.2. Cuando un evento es detectado, el *NanoController* deberá enviarle al *TTA*, en el momento adecuado, una señal para encenderlo y una señal de inicio para habilitar su ejecución. La señal de encendido deberá ser enviada con el suficiente tiempo de antelación para que el *TTA* pueda leer su programa de la memoria de instrucción, antes de recibir la señal de inicio. Por otro lado, una vez producida la detección del evento, el *NanoController* continúa procesando muestras hasta que el evento se encuentre centrado dentro de la FIFO. Cuando esto ocurre, el *NanoController* envía al *TTA* la señal de inicio.

Una vez que el *TTA* se inicia, éste debe cargar, por un lado, el estado actual de sus parámetros internos, las cuales están en parte almacenados en una memoria de datos internos y el resto son transmitidas por el *NanoController*, y por el otro, la señal envolvente y la señal de energía por tramos almacenadas en las memorias FIFOs. A continuación, el *TTA* se encarga de delimitar donde se encuentra el posible evento detectado (etapa 3), extraerle sus características (etapa 4), determinar si realmente se trata de un evento, y en caso afirmativo, clasificarlo (etapa 5) y almacenarlo junto con una marca de tiempo en la memoria FIFO externa al microprocesador. Luego, se realiza un ajuste de los parámetros internos (etapa 6), los cuales son en parte almacenados en la memoria de datos internos y en parte almacenados en la memoria de datos compartidos para ser utilizados por el *NanoController*. Finalmente, el *TTA* se apaga. El CBEBA implementado utiliza un árbol de decisión para realizar la clasificación de eventos masticatorios aprovechando su bajo costo computacional. La arquitectura del microprocesador propuesta aprovecha las características favorables que presentan el *NanoController* y el *TTA* individualmente al embeber el CBEBA en forma conjunta, las cuales son descritas a continuación.

NanoController

La mayoría de los sistemas embebidos están basados en el uso de microcontroladores de propósito general, los cuales permiten reducir el consumo energético al entrar en modo dormido. Sin embargo, durante este modo de operación el núcleo del microcontrolador siempre debe permanecer energizado para poder despertar al microcontrolador, causando un consumo de energía continuo debido a la corriente de fuga. A su vez, los microcontroladores de propósito general poseen una arquitectura de procesador con un camino de datos (datapath) y un set de instrucciones establecido, por lo que su consumo energético no se encuentra optimizado para una aplicación específica. En este sentido, los grupos “*Chair for Chip Design for Embedded Computing (C3E)*” de la *Technische Universität Braunschweig, Alemania* e “*Institute of Microelectronic Systems (IMS)*” de la *Leibniz Universität Hannover, Alemania* desarrollaron un procesador llamado *NanoController* pensado como una arquitectura mínima para ser utilizado en sistemas de ultra bajo consumo encargado de realizar tareas simples en forma periódica, como por ejemplo, procesar muestras de audio a una determinada frecuencia fija. El *NanoController* se encuentra pensado para ser utilizado en conjunto con un microcontrolador o microprocesador auxiliar encargado de realizar tareas más complejas que ocurran en pocas ocasiones. Esta configuración presenta la ventaja de que permite mantener apagado al microcontrolador o microprocesador auxiliar gran parte del tiempo logrando reducir el consumo promedio total del sistema embebido. Además, la arquitectura programable del *NanoController* utiliza mecanismos de optimización de codificación del set de instrucciones de longitud variable para incrementar la densidad de código, reducir el tamaño de la memoria de instrucción, las corrientes de fugas y el área de silicio [88].

Procesador TTA

Un procesador con arquitectura activada por transporte (*TTA*), se puede describir como una arquitectura de palabra de instrucción muy larga (VLIW, del inglés Very Long Instruction Word) en la cual los datapaths internos están expuestos en el set de instrucciones [89]. Un procesador *TTA* está constituido por distintos *archivos de registros* (RF) y *unidades funcionales* (FU). Cada FU implementa internamente una o más operaciones computacionales. Las FUs y los RFs están conectados a través de una interconexión de *buses* de transporte programable creando datapaths programables. Mientras que en la mayoría de las arquitecturas de procesadores la palabra de instrucción especifica la operación computacional a realizar, en una arquitectura *TTA* ésta indica cómo se deben mover los datos entre las distintas FUs por los distintos buses. Los operandos se transfieren a una FU a través de *sockets* programables, uniendo una FU a los buses mediante *conexiones*. La operación computacional de

una FU se activa como un efecto secundario de escribir datos en ella. Las FUs que requieren acceder a la memoria de datos o comunicarse hacia afuera del procesador reciben el nombre especial de *unidad de carga/almacenamiento*. La FU que controla la ejecución del programa accediendo a la memoria de programa recibe el nombre de *unidad de control global*. En la Figura 5.8 se muestra un *TTA* detallando las partes que lo componen. El concepto de arquitectura configurable permite agregar nuevas FUs con funcionalidad personalizada, modificar la cantidad de buses para permitir la transferencia de datos en paralelo, y configurar selectivamente la interconexión de los sockets con los buses, considerando un compromiso entre la flexibilidad de programación y la complejidad de la interconexión. Se utilizaron las herramientas de diseño de código abierto *TTA HW/SW Co-Design Environment (TCE)* durante el proceso de exploración del espacio de diseño de este microprocesador [90].

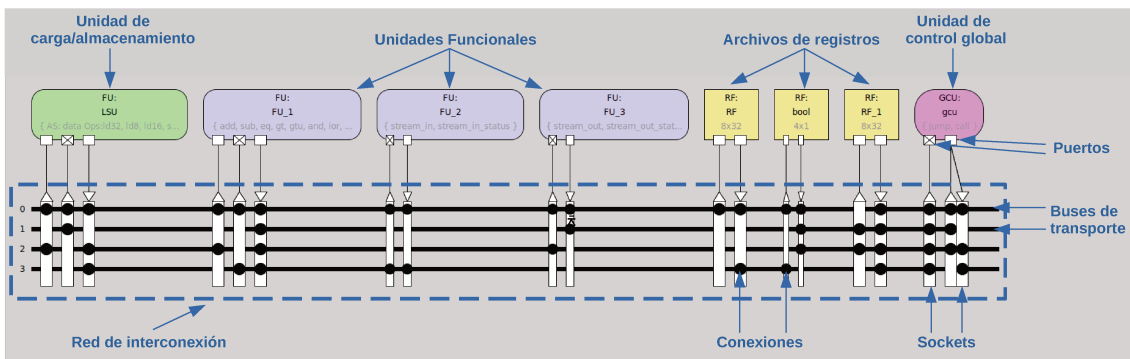


Figura 5.8: Partes de un procesador TTA.

Resultados y discusión

El CBEBA fue desarrollado con una precisión del tipo doble para representar los datos. Para poder representar estos datos con formato de punto fijo en el microprocesador, se realizó un análisis de la resolución requerida y una conversión del algoritmo utilizando representación y aritmética de punto fijo.

Para elegir la mejor arquitectura del *TTA* se exploró un espacio de diseño, variando la cantidad de *buses*, *sockets*, *FUs* y *RFs*. En la Figura 5.9 se muestra la arquitectura minimalista elegida, lograda gracias al bajo costo computacional requerido por las etapas del CBEBA implementado (etapas 3 a 6). El largo de la palabra de instrucción del *TTA* es de 18 bits y su memoria de instrucción tiene una profundidad de 4480 posiciones. La frecuencia mínima de operación requerida es de 1.980.800 Hz.

Cabe señalar de que todos los datos y registros utilizados en el *TTA* fueron representado con un máximo de 16 bits, a pesar de que el *TTA* es un procesador de 32 bits. Esto permitió reducir el tamaño de la memoria de datos del *TTA*, y ade-

más, permitirá migrar dicha implementación a futuros procesadores *TTA* de 16 bits, logrando reducir aún más la superficie y el consumo energético del microprocesador.

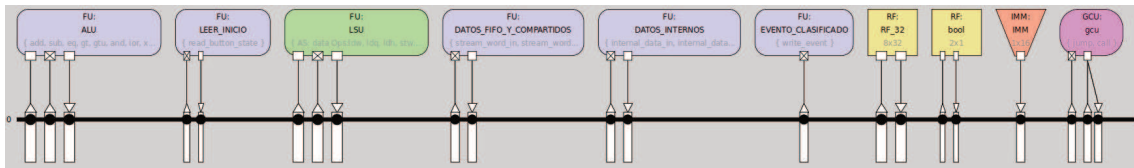


Figura 5.9: Arquitectura del procesador TTA.

Se realizó un ensayo para evaluar el rendimiento del reconocedor de eventos masticatorios. El experimento fue realizado mediante una combinación de simulación y emulación. El funcionamiento del *NanoController* fue simulado, tomando como entrada el audio digitalizado y almacenando los resultados en un archivo binario. Cada vez que el *NanoController* debía iniciar el *TTA*, los datos eran transferidos vía UART por el puerto USB a una FPGA Artix-7 XC7A35T-L1CSG324I (Xilinx Inc., San José, CA, USA), en la cual se encontraba emulado el *TTA*. La comunicación bidireccional USB mantenía el lazo cerrado necesario entre los procesadores.

La Tabla 5.2 muestra las tasas de detección y clasificación de eventos masticatorios del sistema implementado y los valores de referencia obtenido en la implementación en punto flotante, presentados previamente en el capítulo 3. Se comprobó si la discrepancia de los resultados obtenidos eran estadísticamente significativa ($p < 0,05$) usando la prueba de los rangos con signo de Wilcoxon [86]. Los valores con fondo verde dentro de una misma fila y para un mismo dataset difieren significativamente, mientras que los valores con fondo rosa, en iguales condiciones, no muestran diferencias significativas. Al comparar los resultados obtenidos respecto al valor alcanzado en la versión implementada en punto flotante, se observa que en la detección de eventos masticatorios, el *F1-score* presenta una leve disminución del 0,5% en el dataset sin ruido DS1. Sin embargo, esta disminución aumenta al 3,0% en el dataset con ruido DS2. Además, el *recall* también disminuye en un 3,0% en DS2, indicando un aumento en la cantidad de eventos masticatorios no detectados. Por otro lado, no es posible afirmar que exista una disminución en la tasa de clasificación de eventos, ya que los resultados no muestran ser estadísticamente significativos.

Estos resultados experimentales permiten afirmar que esta implementación del CBEBA embebida en el microprocesador obtendrá muy buenas tasas de reconocimiento de eventos masticatorios al ser fabricado el circuito integrado e incorporado al sensor acústico.

		Dataset DS1		Dataset DS2	
		Punto Flotante	Punto Fijo	Punto Flotante	Punto Fijo
Detección	<i>Recall</i>	98,4 ± 1,0	97,4 ± 1,1	98,0 ± 0,9	95,0 ± 1,3
	<i>Precision</i>	99,2 ± 0,5	99,3 ± 0,8	95,8 ± 2,1	92,8 ± 2,4
	<i>F1 – score</i>	98,8 ± 0,6	98,3 ± 0,6	96,9 ± 1,4	93,9 ± 1,7
Clasificación	<i>Recall_M</i>	92,0 ± 2,6	91,3 ± 3,1	89,8 ± 4,3	90,0 ± 4,6
	<i>Precision_M</i>	92,2 ± 2,8	91,8 ± 3,4	87,0 ± 5,5	87,9 ± 5,6
	<i>F1 – score_M</i>	92,1 ± 2,4	91,5 ± 3,1	88,4 ± 4,7	88,9 ± 4,9
	<i>Recall_G</i>	88,9 ± 3,7	88,1 ± 4,5	87,9 ± 4,8	87,2 ± 6,3
	<i>Precision_G</i>	89,3 ± 4,3	89,3 ± 4,6	81,5 ± 8,6	83,3 ± 8,8

Tabla 5.2: Comparación de resultados (media ± desvío estándar; %) entre CBEBA utilizando representación en punto flotante y punto fijo en dos datasets.

6 Conclusiones y Trabajo Futuro

6.1. Conclusiones

En esta tesis se abordó el diseño, desarrollo e implementación de un sistema que permite el monitoreo en tiempo real del comportamiento alimentario de rumiantes en condiciones de pastoreo libre. Para ello se desarrollaron e implementaron diversos algoritmos que utilizan herramientas de aprendizaje maquina y procesamiento de señales. Además se desarrollaron dos sensores acústicos como parte del sistema.

Dado que los algoritmos existentes en el estado del arte no son capaces de reconocer eficazmente eventos masticatorios (*chew*, *bite* y *chew-bite*) en condiciones de ruido, se buscó un nuevo conjunto de características robustas asociadas a eventos masticatorios ante diversas condiciones operativas. Este conjunto de características son extraídas y combinadas con técnicas de aprendizaje maquina para reconocer eventos masticatorios en un nuevo algoritmo (CBEBA). El CBEBA es un sistema de reconocimiento de patrones basado en el uso de la potencia instantánea de la señal de entrada y es el primer método utilizado en bovinos capaz de distinguir entre los eventos *chew* producidos durante el pastoreo o durante la rumia (*grazing-chew* y *rumination-chew*, respectivamente), además de también reconocer los eventos *bite* y *chew-bite*. El CBEBA se basa internamente en etapas de preprocesamiento, cálculo de señales y detección de posibles eventos candidatos usando umbrales variables. Cuenta además con una etapa de extracción de características y una etapa de clasificación basada en métodos de aprendizaje maquina. Además, hay una etapa encargada de actualizar variables internas utilizando métodos adaptativos de procesamiento de señales para mejorar su desempeño en ambientes ruidosos o condiciones operativas cambiantes en el tiempo. Dicho algoritmo demostró ser robusto al evaluarse con dos bases de datos obtenidas en diferentes condiciones de ruido, logrando resultados superiores al algoritmo de referencia del estado del arte (CBIA), con un leve incremento en su costo computacional, siendo factible su implementación en un sensor acústico de bajo consumo energético que opere en tiempo real.

Se desarrolló un sistema general para el reconocimiento y estimación de tiempos de actividades alimentarias (rumia y pastoreo). Estas actividades se encuentran a una escala temporal mayor que los eventos masticatorios que las componen. Se evaluaron dos algoritmos, y en sendos algoritmos sus respectivas variantes. Uno de los algoritmos (BUFAR) se basa en la extracción de características en bloques de duración fija utilizando información de los eventos masticatorios reconocidos. Se evaluó

una variante utilizando el reconocedor de eventos masticatorios CBIA, y otra variante utilizando el CBEBA. El otro algoritmo (JMFAR) no utiliza información de los eventos masticatorios reconocidos, logrando independizarse del error cometido por el clasificador de eventos. En este caso, el algoritmo hace uso de un conjunto de características con información temporal, estadística y espectral de los eventos masticatorios detectados. Además de la variante que utiliza todo el conjunto de características, se propusieron dos variantes reduciendo la cantidad de características utilizadas por cada una, disminuyendo el costo computacional asociado. En la evaluación del desempeño, ambos algoritmos obtuvieron mejores resultados que el algoritmo del estado del arte de referencia (RAFAR). Particularmente la variante BUFAR-CBEBA alcanzó los mejores resultados, teniendo muy poca confusión entre rumia y pastoreo. Además, se evaluó la robustez de los algoritmos al agregar ruido blanco artificial en las señales, mostrando que la variante BUFAR-CBEBA tiene mayor tolerancia al ruido como consecuencia de la robustez mostrada por el clasificador de evento CBEBA y a que éste distingue entre las masticaciones producidas durante el pastoreo y la rumia (eventos *grazing-chew* y *ruminaton-chew*, respectivamente). El costo computacional de todas las variantes fue dos órdenes de magnitud menor que RAFAR, permitiendo una posible implementación en un sensor acústico de cualquier de ellas.

Finalmente como parte del sistema y para posibilitar la ejecución de un reconocedor de eventos masticatorios, se diseñaron e implementaron dos sensores acústicos a medida para la aplicación. Uno de los sensores está diseñado con circuitos integrados comerciales, el cual realiza el procesamiento de los sonidos y almacena solo los resultados del mismo, junto con información provista por un receptor de GPS presente en la placa. Toda esta información se va almacenando en forma de paquetes en memorias flash y cada un determinado periodo fijo de tiempo son transmitidos en forma inalámbrica a una estación base remota. El otro sensor está diseñado utilizando en forma combinada circuitos integrados comerciales y de aplicación específica, el cual almacena los eventos masticatorios reconocidos y una marca de tiempo temporalmente en una memoria, y cada vez que la memoria se encuentra llena, sus datos junto con información del GPS, son transmitidos en forma inalámbrica a una estación base.

En ambos sensores, el sistema está especialmente diseñado, a nivel de *hardware* y *firmware*, para tener un bajo consumo de energía. Además, cuentan con un sistema de alimentación de alta eficiencia energética y dos baterías que son cargadas por un panel solar. Un análisis energético mostró que el primer sensor es capaz de operar en forma autónoma sin necesidad de reemplazar en forma manual las baterías.

Con estas herramientas de análisis para el monitoreo alimentario de rumiantes en condiciones de pastoreo libre, queda demostrado que el método acústico posibili-

ta tanto la estimación de los tiempos de rumia y pastoreo, como el reconocimiento individual de movimientos mandibulares y su completa caracterización en diversas condiciones operativas de ruido. A su vez, el bajo costo computacional de los algoritmos propuestos permite pensar en una escalabilidad de los sensores acústicos a nivel de rodeos completos sin la necesidad de dispositivos muy complejos y con la ventaja de operar en forma autónoma permitiendo monitorear patrones de comportamiento de larga duración.

6.2. Artículos y patente

Los resultados obtenidos durante la realización de la presente tesis fueron publicados en:

Revistas:

1. Chelotti, J.O., Vanrell, S.R., Martínez-Rau, L.S., Galli, J.R., Utsumi, S.A., Planisich, A.M., Almirón S.A., Giovanini, L.L. & Rufiner, H.L.. “Using segment-based features of jaw movements to recognize foraging activities in grazing cattle”. *Biosystems Engineering*, 2022. En revisión.
2. Martínez-Rau, L.S., Chelotti, J.O., Vanrell, S.R., Galli, J.R., Utsumi, S.A., Planisich, A.M., Rufiner, H.L. & Giovanini, L.L. . “A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals”. *Computers and Electronics in Agriculture*, 2021.
3. Chelotti, J.O., Vanrell, S.R., Martínez-Rau, L.S., Galli, J.R., Planisich, A.M., Utsumi, S.A., Milone, D.H., Giovanini, L.L., Rufiner, H.L.. “An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle”, *Computers and Electronics in Agriculture*, vol. 173, pp. 105443, 2020.

Congresos:

1. Martínez-Rau, L.S., Chelotti, J.O., Vanrell, S.R. & Giovanini, L.L.. “Developments on real-time monitoring of grazing cattle feeding behavior using sound”. *IEEE International Conference on Industrial Technology (ICIT)*, pp. 771-776, 2020.
2. Martínez-Rau, L.S., Deniz, N. N., Chelotti, J.O., Giovanini, L.L. & Kler, P.A.. “Acoustic real-time sensor for ingestive behaviour of grazing cattle”. *IX Congreso de Microelectrónica Aplicada*, San Fernando del Valle de Catamarca, 2018.

Además, debido al impacto tecnológico de las investigaciones y desarrollos implicados, se protegió su propiedad intelectual a partir de la siguiente solicitud de patente:

1. Patente Argentina N° 20220100910. Inventores: L. L. Giovanini, H. L. Rufiner, L. S. Martínez-Rau, J. R. Galli y J. O. Chelotti, “Sistema para el seguimiento y evaluación de actividades alimentarias en producción animal basado en la energía del sonido”. Cp-titulares: UNL, UNR, CONICET.

6.3. Trabajo futuro

Se buscará mejorar las tasas de reconocimiento de los algoritmos desarrollados en la presente tesis. Se incorporará información de movimiento proveniente de diferentes sensores (acelerómetros, unidades de medición inercial, entre otros) para mejorar el desempeño en el reconocimiento de actividades alimentarias. Además, por un lado, se exploraran distintas técnicas de cancelación activa de ruido que permitan utilizar el reconocedor de eventos masticatorios desarrollado en escenarios más desafiantes, como establos, corrales y galpones, donde hay ruidos más fuertes y variados que en condiciones de pastoreo libre. Por otro lado, se buscará mejorar el reconocimiento de actividades a través de la incorporación de una etapa de segmentación dinámica de bajo costo computacional mediante la detección derivada de cambios dinámicos o de las secuencias de eventos.

Finalmente se seguirá trabajando en el perfeccionamiento de los sensores acústicos. El sensor ya fabricado será evaluado en el campo de aplicación por largos periodos de tiempo en distintos animales, contemplando variaciones inter-animal y climáticas. Por otro lado, en el otro sensor acústico, se propone incorporar el manejo de los periféricos GPS y LoRa en el microprocesador diseñado, evitando así el uso de un microcontrolador y reduciendo aún más su consumo y tamaño. También se propone enviar a fabricar un pequeño lote de este microprocesador para verificar su correcto funcionamiento y poder incorporarlo en la fabricación del sensor, y así poder realizar distintas pruebas de campo.

Anexos

Referido a los artículos incluidos en los Anexos A, B, C, D y E.

- Martínez-Rau, L.S., Chelotti, J.O., Vanrell, S.R., Galli, J.R., Utsumi, S.A., Planisich, A.M., Rufiner, H.L. & Giovanini, L.L. . “A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals”. *Computers and Electronics in Agriculture*, 2021. En revisión.
- Chelotti, J.O., Vanrell, S.R., Martínez-Rau, L.S., Galli, J.R., Planisich, A.M., Utsumi, S.A., Milone, D.H., Giovanini, L.L., Rufiner, H.L.. “An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle”, *Computers and Electronics in Agriculture*, vol. 173, pp. 105443, 2020.
- Chelotti, J.O., Vanrell, S.R., Martínez-Rau, L.S., Galli, J.R., Utsumi, S.A., Planisich, A.M., Almirón S.A., Giovanini, L.L. & Rufiner, H.L.. “Using jaw movement-agnostic features to recognize foraging activities in grazing cattle”. *Computers and Electronics in Agriculture*, 2021. En revisión.
- Martínez-Rau, L.S., Chelotti, J.O., Vanrell, S.R. & Giovanini, L.L.. “Developments on real-time monitoring of grazing cattle feeding behavior using sound”. *IEEE International Conference on Industrial Technology (ICIT)*, pp. 771-776, 2020
- Martínez-Rau, L.S., Deniz, N. N., Chelotti, J.O., Giovanini, L.L. & Kler, P.A.. “Acoustic real-time sensor for ingestive behaviour of grazing cattle”. *IX Congreso de Microelectrónica Aplicada*, San Fernando del Valle de Catamarca, 2018

el tesista declara haber contribuido principalmente en el diseño, implementación y evaluación de los algoritmos descritos y los experimentos realizados para obtener los resultados que allí se presentan. Estas tareas fueron realizadas bajo la guía y supervisión del director Dr. L. L. Giovanini y codirector de tesis Dr. P. A. Kler. En cuanto a la escritura de los artículos, el tesista ha sido el autor principal de los artículos de los Anexos A, D y E, y ha colaborado en la escritura de los artículos de los Anexos B y C guiado por los comentarios, sugerencias y revisiones del director,

codirector de tesis y otros coautores que en cada artículo se indican. Los abajo firmantes avalan esta declaración.

Dr. L. L. Giovanini

Director

Dr. P. A. Kler

Co-director

Referido a la solicitud de patente incluida en el Anexo F.

- Patente nacional. Inventores: L. L. Giovanini, H. L. Rufiner, L. S. Martínez-Rau, J. R. Galli, J. O. Chelotti y S. R. Vanrell, “Sistema para el seguimiento y evaluación de actividades alimentarias en producción animal basado en la energía del sonido”. Titulares: UNL, CONICET. En Trámite.

el tesista declara haber contribuido principalmente en el diseño, implementación y evaluación de los algoritmos descriptos y los experimentos realizados para obtener los resultados que allí se presentan. Estas tareas fueron realizadas bajo la guía y supervisión del director Dr. L. L. Giovanini y codirector de tesis Dr. P. A. Kler.

En cuanto a la solicitud de patente presentada en el Anexo F guarda mayor relación con los resultados presentados en los artículos de los Anexos A, D y E, en los cuales el tesista es autor principal. Los abajo firmantes avalan esta declaración.

Dr. L. L. Giovanini

Director

Dr. P. A. Kler

Co-director

Anexo A

**A robust computational approach for
jaw movement detection and
classification in grazing cattle using
acoustic signals**

A robust computational approach for jaw movement detection and classification in grazing cattle using acoustic signals

Luciano S. Martinez Rau^a, José O. Chelotti^a, Sebastián R. Vanrell^a,
Julio R. Galli^{b,c}, Santiago A. Utsumi^d, Alejandra M. Planisich^c,
H. Leonardo Rufiner^{a,e}, Leonardo L. Giovanini^a

^a*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i),
FICH-UNL/CONICET, Argentina*

^b*Instituto de Investigaciones en Ciencias Agrarias de Rosario, IICAR, UNR-CONICET,
Argentina*

^c*Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Argentina*

^d*W.K. Kellogg Biological Station and Department of Animal Science, Michigan State
University, United States*

^e*Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina*

Abstract

Monitoring behaviour of the grazing livestock is a difficult task because of its demanding requirements (continuous operation, large amount of information, computational efficiency, device portability, precision and accuracy) under harsh environmental conditions. Detection and classification of jaw movements (JM) events are essential for estimating information related with foraging behaviour. Acoustic monitoring is the best way to classify and quantify ruminant events related with its foraging behaviour. Although existing acoustic methods are computationally efficient, a common failure for broad applications is the deal with interference associated with environmental noises. In this work, the acoustic method, called Chew-Bite Energy Based Algorithm (CBEBA), is proposed to automatically detect and classify masticatory events of grazing cattle. The system incorporates computations of instantaneous power signal for JM-events classification associated with chews, bites and composite chew-bites, and additionally between two classes of chew events: i) low energy chews that are associated with rumination and ii) high

Email address: lmrau@sinc.unl.edu.ar (Luciano S. Martinez Rau)

energy chews that are associated with grazing. The results demonstrate that CBEBA achieve a recognition rate of 91.9% and 91.6% in noiseless and noisy conditions, respectively, with a high classification precision and a marginal increment of computational cost compared to previous algorithms, suggesting feasibility for implementation in low-cost embedded systems.

Keywords: Acoustic monitoring, Cattle grazing behaviour, Jaw movement classification, Noise robustness, Pattern recognition, Sound energy analysis.

1. Introduction

Precision livestock farming typically integrates smart animal monitoring technologies to aid farmers with relevant management decisions regarding animal nutrition, health and welfare (Michie et al., 2020). The deployment of animal monitoring dashboards has been enhanced recently by improved sensors (Andriamandroso et al., 2016), advanced communication technologies and enhanced visualization tools allowing for rapid inspection of production traits and behaviours associated with specific activities, changes of location and body posture (Berckmans, 2014).

The use of modern technologies based on fixed video cameras allow for individual or group behaviour monitoring in an automatic, continuous and non-intrusive way in a given fixed area (Fuentes et al., 2020). Their use is limited to small farming areas such as pens and stables. On the other hand, the use of small wearable video cameras on animals would allow to expand the operating region, although their application still needs further development (Saitoh and Kato, 2021). Thus, wearable sensors are the most widely used acquisition method to cover large farm and field areas. However, their operational requirements, primarily device portability, robustness and power capabilities, along with the computational cost and complexity of analytical components often represents an obstacle for further technological progress and adoption (Stone, 2020).

Ones of the most frequently used monitoring techniques are the position and motion sensors, which allows the surveillance of cattle and sheep movements (Andriamandroso et al., 2016). Nose band sensors (Nydegger et al., 2010; Werner et al., 2018; Zehner et al., 2017), multidimensional accelerometers (Andriamandroso et al., 2017; Greenwood et al., 2017; Smith et al., 2016) and jaw recorders have been applied to monitor animal locomotion as well as feeding and rumination activities, being used to alert farmers on

behavioural changes associated with diseases, estrus and labor. On the other hand, acoustics methods have been used for monitoring livestock feeding behaviours. Laca et al. (1992) used directional microphones, attached to the forehead of animals, for analysing masticatory sounds in cattle (Galli et al., 2018) and sheep (Galli et al., 2011), as well as for accurate discrimination between feeding and rumination bouts (Chelotti et al., 2020; Vanrell et al., 2018) and for feed intake prediction based on sound energy fluxes (Galli et al., 2018; Laca et al., 2000).

Masticatory sounds are the result of JM-events associated with bites, chews and composite chew-bites (Laca et al., 1992). A grazing bite includes the apprehension and severance of herbage. Chews include the crushing, grinding and processing of herbage during ingestion or rumination. Finally, chew-bites include the combination of chewing and biting in the same JM. Thus, the attributes and statistics of JM-events provide a reliable measure for identification of grazing (include bites, chews and chew-bites) and rumination (chews only) activities and related behaviour events (Chelotti et al., 2016; Milone et al., 2012). Rumination frequency is related to digestion processes and serves as indicator of a suitable rumen health (Sauvant, 2000). The sound properties of bites and chew-bites typically relate to common plant traits and feed structural characteristics (Laca et al., 2000), providing insights of short-term intake rate (Galli et al., 2018) and daily grazing time (Chelotti et al., 2020), two major determinants of the daily feed intake (Hodgson, 1990). A declining rumination time has been shown to be correlated with a declining feed intake (Watt et al., 2015), an acute stressors (Schirmann et al., 2011), an onset of diseases (DeVries et al., 2009), as well as the beginning of estrus (Schirmann et al., 2009) and parturition (Schirmann et al., 2013).

The acquisition of masticatory sounds is the first step of a good acoustic method because registered signals require further processing, analysis and information weighing, and extraction to become useful and insightful for animal monitoring. The analysis of masticatory sounds has been significantly improved in recent years. Milone et al. (2012, 2009) used concepts from automatic speech recognition to develop an algorithm based on hidden Markov models capable of identifying chew, bite, and chew-bite in sheep and cattle. It combines spectral analysis with language-based analysis to detect and classify JM-events. These algorithms achieved an average recognition rate of 80% of successful classifications of bites, chews and chew-bites when tested in controlled experiments lasting a few minutes under good signal-to-noise ratio (SNR) conditions. Chelotti et al. (2016) proposed an alternative algorithm

based on time-domain features of sound signals, achieving similar JM-events recognition rate success at a lower computational cost, compared to Milone et al. (2012), such that it can be implemented in portable microcontroller-based embedded systems (Deniz et al., 2017). More recently, Chelotti et al. (2018) modified and improved this algorithm (Chelotti et al., 2016) using concepts and tools derived from signal processing, pattern recognition and artificial intelligence areas without significantly increasing the computational cost. This algorithm, called Chew Bite Intelligent Algorithm (CBIA), attenuates the effects of time-varying noises and trends, and it achieves a 90% recognition rate. Although CBIA showed good performance for moderate SNR, significant limitations arose when it was employed in farming and husbandry environments, which typically involved louder and time-varying noise and disturbance sources. These negatively affect the detection, features extraction and classification of JM-events since they can not be completely removed from processes internal signals (see Chelotti et al., 2018).

The present work documented and tested the integration of a new set of tools for pattern recognition analysis and artificial intelligence for robust on-line analyses of masticatory sound signals collected from grazing livestock. The main objective is to achieve a more robust detection and classification of JM-events than the CBIA algorithm. Specifically, the new algorithm was especially designed to: i) attenuate distorting effects of environmental noises on masticatory signals likely associated with typical farming conditions and animal handling procedures; and ii) improve the detection and classification of JM-events without significantly increasing the computational cost. The computational cost and a cost-benefit analysis of the new algorithm and CBIA were also evaluated to assess future feasibility for real-time execution in low-cost embedded systems.

The paper is organised as follows: Section 2 analyses the CBIA and presents the new classification features intended to attenuate effects of noises on detection and classification tasks. Then, the proposed algorithm is introduced. Also it is described the acquisition of datasets, the performance measures and the experimental setup used to validate the algorithms. Section 3 shows the comparative results for the proposed algorithm and CBIA. A focused discussion and main conclusions follow in Section 4 and Section 5, respectively.

2. Materials and Methods

In the following section, the operation and weakness of the former CBIA are briefly described. Then, the approach of the proposed algorithm is presented. It is based on the use of instantaneous power signal and it incorporates a novel set of features that are more robust to withstand against likely distortions and interferences produced by environmental noises. The algorithm recognises and discriminates JM-events into four classes (Fig. 1): exclusive rumination-chews (RC - chews related to rumination only), exclusive grazing-chews (GC - chews related to ingestive grazing only), exclusive grazing bites (B - bites taken during grazing) and composite chew-bite (CB - compound chew-bite taken during grazing).

2.1. CBIA review

The CBIA is a real-time pattern recognition system that detects and classifies masticatory sounds produced by ruminants into three JM-event classes (chew (C), bite (B) and chew-bite (CB)) using heuristic features (Chelotti et al., 2018). It is characterised by a sequence of generic stages that allows the processing, description and analysis of the sound signal: conditioning, pre-processing, segmentation, feature extraction and classification. In the **signal conditioning stage** the sound signal is conditioned and filtered to improve its SNR using an adaptive low-pass filter. Then, the pre-processing stage computes the sound envelope, from the filtered signal, and decimates it to reduce the computational cost of following stages. The **segmentation stage** identifies the candidate JM-events (i.e. peaks in the envelope signal) using a time-varying threshold. Once a candidate JM-event is detected, the following **features** are extracted from the filtered sound intensity or from the sound envelope: amplitude, duration, shape index and symmetry of the candidate JM-event. Then, the extracted features feed a classifier (**classification stage**) that assigns each candidate JM-event to one of the three possible JM-event classes. Two stages (segmentation and feature extraction) have been specifically designed for CBIA while the remaining ones (conditioning, pre-processing and classification) can be implemented using any algorithm available in the signal processing and computational intelligence literature (Chelotti et al., 2018). This fact allows to implement distinct combinations of algorithms that bring recognition systems with different computational and performance characteristics. The most adequate combination for real-time operation is the algorithm that uses an adaptive least mean

square filter (LMS) for signal pre-processing and a multilayer perceptron (MLP) for classification, leading to CBIA LMS-MLP variant.

The CBIA showed recognition rates of $\sim 90\%$ under controlled experimental conditions (low noise environment). On the same conditions, other known methods achieved lower rates: CBRTA ($\sim 77\%$) and CBHMM ($\sim 80\%$). Therefore the CBIA will be used as baseline in this work.

A comprehensive study of the effect of environmental noises and disturbances in CBIA has not been carried out. Incoming sound signals usually have noises and slow varying patterns superimposed on the masticatory sounds produced by animals, causing JM-events detection and classification mistakes. The tasks performed by signal conditioning and pre-processing stages leave some noises and disturbances unaffected. They can distort the sound envelope, upsetting the segmentation and feature extraction stages and hampering the overall algorithm performance. Therefore, to improve robustness in a noisy environment, alternative processing strategies and new features should be used.

2.2. Proposed features

The key problem of the former CBIA is the sensitivity to changes in the sound envelope. Any noise recorded together with masticatory sounds distorts the envelope, leading to errors in the detection stage and misinterpretations of the features decreasing the successful JM-events classification rate.

Since noises overlap with the target signal, they cannot be completely removed from the incoming signals by the use of real-time low-computational-cost signal processing techniques. The alternative tested in this work is to combine pre-processing techniques with the use of the envelope of the instantaneous power signal, instead of using the sound envelope to detect JM-events and extract their corresponding features. The instantaneous power is computed as the scalar product of the conditioned incoming signal at a given point of time. Thus, the envelope of the instantaneous power signal works as an expander, increasing the event detection sensibility. It is analogous to an asynchronous demodulation technique, which is frequently used to improve the SNR for small signals in measurement systems (Roden, 1996).

In this work, we propose to extend computation to the following set of temporal features, extracted either from the instantaneous power signal or its envelope:

- **Sign of the envelope slope:** This feature represents the shape of the JM-event and is calculated as the number of times that the envelope slope changes from positive to negative. Typically, the envelope changes from positive to negative once for GC and B, two or more times for CB, and one or more times for RC (Fig. 1, row 3).
- **Accumulated envelope speed:** The speed of the envelope slope (or absolute magnitude of change) is related to the sound intensity and variations. This feature is computed as the cumulative sum of the speed of the envelope slope. It distinguishes low intensity RC from any other grazing event. Also, a GC is differentiated from a B and a CB, which present similar values between them. The Fig. 1, row 4, shows how the accumulated envelope speed increases as a function of time.
- **Duration:** event duration is the time that an JM-event takes place. A typical compound grazing CB has a greater duration than others grazing events. In turn, the duration of a RC is similar to or longer than a GC due to the moisture content of the chewed material (Fig. 1, row 5) (Galli et al., 2020).
- **Symmetry:** is related to the shape of the event and is a measure of the length symmetry of a JM-event computed at peak envelope signal. The value is greater for B than CB, while GC and RC typically present variable values between B and CB (Fig. 1, row 5).
- **Total energy:** is the accumulated value of the instantaneous power signal and it is related to the intensity and duration of the JM-event. The value is higher for CB than B, and B usually have higher value than GC and RC. In turn, RC have a much lower value than any grazing event because the ruminated herbage has already been partially chewed and crushed. The Fig. 1, row 6, shows how the event energy increases as a function of time.

The sign of the envelope slope and duration features have been already used by Chelotti et al. (2018) but in the present work the same features are extracted from a different signal: the instantaneous power signal.

2.3. The CBEBA description

In this section the proposed algorithm called *Chew-Bite Energy Based Algorithm* (CBEBA) is introduced. CBEBA performs three tasks: i) JM-event

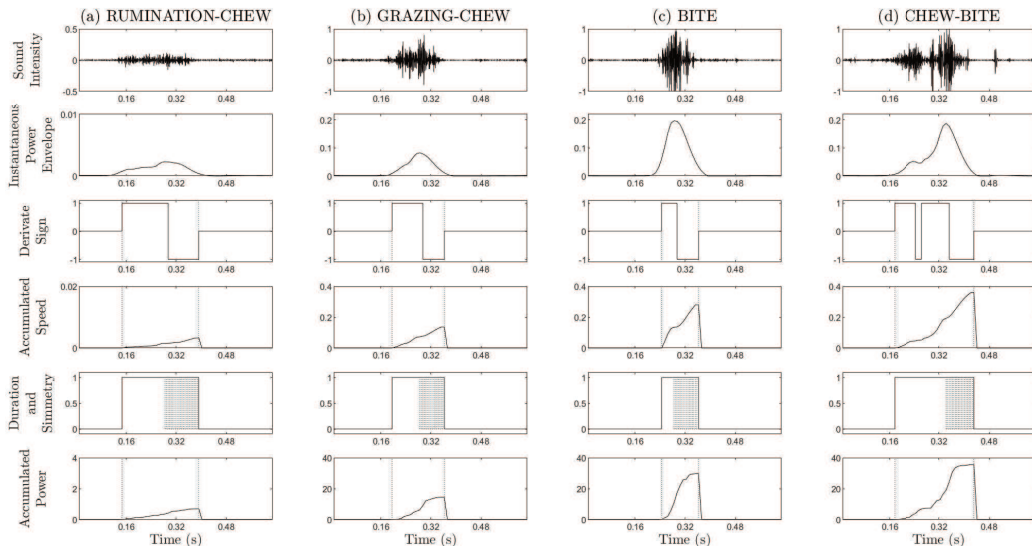


Figure 1: Signals of typical acoustic events produced by jaw movements and their corresponding features for: a) rumination-chew, b) grazing-chew, c) bite and d) chew-bite obtained from the noiseless dataset described in section 2.4. In each row, 1. acoustic signal, 2. instantaneous power envelope, 3. sign of the envelope slope, 4. accumulated envelope slope change rate or speed, 5. duration and symmetry, and 6. accumulated power.

detection, ii) JM-event classification and iii) parameters tuning. JM-event detection is performed by searching for peaks in the instantaneous power signal envelope. Once a possible JM-event has been detected, the features are extracted from the instantaneous power signal and an evaluation is carried out indicating whether the considered event corresponds or not to a true JM. The JM-event classification is carried out using a machine learning technique of low computational cost, to allow future implementation in low-power embedded devices. Finally, parameters tuning is performed. In the time between two consecutive JM-events, the background noise level is estimated and used for the adjustment of internal variables, which significantly improves the algorithm robustness against time-varying conditions.

The online implementation of the algorithm for operation real-time can be internally divided into six stages in a feedback configuration (Fig. 2):

Stage 1 – Signal pre-processing: The input signal usually has noises and disturbances that distort the masticatory sounds produced by animals. Therefore, it is necessary to process the signal in order to remove them. This

task is performed by limiting the signal bandwidth with a second-order Butterworth band-pass filter where the energy spectral density of the JM-events is located. The filtered signal is multiplied by itself to obtain the instantaneous power signal ($p(k)$), where k corresponds to the sampling frequency of the input audio signal. The gain of the acquisition system serves as input to normalize the incoming audio signal for further processing, matching its range with the range of the device where the algorithm is running.

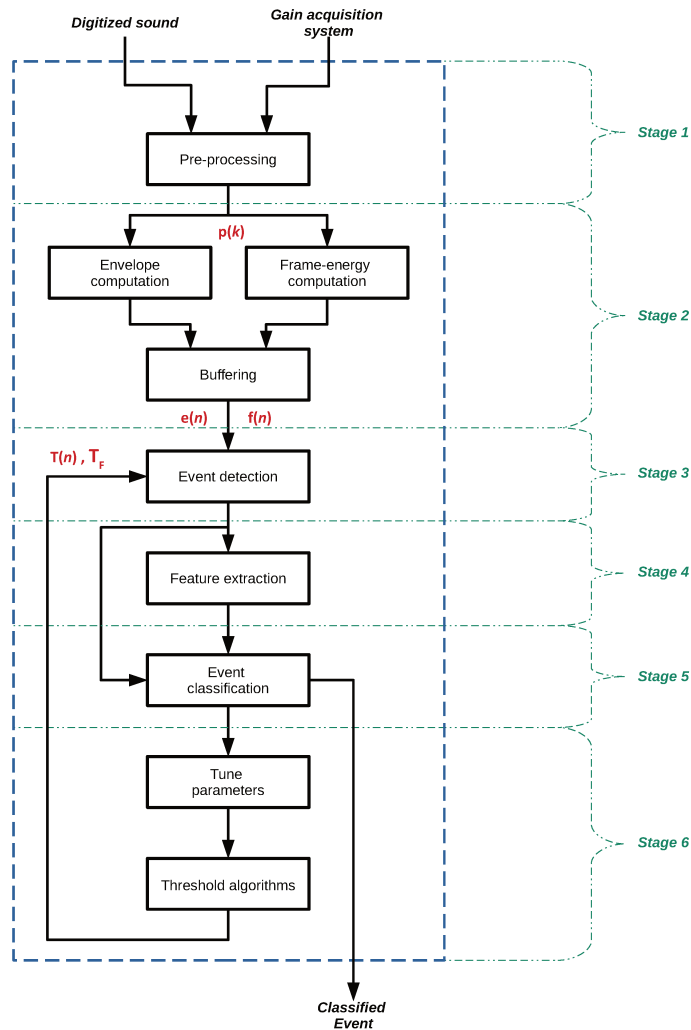


Figure 2: General top-down block diagram of the Chew-Bite Energy Based Algorithm (CBEBA) showing the structure and flow of functions along six operating stages.

Stage 2 – Buffering: The pre-processed instantaneous power signal $p(k)$ follows two parallel paths. In the first path, the envelope computation requires three steps. In the first step, the instantaneous power signal $p(k)$ is filtered with a second-order low-pass Butterworth filter, obtaining the envelope of $p(k)$. In the second step, the low frequency envelope signal is subsampled to reduce the computation cost of the algorithm. Finally, the decimated envelope is saved in the envelope buffer $e(n)$, where n corresponds to the updating frequency of the buffer. In the second path, the frame-energy buffer $f(n)$ stores the energy of $p(k)$ computed by non-overlapping frames, where a “frame” is the cumulative summation of $p(k)$ during a period of time.

Stage 3 – JM-event detection: The buffers $e(n)$ and $f(n)$ contain one second of signals information to be analysed. A peak in $e(n)$ denotes the presence of a possible JM-event. Each peak is detected as a change in the sign of the derivative of the envelope located at the centre position in the $e(n)$. In this case, a second peak is searched beyond the centre position. Each peak must be higher than the value of a given time-varying threshold $T(n)$. If a second peak was found, the middle position of the peaks is centred in $e(n)$.

The time-varying threshold algorithm $T(n)$ is a piecewise linear function that considers both the same anatomical and behavioural characteristics of the animal proposed in the threshold algorithm used in the CBIA and time-varying feeding activities conditions. A description of the parameters involved in the computation of $T(n)$ are available in the supplementary material.

Once a peak has been detected in $e(n)$, the start and end of the candidate JM-event are defined from $f(n)$. The start (n_{START}) and end (n_{END}) of the candidate JM-event correspond to the minimum and maximum n where $f(n) > T_F$ during a hangover period. T_F is the adaptive energy threshold level described in stage 6.

Stage 4 – Features extraction: n_{START} and n_{END} positions of the candidate JM-event are used as boundaries to compute and extract the following features: sign of the envelope slope, accumulated envelope speed, duration, symmetry and total energy, as it was described previously in section 2.2.

Stage 5 – JM-event classification: Once the features have been extracted, the candidate JM-event is analysed to determine if the event should be further classified or not. To be considered a JM, the duration feature must be in a predefined range and the total energy feature must be greater than a certain value of the energy in $f(n)$. Otherwise, the tune threshold algorithm is informed, and the extracted features are discarded. The final classification

step is performed by a MLP.

Stage 6 – Threshold algorithms and parameters tuning: Once the JM-event has been classified, some internal values are extracted from the buffers $e(n)$ and $f(n)$. These values are used to update parameters associated with the time-varying threshold algorithm $T(n)$ and the adaptive energy threshold algorithm T_F . The computation of internal values and parameters related to both algorithms are available in the supplementary material.

2.4. Datasets

Two independent datasets of dairy cow sound records were used to implement and evaluate the proposed algorithms. The first dataset (referred as DS1) was acquired during grazing feeding trials performed at the dairy farm of the W.K. Kellogg Biological Station of Michigan State University, United States, in August 2014. Protocols for animal handling and care were reviewed, approved, and conducted according to the Institutional Animal Care and Use Committee of Michigan State University, as described in Vanrell et al. (2018). Cows were housed and managed on a pasture-based robotic milking system with voluntary cow traffic as described previously in Watt et al. (2015). During six non-consecutive days the foraging behaviour of five lactating multiparous Holstein cows weighing 652 ± 40 kg was continuously monitored. The dairy cows were group grazed on perennial ryegrass (*Lolium perenne*) / white clover (*Trifolium repens*) and orchardgrass (*Dactylis glomerata*) / white clover pastures as part of a larger herd of ~ 140 cows.

The second dataset (referred as DS2) was acquired during grazing feeding trials conducted at the dairy farm of the Campo Experimental J. Villarino of Universidad Nacional de Rosario, Argentina, in October 2014. Project protocols were previously evaluated and approved by the Committee on Ethical Use of Animals for Research of the Universidad Nacional de Rosario. The foraging behaviour of five lactating multiparous Holstein cows weighing 570 ± 40 kg grazing on alfalfa (*Medicago sativa*), fescue (*Festuca arundinacea*) and prairie grass (*Bromus catharticus*) mixed pastures, were continuously monitored during six non-consecutive days. The experimental cows were managed along with a larger dairy herd (150 cows) and milked twice a day (~ 6 am - 6 pm).

The two field studies used the same cow acoustic halters to record masticatory sounds. Each halter included a directional microphone pressed to the forehead of each animal connected to a digital recorder (Sony Digital ICD-PX312, Sony, San Diego, CA, USA). Microphones were held and protected

by an elastic band attached to halters, avoiding microphone movement, friction and scratches. On a given recording day, the five acoustic halters were randomly assigned to each of five cows and were rotated across cows and sampling days. This design allowed to control for any likely device effect (i.e. all cows were tested with all halters). All recordings were saved in waveform audio (WAV) file format, using 16-bit resolution and a sampling rate of 44.1 kHz. The DS1 dataset is composed of eighteen short-term records of ingestive grazing sounds (*grazing-segments*) of 150 s for each one, and eighteen short-term segments of rumination sounds (*rumination-segments*) of 75 s for each one. The selected recordings were captured with minimum or no influence of external environmental noise; thus, represented ideal acoustic conditions for algorithm testing. A typical *grazing-segment* contains more than 170 JM-events whereas a typical *rumination-segment* contains more than 65 JM-events, which results in over 4,200 JM-events in DS1 (32% RC, 18% GC, 9% B and 41% CB). The DS2 dataset contains more than 5,200 JM-events (28% RC, 28% GC, 11% B and 33% CB) collected from sixteen segments of grazing activity and sixteen segments of rumination activity of same duration as DS1. The DS2 recordings were purposely captured in a free-ranging environment; thus, recordings included acoustic signals corrupted by typical external noises, such as blowing wind, cow calls, cow steps, vehicle engines and human voices, among others, and were used to evaluate the proposed algorithm under adverse acoustic conditions.

All segments were labelled aurally by two experts in cows foraging behaviour and with prior experience to identify and classify individual JM-events associated with grazing and rumination. The supervised labelling was done by one of the experts, with results double inspected and checked by the second expert. In most of the cases experts largely agreed with the labelling process, but when there was disagreement, both experts worked together to convey final JM-events classification decisions. The same procedure was applied successfully in Chelotti et al. (2018, 2016, 2020); Vanrell et al. (2018). This supervised JM-event labelling was used as control reference for the purposes of comparing and testing the performance of the algorithm, and to evaluate the quality of sound signals.

The noise effect of environmental sounds on grazing and rumination signals was measured using the JM modulation index (MI_{JM}) and the SNR. The MI_{JM} denotes the suitability of signals for recognition of JM-events in a given segment and it is defined as the ratio between the mean sound intensity during JM-events to the mean sound intensity during the pauses

inter-JM-events. The SNR is used to indicate how the proper JM-events classification is affected by noise. A detailed description of the computation of quality factors is given in Appendix A. Both factors were computed for each segment and then averaged per foraging activity (rumination or grazing) in each dataset. Reported values are given in Table 1.

Table 1: Average value for the noise effect on acoustic signals of foraging activities collected in the DS1 and DS2 datasets

	Rumination		Grazing	
	DS1	DS2	DS1	DS2
$MI_{JM}(\%)$	37.9	21.5	78.1	45.8
SNR (dB)	5.5	2.7	11.7	9.8

2.5. Performance measures

To evaluate the algorithm, comparisons between JM-events recognised by the algorithm vs. JM-events classified aurally were mutually synchronised and compared. This temporal synchronisation and comparison was performed using the HResults tool, which is available as part of the HTK speech analysis toolkit software (HTK 3.4.1, Cambridge University, UK). JM-event synchronisation, matching algorithm recognised and aurally classified JM-event sequences was conducted by an optimal string match using a dynamic programming-based string alignment procedure (Young et al., 2002). The recognition statistics outputs were used to evaluate the detection and classification performance metrics.

The JM-event detection only considers the existence or not of an JM-event, ignoring their corresponding class. The detection performance is collectively affected by the number of correct JM-events detected (true positives), the number of undetected JM-events (false negatives) and the number of incorrect JM-events detected (false positives). The effectiveness of the JM-event detector is reported through metrics for *precision*, *recall* and *F1 – score*.

The JM-event classification considers the type of JM-events detected. The correct number of previously detected JM-events were used to compute the number of true positives (tp_i), the number of true negatives (tn_i), the number of false positives (fp_i), and the number of false negatives (fn_i) JM-events for each JM-event class i , respectively. Classification performances were averaged by JM-event class (macro-averaging) and reported

as F1-score ($F1 - score_M$), arithmetic precision ($precision_M$) and arithmetic recall ($recall_M$), (Sokolova and Lapalme, 2009), and their equivalent geometric precision ($precision_G$) and geometric recall ($recall_G$) (Ballabio et al., 2018). The dispersion of the classification performance per JM-event class tends to affect more negatively geometric averages than arithmetic averages.

2.6. Experimental setup

The proposed CBEBA was coded and the experiments were both carried out using Matlab R2019b (MathWorks, Natick, MA, USA) in a personal computer with an Intel Core i7-4790 3.6 GHz with 16 GB of RAM. The MLP network classifier was implemented using the Levenberg-Marquardt backpropagation algorithm optimising the mean square error (Demuth et al., 2014). Inputs were normalised and output labels binarised to match MLP output. MLPs classifiers with five input neurons (number of input features), four output neurons (number of output labels corresponding to JM-events) and one hidden layer were used. The hyperparameters correspond to the learning rate (from 0.1 to 0.001) and numbers of neurons in the hidden layer (from three to seven) were optimised using the grid search method.

Segments in each dataset were split into two parts. The first part included one-third of the segments and served to tune the internal parameters of the algorithm. The second part included the remaining two-thirds of the segments and were used to validate the algorithm (Fig. 3).

In the first part, a train/test split scheme was performed to tune the internal parameters of CBEBA using a grid search method. Four grazing segments and four rumination segments were used for training the model (light red background segments in Fig. 3). A MLP was trained for each internal parameters configuration optimising its hyperparameters. Additional synthetic JM-events for the minority class (bites) were generated using the adaptive synthetic (ADASYN) algorithm. This step was applied to control for JM-events class imbalances during the algorithm training phase (He et al., 2008). Two grazing segments and two rumination segments were used for evaluating the performance of the system for each configuration (dark red background segments in Fig. 3). The optimal internal parameters configuration that maximised the condition $F1 - score_M$ was founded varying the following parameters: (i) lower (f_{lower}) and (ii) upper (f_{upper}) frequencies during pre-processing stage, (iii) the cut-off frequency (f_{cut}) involved during envelope computation, and (iv) the envelope sub-sampling frequency (S_s). The corresponding optimal values for both datasets were $f_{lower} = 175$ Hz,

$f_{upper} = 900$ Hz, $f_{cut} = 9$ Hz and $S_s = 150$ Hz. The variation range of the parameters and their relationships with JM-event features computations are available in the supplementary material.

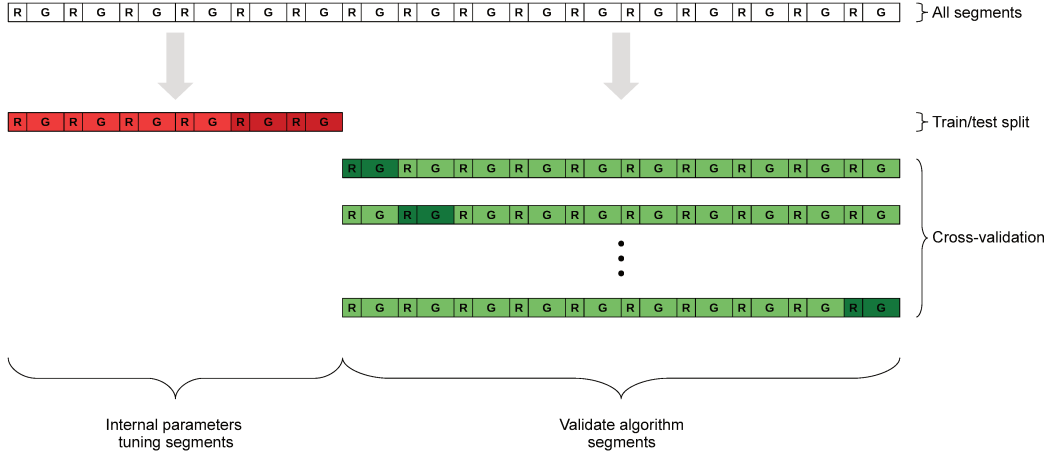


Figure 3: Configuration of the segment fragmentation used to perform the experiments in each dataset. R labels correspond to 75 s *rumination-segments* and G labels correspond to 150 s *grazing-segments*.

The optimal internal parameters obtained in the first part of the split dataset were used in the second part for the final validation of the algorithm in a cross-validation scheme with the corresponding part of the data (Fig. 3). Since grazing and rumination segments are related to different JM-events class, in each iteration, one grazing segment and one rumination segment were selected for the purpose of testing (dark green background segments in Fig. 3), while the remaining segments were used to train the model (light green background segments in Fig. 3). The two segments chosen for testing were only used once and they were not used for testing in subsequent iterations. During the algorithm training phase, synthetic bites were generated with the ADASYN algorithm (He et al., 2008). Performance metrics informed were obtained averaging all iterations. Cross-validation was instrumented as a comprehensive statistical method for evaluation of CBEBA.

Evaluations were conducted according to the following four dataset combinations: i) training with DS1 and testing with DS1 (referred as DS11); ii) training with DS1 and testing with DS2 (referred as DS12); iii) training with DS2 and testing with DS1 (referred as DS21); and iv) training with DS2 and testing with DS2 (referred as DS22).

3. Results

3.1. Qualitative results

A qualitative analysis of the effect on JM-event class separation of the proposed features was performed using a t-distributed stochastic neighbor embedding (t-SNE) analysis (van der Maaten and Hinton, 2008). Fig. 4 shows the result of the dimensionality reduction analysis applied to both sets of JM-events features proposed in this work (Fig. 4a and Fig. 4b) and those sets of features previously applied in the former CBIA (Fig. 4c and Fig. 4d). These figures show the separation of JM-events in four classes for the proposed algorithm (B, CB, RC and GC) and three for CBIA (B, CB and C, where RC and GC are joined as C). Fig. 4a and Fig. 4c show the results for both algorithms in noiseless conditions using the DS1 dataset. B and CB are grouped into two detached clusters with few mixed data points between them. C are also grouped into two clear clusters for both sets of features. In the CBIA case (Fig. 4c) both clusters correspond to the same JM-event class, while in the CBEBA case (Fig. 4a) each cluster corresponds to a different JM-event class. The clusters obtained with the proposed set of features are detached with few mixed data points between them (Fig. 4a). RC and GC are clearly separated, which supports the hypothesis for differences in acoustic properties associated with chewing of fresh (GC) and preprocessed (RC) material. Only B are mixed with both C and CB. CB mixed with GC are negligible. On the other hand, the clusters obtained with the CBIA set of features are detached for some groups and they have mixed data points between them (Fig. 4c). The C cluster is blended with the B cluster. CB form two separated clusters with one of clusters slightly mixed with C events. C form two isolated clusters, which may be explained by the presence of GC and RC (Fig. 4c).

Fig. 4b and Fig. 4d show the results in noisy conditions using the DS2 dataset. Similar separation and characteristics of clusters occurs for the proposed set of features with an increment in clusters overlapping (Fig. 4b). For the CBIA set of features, the clusters are less dense with a considerable number of mixed data points. In this sense, C are grouped in two clusters. One of them presents low density and is mixed with data points corresponding to B and CB. The cluster corresponding to B is low density grouped and mixed with C and CB points. A similar situation occurs with the cluster of CB (Fig. 4d).

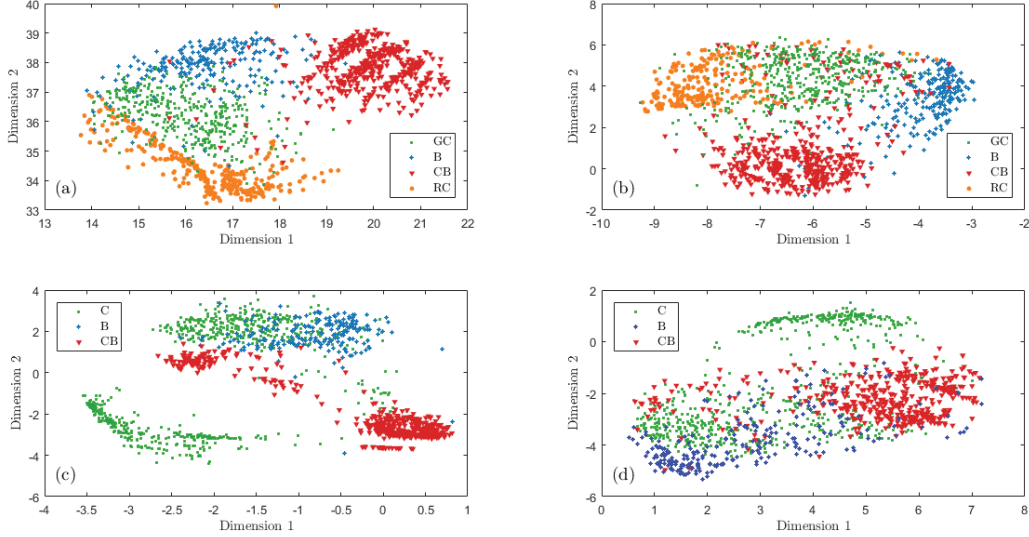


Figure 4: t-SNE analysis of the set of JM-events features used in the proposed CBEBA (panels a and b) and those used in previous CBIA (panels c and d) performed against DS1 noiseless dataset (panels a and c) or DS2 noisy dataset (panels b and d).

3.2. Algorithm performance metrics

The proposed CBEBA was evaluated in each of the four possible dataset combinations, and their performance metrics are summarised in Table 2. The detection reached an excellent performance as shown by a *recall* of $\sim 98.0\%$ in all dataset combinations and a *precision* of $\sim 99.0\%$ for DS11 and DS21. However, *precision* decreased to values of $\sim 96.0\%$ for DS12 and DS22. A similar behaviour was seen for *F1 – score*. The classification performance was the highest and more consistent for DS11 as shown by all classification metrics with higher mean and lower standard deviation. Likewise, classifications for DS22 showed smaller deviations in the mean value ($< 2\%$) and standard deviations ($< 3\%$) than DS11. Conversely, DS12 had lower overall classification performance than DS11. This deterioration was relatively small ($< 4\%$) for *recall_M*, *recall_G* and *F1 – score_M*; and greater ($> 5\%$) for *precision_M* and *precision_G*, respectively. The standard deviation for the classification metrics of DS12 doubled the standard deviation observed for DS11 due to the influence of environmental noise. The classification performance for DS21 had lower *recall_M* (-3.8%), *F1 – score_M* (-1.5%) and *recall_G* (-6.5%), with slighter improvements for *precision_M* (0.9%) and *precision_G* (1.2%) than those corresponding to DS22. The standard deviation for the

Table 2: Comparative metrics for CBEBA detection and classification performance (mean \pm SD; %) for different combinations of training and testing dataset variants.

		DS11	DS12	DS21	DS22
Detection	<i>Recall</i>	98,4 \pm 1,0	98,0 \pm 0,9	97,9 \pm 1,3	98,4 \pm 1,2
	<i>Precision</i>	99,2 \pm 0,5	95,8 \pm 2,1	98,7 \pm 0,8	96,1 \pm 2,6
	<i>F1 – score</i>	98,8 \pm 0,6	96,9 \pm 1,4	98,3 \pm 0,9	97,2 \pm 1,8
Classification	<i>Recall_M</i>	92,0 \pm 2,6	89,8 \pm 4,3	87,5 \pm 3,8	91,3 \pm 3,8
	<i>Precision_M</i>	92,2 \pm 2,8	87,0 \pm 5,5	91,6 \pm 3,1	90,7 \pm 5,1
	<i>F1 – score_M</i>	92,1 \pm 2,4	88,4 \pm 4,7	89,5 \pm 3,2	91,0 \pm 4,2
	<i>Recall_G</i>	88,9 \pm 3,7	87,9 \pm 4,8	81,8 \pm 5,7	88,3 \pm 4,3
	<i>Precision_G</i>	89,3 \pm 4,3	81,5 \pm 8,6	88,9 \pm 4,7	87,7 \pm 6,6

classification metrics of DS21 were smaller than those for DS22 due to reductions of environmental noise associated with the sound signal of DS1.

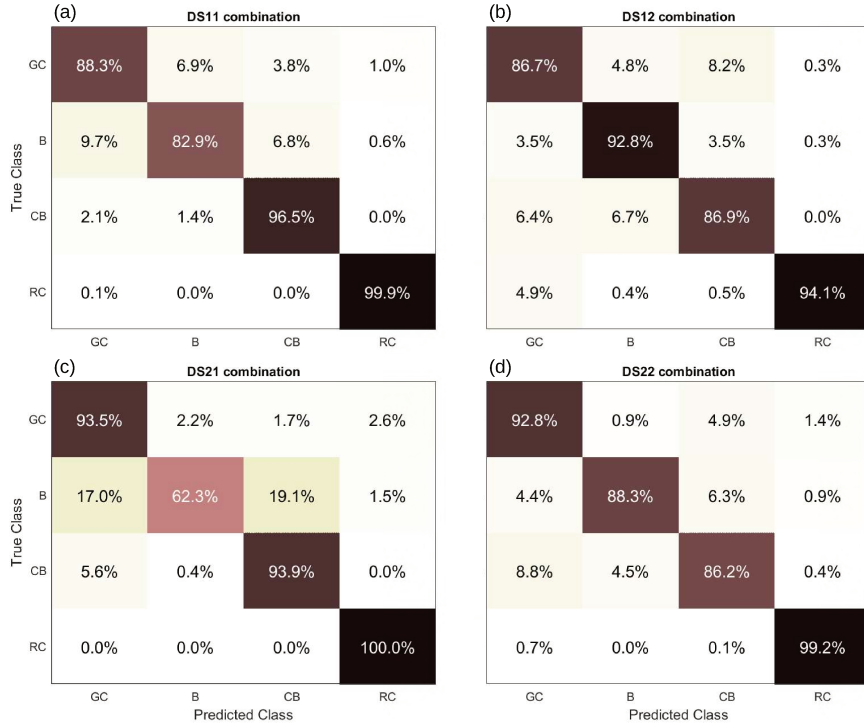


Figure 5: Confusion matrices for classification of grazing-chews (GC), bites (B), chew-bites (CB) and rumination-chews (RC) by CBEBA in the a) DS11, b) DS12, c) DS21, and d) DS22 dataset combinations, respectively.

The recognition results across the four dataset combinations, obtained by accumulating the results of each segment used for testing, are presented in the confusion matrices in Fig. 5. Each row represents the distribution of true JM-events over the event class into which they were classified. The rumination-chew events were almost perfectly classified. The grazing-chew and chew-bite events reached recognition rates higher than 86.2% in all dataset combinations. Grazing-chews and chew-bites were better classified in DS1 (DS11 and DS21) than in DS2 (DS12 and DS22). There was lower confusion between datasets in DS1 (3.8% GC classified as CB and 2.1% CB classified as GC in DS11, and 1.7% GC classified as CB and 5.6% CB classified as GC in DS21) than in DS2 (8.2% GC classified as CB and 6.4% CB classified as GC in DS12, and 4.9% GC classified as CB and 8.8% GC classified as GC in the DS22). Particularly, GC and B were the most affected JM-events in mismatch conditions. The recognition rate of GC increased, whereas the recognition rate of B decreased when CBEBA was trained using the noisy dataset (DS2).

3.3. Comparisons between CBEBA and CBIA

A direct comparison between the CBEBA and the former CBIA is not possible because both algorithms are based on a different classification of JM-events. Therefore, one alternative is to combine the grazing and rumination chews in CBEBA into a single chew (C) class in order to have the same type and number of JM-event classes as in CBIA (C, B and CB classes, respectively).

The comparative performance for detection and classification of JM-events between CBIA and CBEBA are presented in Table 3. The CBIA was configured, trained and tested in the same way as the CBEBA. Performance metrics discrepancy between CBEBA and CBIA were evaluated to be statistically significant ($p < 0.05$), using a Wilcoxon signed-rank test (Wilcoxon, 1945).

CBEBA showed better overall performances than the former CBIA. Both algorithms showed similar *recall* for detection in all dataset combinations ($p > 0.05$). However, the *precision* and *F1 – score* indicated a better overall JM-event detection performance for CBEBA than CBIA, respectively ($p < 0.05$). In noisy testing environments (DS12 and DS22), CBEBA achieved better metrics for *precision* ($\sim 96\%$ vs. $\sim 92\%$) and *F1 – score* ($\sim 97\%$ vs. $\sim 95\%$) for detection than CBIA. This discrepancy between metrics is likely due to the apparent effects of noises on the *precision* and

Table 3: Comparative performance metrics (mean \pm SD; %) for CBIA and CBEBA across the four dataset combinations. Metric values with green background within same rows and dataset combinations differ significantly, whereas metric values with pink background within same rows and dataset combinations show no significant difference ($p < 0.05$; Wilcoxon signed-rank test).

		DS11		DS12		DS21		DS22	
		CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA
Classification	<i>Recall</i>	98.8 \pm 0.7	98.4 \pm 1.0	98.5 \pm 0.7	98.0 \pm 0.9	98.0 \pm 0.9	97.9 \pm 1.3	98.7 \pm 0.6	98.4 \pm 1.2
	<i>Precision</i>	98.8 \pm 0.8	99.2 \pm 0.5	91.7 \pm 2.8	95.8 \pm 2.1	98.0 \pm 0.7	98.7 \pm 0.8	91.9 \pm 2.6	96.1 \pm 2.6
	<i>F1 - score</i>	98.8 \pm 0.5	98.8 \pm 0.6	95.0 \pm 1.7	96.9 \pm 1.4	97.8 \pm 0.5	98.3 \pm 0.9	95.1 \pm 1.5	97.2 \pm 1.8
	<i>Recall_M</i>	91.7 \pm 3.3	91.8 \pm 3.4	72.9 \pm 5.0	90.5 \pm 4.6	82.0 \pm 6.8	92.9 \pm 2.6	86.7 \pm 6.2	90.0 \pm 4.2
	<i>Precision_M</i>	93.7 \pm 2.4	92.1 \pm 3.5	79.6 \pm 6.7	85.9 \pm 5.7	83.0 \pm 4.9	85.1 \pm 4.8	86.8 \pm 5.5	90.2 \pm 5.4
	<i>F1 - score_M</i>	92.5 \pm 1.8	91.9 \pm 2.4	77.5 \pm 5.2	88.0 \pm 4.4	82.9 \pm 5.1	88.8 \pm 3.1	87.3 \pm 5.0	90.0 \pm 4.3
	<i>Recall_G</i>	90.8 \pm 3.9	91.5 \pm 3.8	68.0 \pm 7.6	90.2 \pm 5.1	80.2 \pm 7.8	92.7 \pm 2.8	85.3 \pm 8.2	89.5 \pm 4.7
	<i>Precision_G</i>	93.5 \pm 2.8	91.6 \pm 4.0	78.5 \pm 7.6	84.3 \pm 7.3	81.6 \pm 5.8	83.3 \pm 5.7	85.7 \pm 7.3	89.6 \pm 6.0

F1 - score when each algorithm was trained with a noiseless dataset. This fact is noteworthy for DS12 compared with DS11, where the performance deterioration for CBEBA was roughly half of that observed for CBIA.

The JM-event classification showed similar results. Both algorithms had similar performance metrics for the noiseless train/test dataset combination (DS11), but CBEBA had better classification performance (high mean and low standard deviation of metrics) than CBIA in the presence of noises. Compared to DS11, the performance metrics for JM-event classification for DS12 declined, but the decline for CBEBA was less than half of that observed for CBIA. For the cases of noisy training conditions (DS21 and DS22), CBEBA achieved better overall metrics for JM-event classification performance than CBIA, respectively, indicating a greater sensitivity of CBIA to environmental noise (Table 1). In addition, the difference between *recall_M* and *recall_G* was smaller for CBEBA than CBIA in the same dataset combination, supporting greater dispersion in the recall per JM-event class for CBIA than CBEBA. The same situation occurred for *precision_M* and *precision_G* in DS22, further supporting the proposed set of JM-events features for robust detection and classification of broad sets of JM-event classes using CBEBA.

The confusion matrices in Fig. 6 and Fig. 7 show the JM-events recognition rates obtained by CBEBA and CBIA. Fig. 6 shows comparative results for JM-events classification in noiseless data (DS11) and noisy data (DS12) obtained using noiseless data for training. A similar situation obtained using noisy data for training (DS21 and DS22) is shown in Fig. 7. The first row in figures (Fig. 6a and Fig. 6b, and Fig. 7a and Fig. 7b) show the results for CBEBA, whereas the second row in figures (Fig. 6c and Fig. 6d, and Fig. 7c

and Fig. 7d) show the results for CBIA.

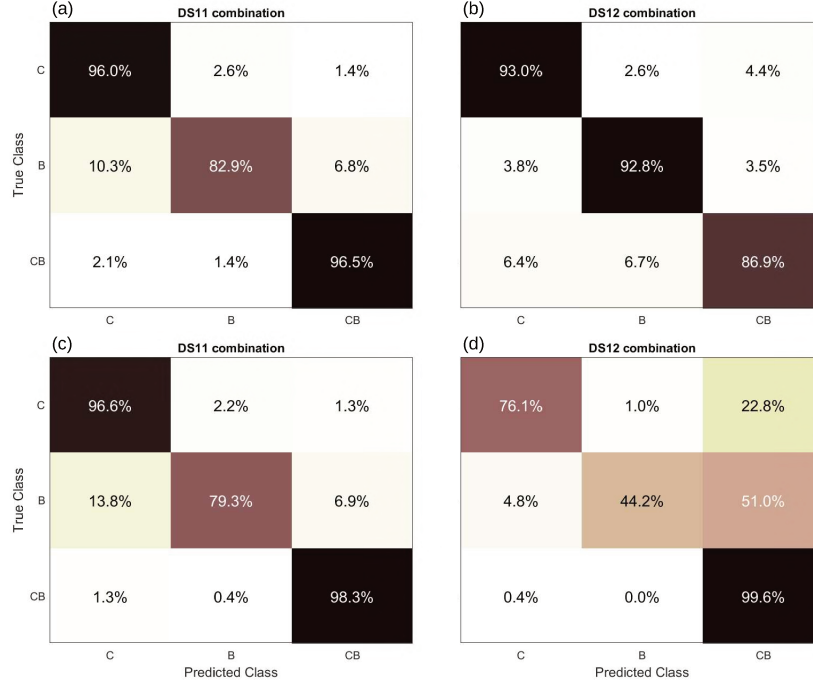


Figure 6: Confusion matrices of the DS11 and DS12 dataset combinations showing recognition rates for chew (C), bite (B) and chew-bite (CB) events classified respectively by the CBEBA (panels a and b) and CBIA (panels c and d) algorithms.

The confusion matrices for DS11 in Fig. 6 showed a slight improvement in the recognition of bites for CBEBA vs. CBIA. Compared to findings shown for CBEBA in DS11, confusion matrices for CBEBA in DS12 showed a similar average JM-events recognition rate (91.8% vs. 90.9%, respectively) with a 9.6% deterioration for recognition of chew-bites and 9.9% improvement for recognition of bites. Compared to results for CBIA in DS11, findings for CBIA in DS12 showed a clear decline in average JM-events recognition rate (91.4% vs. 73.3%, respectively), due to a 20.7% decline in recognition of chews, a 35.1% decline in recognition of bites, and only a 1.3% improvement in recognition of chews-bites as a consequence of the great confusion of chews and bites with chew-bites.

The confusion matrices for DS21 in Fig. 7 showed a slightly better average JM-events recognition rate for CBEBA (84.9%) than CBIA (83.2%). The confusion matrix for DS21 showed that CBEBA correctly recognised more

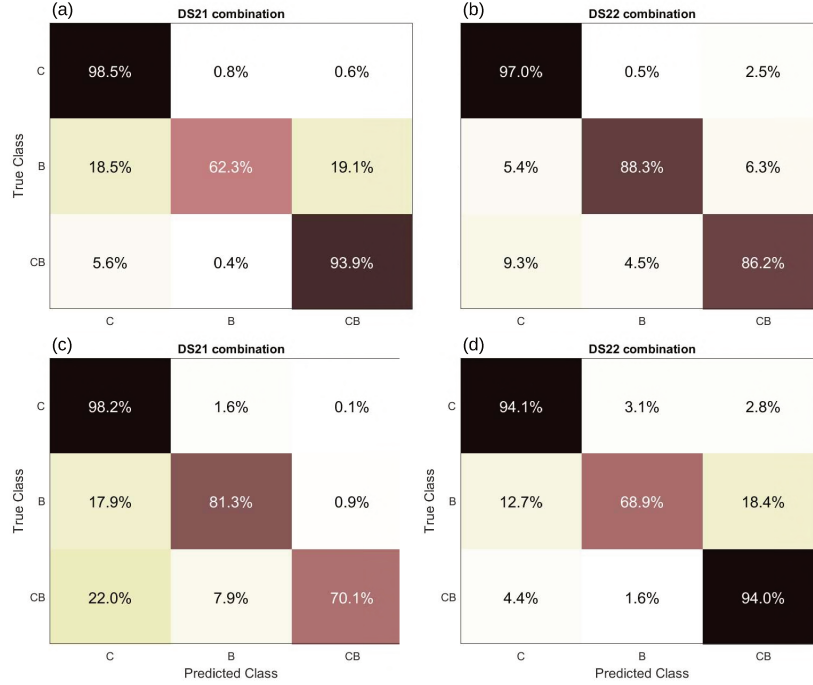


Figure 7: Confusion matrices of the DS21 and DS22 dataset combinations showing recognition rates for chew (C), bite (B) and chew-bite (CB) events classified respectively by the CBEBA (panels a and b) and CBIA (panels c and d) algorithms.

than 93.9% of combined grazing and rumination chews and chew-bites, but only 62.3% of bites. Conversely, CBIA achieved a better recognition of bites (81.3%) at the expense of a lower recognition of chew-bites (70.1%). Furthermore, CBIA presented a much lower recognition of bites in DS22 than CBEBA (68.9% vs. 88.3%, respectively). However, it is noteworthy that regarding DS22 the average JM-events recognition rate for CBEBA declined 5.6% compared to a decline of 2.4% for CBIA in DS21, with a 26% decline in the recognition of bites and despite a 1.5% and 7.7% improvement in the recognition of chews and chew-bites for CBEBA, respectively. Compared to results for CBIA in DS22, findings for CBIA in DS21 showed lower average JM-events recognition rate (85.6% vs. 83.2%, respectively) due to a 23.9% decline in the recognition of chew-bites, and despite the 4.1% and 12.4% improvement in the recognition of chews and bites, respectively.

Table 4 shows the computational cost of CBEBA and CBIA, both ran on a similar MLP neural network for JM-event classification, expressed in

terms of the number of operations per second (ops/s). A detailed description of the operations and assumptions is given in Appendix B. The overall computational cost of CBEBA was 11% higher than that for CBIA, mainly due to the additional costs associated with the signal conditioning, JM-event classification and updating of internal algorithm parameter tasks.

Table 4: Partitioned computational costs for the CBEBA and CBIA expressed in operations per second.

Algorithm computational stage	CBIA ¹ [ops/s]	CBEBA ² [ops/s]
Signal conditioning and pre-processing	10000	16000
Event detection and feature extraction	28800	26110
Event classification	160	690
Internal parameters update	-	378
Global cost	38960	43178

¹ Considering k=2 kHz, n=100 Hz and a MLP classifier.

² Considering k=2 kHz, n=150 Hz and a MLP classifier.

4. Discussion

Previous deployment of acoustic telemetry to monitor livestock foraging behaviour showed a promising future for the technique, but a major bottleneck has been the need for more appropriate and robust analytical procedures to deal with issues associated with environmental noise. The present study builds upon a previous methodology (CBIA) to analyse JM of grazing cattle and incorporates analytics of novel sound signal features combined with machine learning techniques for a better detection and classification of JM-events in noisy environments.

A major advantage for the proposed CBEBA has been the expansion in detection and classification of JM-event classes. While previous attempts (Chelotti et al., 2018, 2016; Deniz et al., 2017; Milone et al., 2012) did not differentiated chews among those associated with ingestive grazing and those for rumination, a qualitative analysis of the proposed and previous (CBIA) set of JM-events features showed that chew events are always clustered into two groups, independently if we explicitly differentiate chews associated with ingestive grazing and rumination (the proposed set of features) or not (the CBIA set of features) (Fig. 4). This organisation of the data indicates that

this desegregation of the chew events cluster is possible and could be proposed to improve the recognition capabilities of the algorithms. Thus, the present CBEBA method successfully classified JM-events into exclusive bites, grazing-chews, composite chew-bites and exclusive rumination-chews. Further, the computation of the instantaneous signal power for determination of JM-event energy as a new discriminative feature allowed a better discrimination of chews associated with ingestive grazing and rumination. A plausible explanation could be related to differences in the water content of grazed and ruminated forages. Moisture content can be classified into internal or external, depending on whether it belongs or not to cell wall contents. Likewise, as masticatory sounds produced by chewing are partly due to a rupture of cell walls and extrusion of internal water and contents, we hypothesise that grazing-chews may have produced sound JM-events of greater total energy than rumination chews because the water and contents of intact cells in the ingested forage material were greater than that for the partially processed and degraded material regurgitation during rumination (Galli et al., 2020).

The suitability and comparative ability of CBEBA for detection and classification of JM-events were evaluated using the same and different dataset combinations for the algorithm training and testing. This approach involved datasets acquired in a noisy and noiseless environment, thus allowing for direct evaluating of noise effects and the sensibility to mismatch condition. CBEBA presents good overall results for detection and classification of JM-events. In all training and testing dataset combinations, including noisy datasets, the *recall* for detection was close to 98%, which indicates a low rate of undetected JM-events. The *precision* for detection reached slightly better values for test with noiseless data (DS11 and DS21) than for test with noisy data (DS12 and DS22) ($\sim 99\%$ vs. $\sim 96\%$), which was associated with a greater number of false JM-events detected when testing under noisy scenarios. This finding supports promising applications of CBEBA for deployment in noisy environments. Furthermore, all classification metrics for DS22 decreased less than 1.6% compared to the best noiseless scenario (DS11) (Table 3) and a recognition rate higher than 82% for all JM-events (with the particularity that rumination-chews were almost perfectly classified) was observed in DS11 and DS22 (Fig. 5), suggesting that the features choices are robust in a noisy environment.

The CBEBA performance using different training data was evaluated. Better results in all classification metrics were observed when CBEBA, tested with noiseless data, was trained using noiseless data (DS11) than those using

noisy data (DS21). Further, similar JM recognition rates of grazing-chews, chew-bites and rumination-chews were observed, but the JM recognition rate of bites decreased considerably to 62.3% in DS21 (Fig. 5). This suggests that the additional information provided to algorithm training by noise had a negative influence on noiseless JM-events classifications success. On the other hand, CBEBA, tested with noisy data, showed a decrease in all classification metrics when it was trained using noiseless data (DS12) compared to those using noisy data (DS22). A plausible explanation could be related to the fact that the noise provided additional useful information during the algorithm training for correct noisy JM-events classifications.

Additionally, changing either the sampling frequency of the input audio (fulfilling the Nyquist–Shannon sampling theorem) or increasing the word-length of the digitised input audio above eight bits had a negligible variation (<0.5%) in all performance metrics for CBEBA, and consistent with previous work by Chelotti et al. (2016). The computational cost of CBEBA was 11% higher than CBIA (43178 ops/s vs 38960 ops/s). Thus, the low computational cost, the low input sampling frequency, and the reduced input word-length representation required by CBEBA make it suitable for use as part of portable components for embedded devices. Microcontroller-based embedded systems can reduce their power consumption working at low clock frequency. This fact allows the devices to operate for extended periods of time. In the case that CBEBA would like to be implemented in a microcontroller to operate in real-time, the amount of processing operations associated to registers and memory access would be taken into account. If it is considered that these operations are approximately three times higher than the previously 43178 arithmetic and logic operations, the total number of processing operations for the microcontroller would be 172,712 ops/s. Furthermore, the execution time depends on the architecture and the operating frequency of the microcontroller. For example, using an ARM Cortex-M4 microcontroller (ATSAM4LS2A, Microchip Technology Inc., Chandler, AZ, USA) operating at 1 MHz and considering that performing an operation may require up to 4 clock cycles, CBEBA would require approximately 173 ms. In the case of CBIA, the execution time would be approximately 156 ms following the same analysis. The idle time available until the next JM-event could be used to perform peripheral management or enter in a low power mode.

Accurate measurement of herbage intake is critical to advance knowledge on foraging mechanisms that can result in a more efficient production and

utilization of herbage resources along with a more profitable and affordable animal production. According to pioneer acoustic works (Laca et al., 2000) and earlier applications of the technique with sheep (Galli et al., 2011) and dairy cattle (Galli et al., 2018), the sound energy enclosed in masticatory (chewing and biting) events serves as robust predictor of herbage intake and could be scaled successfully to monitor changes in animal production, health or fitness, long-term. Therefore, the automated computation of sound energy as a novel analytical property in CBEBA is highly relevant and very appealing. This new property suggests a new opportunity to automate CBEBA to routinely monitor and assess use livestock foraging behaviour to manage feed intake (Galli et al., 2018).

5. Conclusions

In this study, the CBEBA was tested as robust alternative for acoustic detection and classification of JM of grazing cows, typically involving noisy environments. The algorithm integrated six generic stages and incorporated computations based on the instantaneous power signal rather than on the intensity signal to improve the JM-event detection sensibility. A novel and extended set of robust features are extracted and combined using machine learning techniques to recognise JM-events. Specific performance metrics for pattern recognition were used to evaluate the results of CBEBA in noisy, noiseless, and crossed scenarios. Overall, CBEBA showed good results, indicating adequacy and scalability of the algorithm for acoustic analytics in broad sets of acoustic scenarios.

Specifically, notable results were obtained in noisy and crossed algorithm training and testing scenarios outperforming the algorithm from the state-of-the-art. To the best of our knowledge this is the first algorithm that is able to distinguish chew events associated with ingestive and rumination activities in grazing cows while dealing accordingly with augmented noises typical of grazing environments. The increase in the computational cost of CBEBA compared to CBIA remained in the same order of magnitude, suggesting feasibility for efficient implementation in a low-cost embedded device for real-time operation.

Future work must include evaluations of the algorithm in more challenging animal husbandry scenarios such as dairy cattle barns and beef cattle feed yards exhibiting louder and more varied array noises as plausible sources of signal interference and degradation. Likewise, future work is needed to scale

up uses of CBEBA as part of low cost online applications to monitor animal activities with regard to feed intake, health, breeding, parturition or production. Finally, the recognition performance of the algorithm could be improved by including more complex processing techniques such as using a deep learning approach combining the JM features extraction and JM classification at the expense of an increased computational cost.

Acknowledgments

This work has been funded by Universidad Nacional del Litoral, PACT CAID 2011–525, Universidad Nacional de Rosario, projects 2013-AGR216, 2016-AGR266 and 80020180300053UR, Agencia Santafesina de Ciencia, Tecnología e Innovación (ASACTEI), project IO-2018–00082, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), project 2017-PUE-sinc(i). Also, this work was funded under the USDA-NIFA MICL0222 and MICL0406 projects, and direct support from AgBioResearch-MSU. Authors would like to thank the dedication and perceptive help by the KBS Robotic Dairy Farm staff, and Campo Experimental J. Villarino Dairy Farm staff for their assistance and support during the completion of this study.

AUTHORSHIP STATEMENT

LSMR, JRG, LLG, SAU, HLR and AMP participated in conceptualization; LSMR participated in software stage; LSMR, JOC, JRG and AMP participated in the data curation; LSMR and JOC participated in the formal analysis; LSMR and JOC participated in investigation stage; LSMR, JOC and SRV participated in methodology, validation and visualization stages; JRG, LLG, SAU, SRV and HLR participated in the funding acquisition; LSMR, LLG and HLR participated in project administration; LSMR, JOC, SRV, SAU and LLG contributed to the writing and reviewing of the original draft; All the authors reviewed and approved the manuscript.

Appendix A. Signal quality factors

The quality of the signals available in the two datasets depends on both, the JM and the noise levels. The noise level in corrupted signals can affect the recognition and classification performance.

Given that JM during rumination and grazing feeding activities are performed rhythmically, the variations between the mean sound intensity during

JM-events x_E and the mean sound intensity inter-JM-events x_{IE} are used to indicate the suitability for recognition of true JM-events in a given segment by the following JM modulation index $MI_{JM} = (\overline{x_E} - \overline{x_{IE}})/(\overline{x_E} + \overline{x_{IE}})$, with:

$$\overline{x_E} = \frac{1}{l_s} \sum_{k=1}^l x^2[k] \cdot w[k] \quad (\text{A.1})$$

$$\overline{x_{IE}} = \frac{1}{l_n} \sum_{k=1}^l x^2[k] \cdot (1 - w[k]) \quad (\text{A.2})$$

where $x[k]$ is the corrupted segment, l is the length of a segment in samples, l_s and l_n are the number of samples with and without JM activity, respectively, and $w[k]$ is a logical function indicating the presence of an JM-event in the k -th sample.

Regarding the classification task, the noise present during JM-events must be isolated. To estimate a noise-free activity signal $\hat{s}[k]$ and a noisy signal $\hat{n}[k]$, a multiband spectral subtraction algorithm assuming uncorrelated additive noise in signal segments was used (Loizou, 2013). The criterion used to indicate the difficulty to classify JM-events in a given segment is the SNR, which is computed as follow:

$$SNR(\text{dB}) = 10 \log \left(\sum_{k=1}^l \hat{s}^2[k] \right) - 10 \log \left(\sum_{k=1}^l \hat{n}^2[k] \right) \quad (\text{A.3})$$

Appendix B. Computational Cost

The computational cost of the CBEBA stages depends on the selected classifier, the sampling frequency of the input signal k , and the sub-sampling frequency n , which also determines the size of buffers. In order to obtain a simple comparison with other online algorithms, a MLP classifier using two JM-events per second was considered in this analysis. Worst-case scenarios were considered for each stage in order to get a theoretical upper bound. The required number of operations per stage of computation for CBEBA was:

1. Signal pre-processing
 - (a) *Bandwidth limitation*: A second-order band-pass filter is applied to avoid undesired noises, which involves 4 multiplications and 3 sums per sample ($7k$ ops/s).

- (b) *Power computation*: 1 multiplication per sample is required to obtain the instantaneous power signal (k ops/s).
2. Buffering
 - (a) *Envelope computation*: A second-order IIR low-pass filter is applied to the instantaneous power signal to compute the envelope, which requires 5 multiplications and 4 sums per sample ($9k$ ops/s). In order to reduce the computational cost of the next stages the envelope is decimated, so a counter and a comparison per sample is needed ($2k$ ops/s). These values are stored in $e(n)$ at a rate n (n ops/s).
 - (b) *Frame-energy computation*: An internal accumulator computes the sum of the energy signal (k ops/s). At a rate n the accumulator value is stored in $f(n)$ and it is reseted ($2n$ ops/s).
 3. JM-event detection
 - (a) *Possible event*: Finding the first valid peak of the envelope requires 1 subtraction and 1 comparison per JM-event. Searching for a second peak in the envelope requires $0.275n$ subtractions and $0.275n$ comparisons per JM-event. Centring the position between both peaks requires 1 subtraction and 1 comparison per JM-event. Additionally, updating the threshold $T(n)$ requires $0.375n$ operations per JM-event.
 - (b) *Event bouts*: Determining the threshold energy level involves $0.5n + 4$ operations per JM-event. Assuming the worst case scenario in which the start and end positions of the candidate JM-event coincide with the start and end positions of $f(n)$, $0.5n$ comparisons and 8 accumulations per JM-event are required.
 4. Feature extraction: In this scenario, $0.5n - 1$ sums per JM-event are needed to compute the total energy feature. $0.5n$ subtractions and $0.5n$ comparisons per JM-event are required to determine the sign of the envelope slope feature. Computing the duration of the candidate JM-event requires 1 subtraction per JM-event. The symmetry feature involves a maximum of $0.5n - 1$ comparisons, 1 subtraction and 1 division per JM-event. Finally, the accumulated envelope speed feature requires $1.5n - 1$ operations per JM-event.
 5. JM-event classification
 - (a) *Event decision*: A candidate JM-event must meet the two conditions previously established. The first condition requires 1 comparison per JM-event. The second condition requires $n - 1$ sums

per JM-event to compute the total energy in $f(n)$, and 1 multiplication and 1 comparisons per JM-event to evaluate the logic condition.

(b) *Event classifier*: According to the possible values of the optimal hyperparameters, the most expensive computational cost is considered. The required operations in the MLP classifiers depend on the number of input neurons (5), the number of hidden layers (1), the count of neurons in the hidden layer (7) and the number of output neurons (4). The cost of these classifiers is 192 operations per JM-event.

6. Threshold algorithms and parameters tuning: Update and compute all values related to $T(n)$ requires $n + 24$ operations per JM-event. Similarly, the parameters related to T_F involve 15 operations per JM-event.

The required total number of operations per second is:

$$f_{CBEBA,MLP}(k, n) = 20k + 17.85n + 500$$

References

- Andriamandroso, A., Bindelle, J., Mercatoris, B., Lebeau, F., 2016. A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnologie, Agronomie, Société et Environnement* 20.
- Andriamandroso, A.L.H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B., Dumortier, P., Blanchy, G., Blaise, Y., Bindelle, J., 2017. Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors. *Comput. Electron. Agric.* 139, 126–137.
- Ballabio, D., Grisoni, F., Todeschini, R., 2018. Multivariate comparison of classification performance measures. *Chemometrics and Intelligent Laboratory Systems* 174, 33–44.
- Berckmans, D., 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. *Rev. Sci. Tech.* 33, 189–196.
- Chelotti, J.O., Vanrell, S.R., Galli, J.R., Giovanini, L.L., Rufiner, H.L., 2018. A pattern recognition approach for detecting and classifying jaw movements in grazing cattle. *Computers and Electronics in Agriculture* 145, 83–91.
- Chelotti, J.O., Vanrell, S.R., Milone, D.H., Utsumi, S.A., Galli, J.R., Leonardo Rufiner, H., Giovanini, L.L., 2016. A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle. *Comput. Electron. Agric.* 127, 64–75.
- Chelotti, J.O., Vanrell, S.R., Rau, L.S.M., Galli, J.R., Planisich, A.M., Utsumi, S.A., Milone, D.H., Giovanini, L.L., Rufiner, H.L., 2020. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Computers and Electronics in Agriculture* 173, 105443.
- Demuth, H.B., Beale, M.H., De Jess, O., Hagan, M.T., 2014. *Neural network design*. Martin Hagan.
- Deniz, N.N., Chelotti, J.O., Galli, J.R., Planisich, A.M., Larripa, M.J., Rufiner, H.L., Giovanini, L.L., 2017. Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals. *Computers and electronics in agriculture* 138, 167–174.

- DeVries, T.J., Beauchemin, K.A., Dohme, F., Schwartzkopf-Genswein, K.S., 2009. Repeated ruminal acidosis challenges in lactating dairy cows at high and low risk for developing acidosis: Feeding, ruminating, and lying behavior. *Journal of Dairy Science* 92, 5067–5078.
- Fuentes, A., Yoon, S., Park, J., Park, D.S., 2020. Deep learning-based hierarchical cattle behavior recognition with spatio-temporal information. *Computers and Electronics in Agriculture* 177, 105627.
- Galli, J.R., Cangiano, C.A., Milone, D.H., Laca, E.A., 2011. Acoustic monitoring of short-term ingestive behavior and intake in grazing sheep. *Livestock Science* 140, 32–41.
- Galli, J.R., Cangiano, C.A., Pece, M.A., Larripa, M.J., Milone, D.H., Utsumi, S.A., Laca, E.A., 2018. Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *animal* 12, 973–982.
- Galli, J.R., Milone, D.H., Cangiano, C.A., Martínez, C.E., Laca, E.A., Chelotti, J.O., Rufiner, H.L., 2020. Discriminative power of acoustic features for jaw movement classification in cattle and sheep. *Bioacoustics* 29, 602–616.
- Greenwood, P.L., Paull, D.R., McNally, J., Kalinowski, T., Ebert, D., Little, B., Smith, D.V., Rahman, A., Valencia, P., Ingham, A.B., et al., 2017. Use of sensor-determined behaviours to develop algorithms for pasture intake by individual grazing cattle. *Crop and Pasture Science* 68, 1091–1099.
- He, H., Bai, Y., Garcia, E.A., Li, S., 2008. Adasyn: Adaptive synthetic sampling approach for imbalanced learning, in: 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence), IEEE. pp. 1322–1328.
- Hodgson, J.G., 1990. *Grazing management: science into practice*. John Wiley & Sons Inc.
- Laca, E.A., Ungar, E.D., Seligman, N.G., Ramey, M.R., Demment, M.W., 1992. An integrated methodology for studying short-term grazing behaviour of cattle. *Grass and forage science* 47, 81–90.

- Laca, E.A., WallisDeVries, M.F., et al., 2000. Acoustic measurement of intake and grazing behaviour of cattle. *Grass and Forage Science* 55, 97–104.
- Loizou, P.C., 2013. *Speech enhancement: theory and practice*. CRC press.
- Michie, C., Andonovic, I., Davison, C., Hamilton, A., Tachtatzis, C., Jonsson, N., Duthie, C.A., Bowen, J., Gilroy, M., 2020. The internet of things enhancing animal welfare and farm operational efficiency. *Journal of Dairy Research* 87, 20–27.
- Milone, D.H., Galli, J.R., Cangiano, C.A., Rufiner, H.L., Laca, E.A., 2012. Automatic recognition of ingestive sounds of cattle based on hidden markov models. *Comput. Electron. Agric.* 87, 51–55.
- Milone, D.H., Rufiner, H.L., Galli, J.R., Laca, E., Cangiano, C.A., 2009. Computational method for segmentation and classification of ingestive sounds in sheep. *Computers and Electronics in Agriculture* 65, 228–237.
- Nydegger, F., Gygax, L., Egli, W., 2010. Automatic measurement of rumination and feeding activity using a pressure sensor, in: *International Conference on Agricultural Engineering-AgEng 2010: towards environmental technologies*, 6-8 September 2010, Clermont-Ferrand, France, pp. 1–8.
- Roden, M.S., 1996. *Analog and digital communication systems*. Prentice Hall.
- Saitoh, T., Kato, Y., 2021. Evaluation of wearable cameras for monitoring and analyzing calf behavior: A preliminary study. *Animals* 11, 2622.
- Sauvant, D., 2000. Granulométrie des rations et nutrition du ruminant. *Productions animales* 13, 99–108.
- Schirmann, K., Chapinal, N., Weary, D.M., Heuwieser, W., von Keyserlingk, M.A.G., 2011. Short-term effects of regrouping on behavior of prepartum dairy cows. *Journal of Dairy Science* 94, 2312 – 2319.
- Schirmann, K., Chapinal, N., Weary, D.M., Vickers, L., von Keyserlingk, M.A.G., 2013. Short communication: Rumination and feeding behavior before and after calving in dairy cows. *Journal of Dairy Science* 96, 7088 – 7092.

- Schirmann, K., von Keyserlingk, M.A.G., Weary, D.M., Veira, D.M., Heuwieser, W., 2009. Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.* 92, 6052–6055.
- Smith, D., Rahman, A., Bishop-Hurley, G.J., Hills, J., Shahriar, S., Henry, D., Rawnsley, R., 2016. Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Comput. Electron. Agric.* 131, 40–50.
- Sokolova, M., Lapalme, G., 2009. A systematic analysis of performance measures for classification tasks. *Inf. Process. Manag.* 45, 427–437.
- Stone, A.E., 2020. Symposium review: The most important factors affecting adoption of precision dairy monitoring technologies. *Journal of Dairy Science* 103, 5740–5745.
- van der Maaten, L., Hinton, G., 2008. Visualizing data using t-sne. *Journal of machine learning research* 9, 2579–2605.
- Vanrell, S.R., Chelotti, J.O., Galli, J.R., Utsumi, S.A., Giovanini, L.L., Rufiner, H.L., Milone, D.H., 2018. A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *Computers and Electronics in Agriculture* 151, 392–402.
- Watt, L.J., Clark, C.E.F., Krebs, G.L., Petzel, C.E., Nielsen, S., Utsumi, S.A., 2015. Differential rumination, intake, and enteric methane production of dairy cows in a pasture-based automatic milking system. *Journal of dairy science* 98, 7248–7263.
- Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., O’Brien, B., 2018. Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows. *Journal of Neuroscience Methods* 300, 138–146.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* 1, 80.
- Young, S., Evermann, G., Gales, M., Hain, T., Kershaw, D., Liu, X., Moore, G., Odell, J., Ollason, D., Povey, D., et al., 2002. The htk book. Cambridge university engineering department 3, 12.

Zehner, N., Umstätter, C., Niederhauser, J.J., Schick, M., 2017. System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *Computers and Electronics in Agriculture* 136, 31–41.

Supplementary Material: Internal parameters choice

The spectral energy distribution of the JM-events corresponding to the DS1 and DS2 datasets were calculated with a linear autoregressive model of order 2^{10} . The linear prediction coefficients were computed to obtain a frequency resolution of 1.35 Hz using the Levinson-Durbin algorithm (Galli et al., 2020). Figure 1 shows the relative energy distribution for each JM-event in each dataset.

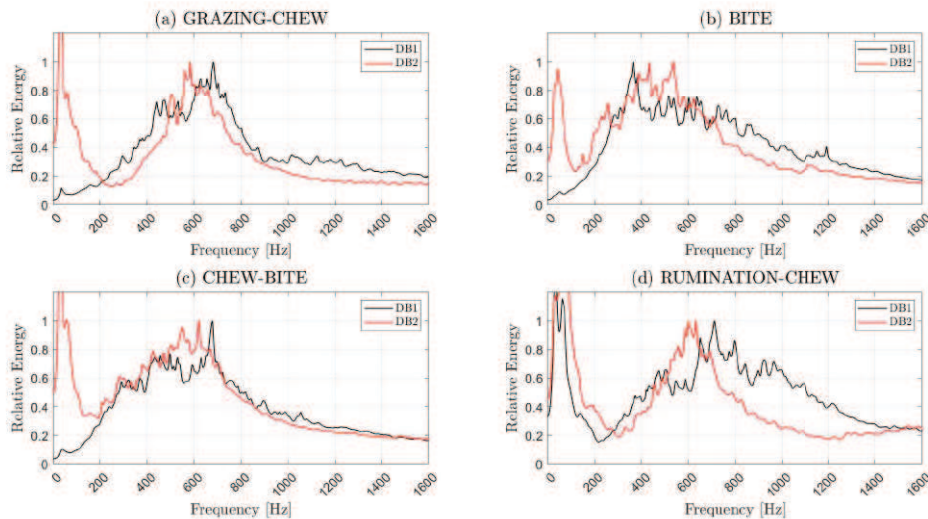


Figure 1: Spectral energy distribution for grazing-chews, bites, composite chew-bites and rumination-chews events computed from noiseless (DS1) and noisy (DS2) datasets.

The internal parameters of the algorithm to be optimised do not show correlation between each other. Modifying the values for f_{lower} and f_{upper} changes the bandwidth of the input signal, and consequently the SNR. The value of f_{cut} affects some features extracted from the JM-events: the sign of the envelope slope, the accumulated envelope speed and the symmetry. As the frequency increases, the sign of the envelope slope and the symmetry features are more susceptible to being affected by noise, and the envelope speed response increases allowing higher levels of discriminatory information. Therefore, there is an optimal cut-off frequency for JM-event classification. The value of n is related to the resolution of the boundary positions n_{START} and n_{END} used to extract the features. The limits and steps values of the internal parameters used in the grid search method and optimal values adopted in each dataset are shown in Table 1.

Supplementary Material: Computation of threshold algorithms

Internally, the buffers $e(n)$ and $f(n)$ are divided into three fixed length segments. The first and third part of the buffers are used to extract the background noise, whereas the second part is used to extract information related to JM-events (see Figure 2).

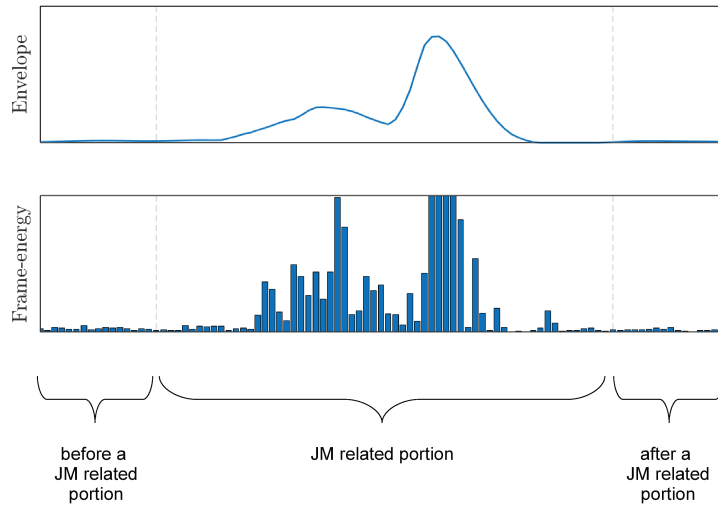


Figure 2: Internal division of the *envelope* $e(n)$ and *frame-energy* $f(n)$ buffers.

Time-varying threshold algorithm

Once an JM-event has been classified by the CBEBA, the time-varying threshold algorithm $T(n)$ tunes the value in three successive steps:

1. The time elapsed between the previous and the current classified JM-event is calculated, and the maximum and minimum envelope values are computed from $e(n)$.

Table 1: Description of the variation of internal parameters of the CBEBA used in the grid search method for noiseless (DS1) and noisy (DS2) datasets.

Parameter	Range	Step	Optimal values	
			DS1	DS2
f_{lower} [Hz]	0 - 250	25	175	175
f_{upper} [Hz]	700 - 1000	25	900	900
f_{cut} [Hz]	2 - 20	0.5	9	9
n [Hz]	50 - 250	25	150	150

2. These values are used to compute the parameters associated with $T(n)$ according to:

- Unresponsive period (T_U): period of time after detecting the end of a classified JM-event in which the algorithm is no longer searching for a new JM-event. T_U is computed for each event as 25% of the average time between two classified JM-events for the last four JM-events.
- Maximum period (T_M): maximum time that a JM-event can last within the same activity. T_M is computed for each JM-event as 50% of the average time between two classified JM-events for the last four JM-events.
- Threshold Maximum (T_{MAX}): maximum value expected for a new possible JM-event. T_{MAX} is computed as the 50% of the average maximum envelope value for the last four JM-events.
- Threshold Minimum (T_{MIN}): minimum value expected for a new possible JM-event. T_{MIN} is computed as:

$$T_{MIN} = \min \left\{ \left(T_{MIN,avg} + \frac{T_{MAX} - T_{MIN,avg}}{4} \right), 4 \cdot T_{MIN,avg} \right\}$$

where, $T_{MIN,avg}$ is the average minimum envelope value for the last four JM-events.

- Threshold step (ΔT): is the decrease level of threshold $T(n)$ from T_{MAX} to T_{MIN} once after the unresponsive period T_U expires, and serves therefore to significantly improve the JM-event detection sensitivity. The ΔT is computed as $(T_{MAX} - T_{MIN})/T_M$.
- Decrease period (T_D): when the animals resume grazing or rumination after a long period of JM inactivity, it is possible that the previously calculated threshold level remains high to detect new coming JM-events. Therefore, T_D is the lag time after the last classified JM-event. After that time the threshold $T(n)$ decreases its value to improve the JM-event detection sensitivity. T_D is computed for each JM-event as eight times the time between two classified JM-events for the last four JM-events.

- Increase period (T_I): it is possible that during a period of low-intensity in the envelope $e(n)$ (e.g. silence) the threshold level remains very sensitive to false JM-events likely to noises or interference trends. Therefore, T_I is the period of time since the last classified JM-event after which the threshold $T(n)$ increases its value if a false JM-event detection occurs. T_I is computed for each JM-event as four times the average time between two classified JM-events for the last four JM-events.

3. The new value of $T(n)$ changes according to:

$$T(n) = \begin{cases} T(n-1) - \Delta T, & \forall T_U < n < T_M \\ 0.5 \cdot T(n-1), & \text{if } n = p \cdot T_D, \text{ with } p = 1, 2, 3, \dots \\ 1.5 \cdot T(n-1), & \text{if } n > T_I \text{ and a false JM detection occurs} \\ T(n-1), & \text{otherwise} \end{cases}$$

Energy threshold algorithm

Once an JM-event has been classified by CBEBA, the noise power and the signal power of the JM-event are estimated. The background noise power is computed as the average value of the first and third parts of $f(n)$. The signal power is computed as the average energy level of the JM-event located in the second part (Figure 2). Then, the energy threshold value of T_F updates value according to:

$$T_F = k \cdot N_{FEL}$$

where k is the sensibility parameter which is modified in discrete values according to the average value of the SNR level for the last four classified JM-events, and N_{FEL} is the average of the background noise power for the last four classified JM-events.

References

Galli, J.R., Milone, D.H., Cangiano, C.A., Martínez, C.E., Laca, E.A., Chelotti, J.O., Rufiner, H.L., 2020. Discriminative power of acoustic features for jaw movement classification in cattle and sheep. *Bioacoustics* 29, 602–616.

Anexo B

**An online method for estimating
grazing and rumination bouts using
acoustic signals in grazing cattle**

An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle

José O. Chelotti^a, Sebastián R. Vanrell^a, Luciano Martinez-Rau^a,
Julio R. Galli^{b,e}, Alejandra M. Planisich^b, Santiago A. Utsumi^c,
Diego H. Milone^a, Leonardo L. Giovanini^a, H. Leonardo Rufiner^{a,d}

^a*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i), FICH-UNL/CONICET, Argentina*

^b*Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Argentina*

^c*W.K. Kellogg Biological Station and Department of Animal Science, Michigan State University, United States*

^d*Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina*

^e*Instituto de Investigaciones en Ciencias Agrarias de Rosario, IICAR, UNR-CONICET, Argentina*

Abstract

The growth of the world population expected for the next decade will increase the demand for products derived from cattle (i.e., milk and meat). In this sense, precision livestock farming proposes to optimize livestock production using information and communication technologies for monitoring animals. Although there are several methodologies for monitoring foraging behavior, the acoustic method has shown to be successful in previous studies. However, there is no online acoustic method for the recognition of rumination and grazing bouts that can be implemented in a low-cost device. In this study, an online algorithm called bottom-up foraging activity recognizer (BUFAR) is proposed. The method is based on the recognition of jaw movements from sound, which are then analyzed by groups to recognize rumination and grazing bouts. Two variants of the activity recognizer were explored, which were based on a multilayer perceptron (BUFAR-MLP) and a decision tree (BUFAR-DT). These variants were evaluated and compared under the same conditions with a known method for offline analysis. Compared to the former

Email address: jchelotti@sinc.unl.edu.ar (José O. Chelotti)

Preprint submitted to Computers and Electronics in Agriculture

method, the proposed method showed superior results in the estimation of grazing and rumination bouts. The MLP-variant showed the best results, reaching F1-scores higher than 0.75 for both activities. In addition, the MLP-variant outperformed a commercial rumination time estimation system. A great advantage of BUFAR is the low computational cost, which is about 50 times lower than that corresponding to the former method. The good performance and low computational cost makes BUFAR a highly feasible method for real-time execution in a low-cost embedded monitoring system. The advantages provided by this system will allow the development of a portable device for online monitoring of the foraging behavior of ruminants. Web demo available at: <https://sinc.unl.edu.ar/web-demo/bufar/>

Keywords: Acoustic monitoring, activity recognition, ruminant foraging behavior, precision livestock farming, pattern recognition, machine learning.

1. Introduction

Accurate monitoring of animal foraging behavior is a complex but essential task to optimize livestock production systems (Hodgson and Illius, 1998). Changes in the ruminant foraging behavior are indicators of animal health and welfare and can be useful in early detection and prevention of several diseases. For example, an increment in rumination time can be associated with an increment of saliva production and improvements in rumen health (Beauchemin, 1991). Conversely, a reduction of rumination can be interpreted as an indicator of stress (Herskin et al., 2004), anxiety (Bristow and Holmes, 2007), or a disease (Hansen et al., 2003; Paudyal et al., 2018; Welch, 1982). In the last decade, precision livestock farming has been presented as a useful approach to tackle these problems, using advanced technology to monitor each animal. In this sense, recent technological developments have facilitated the use of sensors to monitor many physical variables both for animal science research and for practical farm level applications (Berckmans, 2014).

Foraging behavior of ruminants can be characterized by jaw movements (short timescale) and activities (long timescale). Jaw movements (JM) have a duration close to 1 s, whereas activity bouts can last from minutes to hours. The JM (or masticatory events) are biting, when herbage is apprehended and severed; chewing, when herbage is comminuted; and a combination of chewing and biting in a single JM, which is called chew-bite (Galli et al., 2018;

Laca et al., 1992; Ungar and Rutter, 2006). Main foraging activities are grazing and rumination. Their duration widely fluctuates in the day. Grazing can cover from 25 to 50% of the day and rumination from 15 to 40% (Hodgson, 1990; Kilgour, 2012; Phillips, 1993). The grazing process involves searching, apprehending, chewing, and swallowing herbage. Rumination involves bolus regurgitation, chewing, and deglutition, in a periodic cycle that typically last 1 min. During both activities, JM are performed rhythmically with a frequency that ranges from 0.75 to 1.20 JM per second (Andriamandroso et al., 2016). While grazing, the three types of JM are present (i.e., chew, bite and chew-bite), whereas only chews are present during rumination (Hodgson and Illius, 1998).

An automatic monitoring system should be reliable, insightful, and practical to implement. For instance, these goals imply that recorded signals should be analyzed without human assistance, that the methodology should be scalable to large herds (even in pasture-based production systems), that the device autonomy should facilitate the collection of data over long periods of time (from days to weeks), and that data should be processed online to reduce in-device data-storing and communication requirements. Thus, an ideal methodology to be deployed in the field is one that is powerful at characterizing the foraging behavior as well as it is efficient at data processing.

Different sensing technologies have been used in the development of automatic monitoring systems, such as motion sensors, noseband pressure sensors, and microphones (Andriamandroso et al., 2016). Among motion sensors it is widespread the use of accelerometers (Arcidiacono et al., 2017; Giovanetti et al., 2017; González et al., 2015; Martiskainen et al., 2009) and inertial measurement units (Andriamandroso et al., 2017; Greenwood et al., 2017; Smith et al., 2016). These sensors have been used to recognize a broader set of activities such as rumination, grazing, resting, drinking and walking. An activity is determined by a postural analysis of the animal, where the sensors are used to estimate the position and motion of its head and body. However, this strategy can confuse activities that share the same posture. A better strategy for recognizing ruminating, eating and drinking activities is the use of noseband pressure sensors (Nydegger et al., 2010; Rutter, 2000; Rutter et al., 1997; Werner et al., 2018; Zehner et al., 2017). They have been used in the analysis of housed and free-grazing cows during one- to two-hour sessions. This yielded very good results, but further studies are required for continuous long-term monitoring. A limitation of this approach is that does not discriminate between JM (i.e., they are not classified) which is a require-

ment for a more detailed analysis such as herbage intake estimation (Galli et al., 2018).

Acoustic monitoring has proven to be reliable for recognizing short-term JM in free-ranging cows (Chelotti et al., 2018; Clapham et al., 2011; Laca et al., 1992; Milone et al., 2012; Navon et al., 2013). In particular, the chew-bite intelligent algorithm (CBIA) performs an online processing of the sound signal and has achieved very good results (Chelotti et al., 2018). A related commercial monitoring system is the Hi-Tag system (SCR Engineers Ltd., Netanya, Israel). Its design is focused on the autonomy, portability and hardware robustness required by the application. Besides it is based on microphones, the analysis of the signal is exclusively focused on rumination monitoring (Goldhawk et al., 2013; Schirmann et al., 2009). Recently, acoustic monitoring has also been successful on long-term recognition of foraging activities in free-ranging cows (Vanrell et al., 2018). The regularity-based acoustic foraging activity recognizer (RAFAR) was able to identify grazing and rumination bouts from sound recordings. The success of RAFAR relies on an offline analysis of long recordings (several hours), which clearly expose the regularities of foraging activities. Those recordings are acquired in each animal of the herd and then analyzed in a desktop computer. However, there are some practical limitations with this approach. A portable device, has limited storage capacity, processing capability, and power supply. These limitations becomes more relevant when the application on large herds is desired.

In this study, the acoustic monitoring strategy is taken one step further. The main point to explore is the potential of identifying the foraging activities from a prior recognition of JM following a bottom-up approach. The proposed method is focused on an online processing of the acoustic signals, i.e. the input signal is processed sample-by-sample, as it is received. In addition, the method should have relatively low computational cost and be focused on its real-time implementation in a low-cost embedded system. This would contribute to establish the acoustic monitoring as a non-invasive alternative that could handle the requirements of the application and can provide insights about natural foraging behavior of ruminants.

2. Material and methods

2.1. Proposed method

An online method for detection and classification of the most important foraging activities of ruminants is presented in this section. The method can process the signal sample-by-sample (online fashion). The bottom-up foraging activity recognizer (BUFAR) has two levels of recognition. First, JM are recognized and then this information is used to estimate rumination and grazing bouts. As a result, the information about nutritional status can be enhanced by providing statistics of both JM and activity bouts.

Fig. 1 shows typical sound recordings during (a) grazing and (b) rumination. The amplitude of the sound signals might be seen as an obvious measure for discrimination. However, variations in the amplitude across microphones, recording devices, sessions, and cows have not allowed a reliable classification. By contrast, the rate of JM of both activities is very similar and it helps to distinguish activity bouts from noisy segments in the recordings. A clear difference between the activities is the proportion of JM. For example, in these recordings, grazing has 25% of chews, 10% of bites, and 65% of chew-bites, whereas rumination has a 100% of chews. Thus, the rate and the proportion of JM are the keys of the proposed method.

A diagram of the proposed system BUFAR is shown in Fig. 2. It has five stages that perform the required processing of data to recognize JM and foraging activities. For the sake of a low computational cost, tasks within each stage have been simplified whenever it was possible. The input of the system is the sound signal produced during foraging activities. Three activities are considered: rumination, grazing, and other activities. Other activities include any activity other than rumination or grazing (i.e., milking, silence, confusing sounds, etc.). Detection and classification of JM are performed with the CBIA algorithm (Chelotti et al., 2018). CBIA comprises three stages: signal pre-processing, jaw-movement detection, and jaw-movement classification. In signal pre-processing stage, the raw signal is conditioned and filtered to improve the signal-to-noise ratio (SNR) and remove slow varying trends. Jaw-movement detection stage spots these movements by analyzing the filtered signal with an adaptive threshold. Each JM is assigned with a timestamp and a set of features (duration, maximum amplitude, shape index, and symmetry). The timestamp is saved in the segment buffer and it will be used for activity recognition. In the classification stage, the features

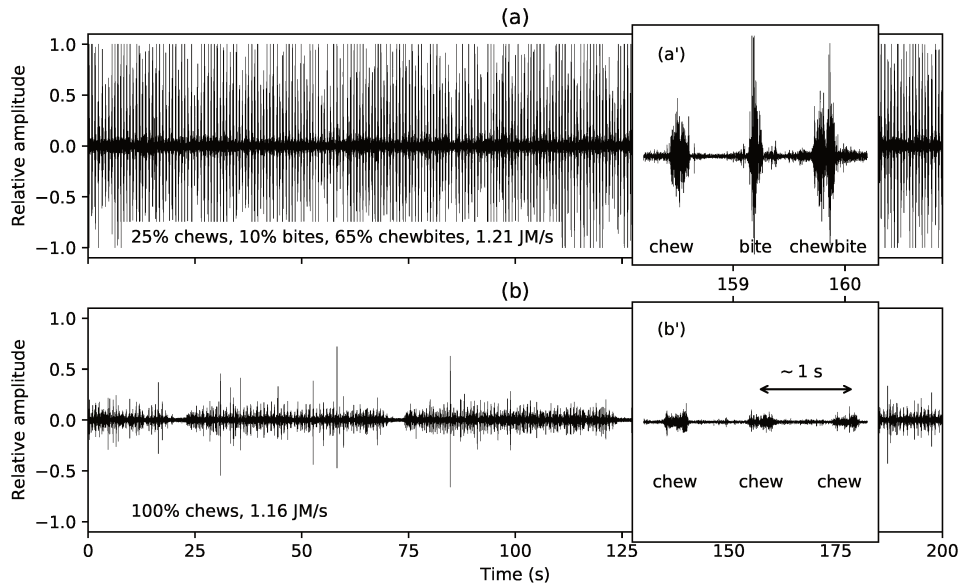


Figure 1: (a) Grazing and (b) rumination activities. Typical percentages and rate of jaw movements by activity. The jaw movement included in each activity are zoom-in.

of each JM are taken by a neural network model to assign an event label: bite (b), chew (c), or chew-bite (cb).

The proposed system performs activity recognition by analyzing fixed-length segments of the acoustic signal. JM that are detected and classified within a segment are stored in a segment buffer. The rate of JM in a segment and the proportions of their types are computed to feed the last processing stage. At this point, activity classification could be seen as a simple task, but an exploratory data analysis on the training set has shown a complex underlying distribution of the segment features (rate, %c, %b, %cb). The rate of recognized JM during rumination and grazing is expected to be in the range from 0.75 to 1.40 Hz (Fig. 3). By contrast, the rate of JM identified during other activities presents a lower frequency. The overlapping among rate distributions of activities is part of the problem.

The triangle plot in Fig. 4 shows the proportions of the identified JM by CBIA for several segments of the training set. Proportions of a single segment %c, %b, and %cb always sum to 1.0. The top corner corresponds to 100% of chews, the bottom left corner corresponds to 100% of chew-bites, and the bottom right corner corresponds to 100% of bites. Points inside the

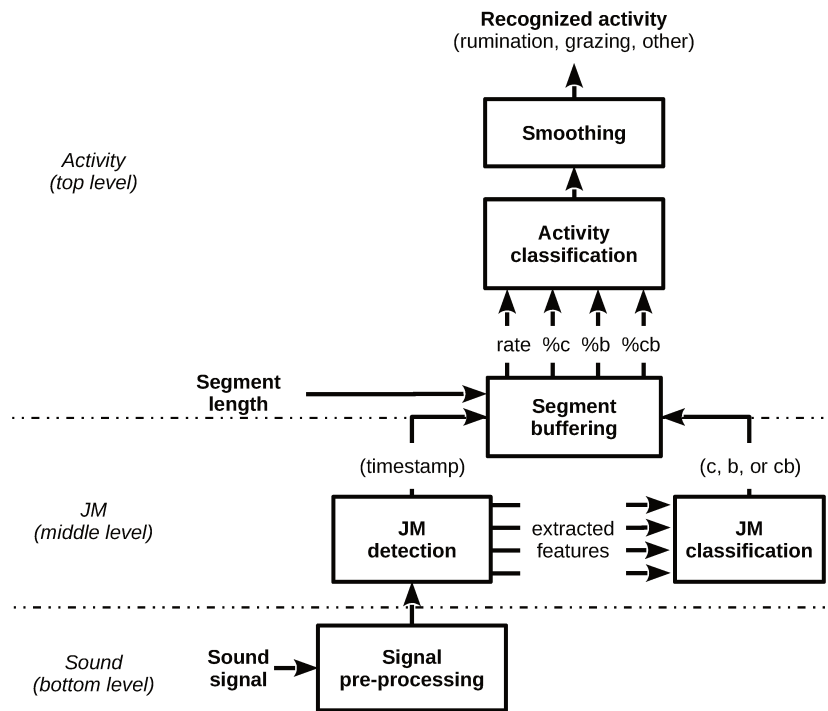


Figure 2: General diagram of the bottom-up foraging activity recognizer (BUFAR). Activity classification uses information of jaw movements (JM) within a segment. JM include: chew (c), bite (b), and chew-bite (cb).

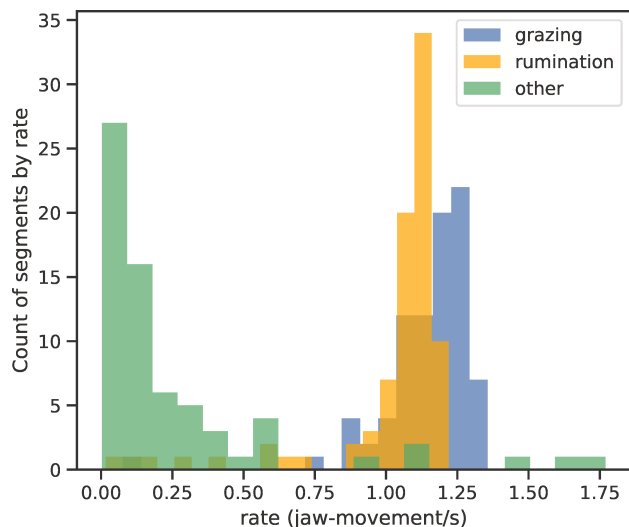


Figure 3: Distributions of jaw movements rate for grazing, rumination, and other activities on segments of the training set.

triangle correspond to segments composed by more than one type of JM. For example, while rumination is mainly composed by chews (orange diamonds are on the top corner), grazing has a diversity of JM compositions (blue circles are dispersed in the triangle). During other activities, bites are the most assigned type of JM (green squares located on the bottom right corner). However, they are mostly false positives for class b.

Distributions of segment features show that the recognition of JM within grazing and rumination activities is not perfect. For example, CBIA detects a few bites during rumination, which is not actually true. Thus, the problem of distinguishing between activities requires a powerful method to handle these errors. In this study, the use of a simple method of machine learning is proposed. Activity classification is performed by a trainable model, such as a multilayer perceptron or a decision tree, which assigns an activity label to the segment. A multilayer perceptron (MLP) is a feed-forward artificial neural network that can deal with non-linearly separable data (Bishop, 2006). It consists of several layers of nodes (simple perceptrons) in a directed graph, with each layer fully connected to the next one, but without connections between nodes in the same layer. Decision Trees (DTs) have the ability of learning simple decision rules and systematizing them in order to arrive at complex decisions (Bishop, 2006). For numerical attributes, DTs divide the

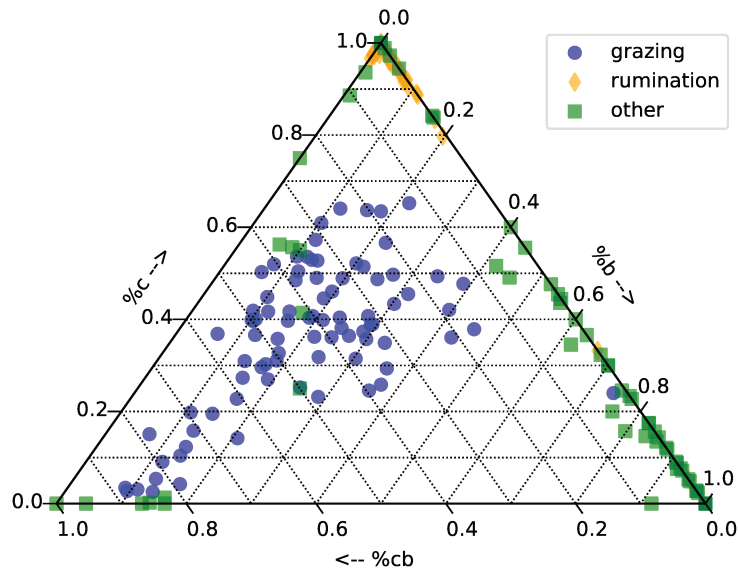


Figure 4: Proportions of bite (%b), chew (%c), and chew-bite (%cb) as labeled by Chew-Bite Intelligent Algorithm (CBIA) for grazing, rumination, and other activities. The top corner corresponds to 100% of chews, the bottom left corner corresponds to 100% of chew-bites, and the bottom right corner corresponds to 100% of bites.

feature space into axis-parallel rectangular regions and label each region with the correspondent class. In addition, a DT provides solutions which are easy to implement and understand.

At the end of the processing stages, each segment of the input signal has a label that indicates if it corresponds to rumination, grazing, or other activity. Finally, a smoothing process is applied over the sequence of labeled segments in order to remove short gaps and thus reduce fragmentation of activity bouts. Thus, long recognized bouts are encouraged, which mimics the typical length of activity bouts.

2.2. Acoustic database

Acoustic signals were collected in August of 2014 at the dairy facility in the Kellogg Biological Station Robotic and Grazing Farm, operated by the Michigan State University. As described in (Vanrell et al., 2018), the code for animal use by the Institutional Animal Care and Use Committee of the Michigan State University was reviewed, approved, and conducted according to protocols for animal handling and care. SONY ICDPX312 recorders were used to record the signals (Fig. 5a). A microphone was placed facing inwards

on the forehead of cows (Fig. 5b) and was protected by a rubber foam (Milone et al., 2012). All recordings were saved in WAV file format, considering a 44.1 kHz sampling rate and 16-bit resolution.

Cows were rotationally grazed on a pasture-based robotic milking system with voluntary cow traffic as described previously in Watt et al. (2015). Briefly, the five multiparous experimental cows (parity 2.6 ± 0.5 ; days in milk 108 ± 34 ; body weight 654 ± 21 kg; milk yield 39 ± 4 kg; milkings/d 3 ± 1) were group housed and managed together as part of a larger robotic and grazing herd of 146 Holstein cows, allocated to two Lely A3-Robotic Milkers (Lely Industries N.V., Maassluis, the Netherlands). Cows were raised and grazed previously on same pasture so they were properly adapted to the farming system and diets before this study commenced. Milking was conducted according to milk table permissions set by a minimum expected milk yield/milking of 9.1 kg or 6 h of minimum interval. During milking cows were fed a grain based concentrate (GBC) at a rate of 1 kg per 6 kg of milk production (cap 12 kg/ cow d^{-1}). The average crude protein (CP), neutral detergent fiber (NDF), and net energy for lactation (NEL) of the GBC pellet offered (Cargill Inc, Big Lake, MN) was 19.30%, 9.94%, and 2.05 Mcal/kg, respectively. Cows had 24 h access to pasture dominated either by perennial ryegrass (*Lolium perenne*) and white clover (*Trifolium repens*), or orchardgrass (*Dactylis glomerata*), tall fescue (*Festuca arundinacea*) and the same white clover. Cows were grazed at an average herbage allowance of 30 kg of DM/cow d^{-1} split evenly into an AM and PM break of fresh pasture (15 kg of DM/cow) freely accessible at opposite locations of the farm (north and south) from 10:00 h to 22:00 h and from 22:00 h to 10:00 h, respectively. Herbage allowance was adjusted according to changes in pasture growth rates and measurements of pregrazing herbage cover (Y ; measured to ground level) by a plate meter ($Y = 125x$; $r^2 = 0.96$), using 30 readings of sward height (SH; x) taken alongside allocations. At the time of the study the average pregrazing and postgrazing herbage mass ($n = 16$ paddocks) was 2387 ± 302 kg/ha (19.2 ± 2.5 cm SH) and 1396 ± 281 kg/ha (11.2 ± 2.2 cm SH), respectively. The average CP (4010 CN combustion system, Costech Analytical Technologies Inc., Valencia, CA), NDF and acid detergent fiber (ADF) (200 Fiber Analyzer, Ankom Technology Corp., Fairport, NY), and acid detergent lignin (ADL) and 48 h in vitro DM digestibility (Daisy II, Ankom Technology Corp.) of hand pluck pasture samples ($n = 16$) was $18.7 \pm 2.5\%$, $49.3 \pm 4.5\%$, $25.7 \pm 2.0\%$, $3.3 \pm 0.8\%$, and $78.1 \pm 3.0\%$, respectively.



Figure 5: Box located on the cow’s neck: (a) microphone plug-in and (b) recording device. Internal view of (c) the headband and (d) the microphone protected by a rubber foam on the cow’s forehead.

Expert labeling was used as a control reference for comparison and evaluation against algorithms results. Two experts with prior experience on animal behavior scouting, and digital analysis of acoustic signals, viewed the plot of the sound waveform and listened to the recordings to make a decision. Experts were able to identify, classify, and label the activity blocks, either as grazing, rumination, or neither of these activities. Experts agreed 100% on the labeling, but there were small differences in the limits of each label (start and end). These marks were carefully revised further until an agreement was made. This type of labeling was already used in previous works and validated with visual check, in-situ and using videos (Chelotti et al., 2018, 2016; Vanrell et al., 2018).

For comparison purposes on rumination time estimation, the animals were continuously monitored during the experiments with the Hi-Tag rumination monitor system. This system consists of rumination loggers on the collar of the animal, stationary or mobile readers, and software for processing electronic records (Schirmann et al., 2009). Animal behavior was monitored with this system and summarized as the total time spent ruminating during two-hour chunks.

2.3. Performance metrics

In continuous activity recognition, performance evaluation requires a comparison between a reference sequence and a recognized sequence. The activity

blocks of the reference sequence and the recognized sequence may not be in a one-to-one correspondence. For example, a single block (an activity bout) of the reference sequence can be partially detected by three shorter blocks in the recognized sequence. A comprehensive set of performance metrics for continuous activity recognition has been proposed by Ward et al. (2011) and has been recently used in a related study (Vanrell et al., 2018). These metrics are based on two complementary short- and long-term timescales. They present a multidimensional and detailed description instead of a single performance number. In this way, the strengths and weaknesses of a recognizer can be assessed, avoiding ambiguity in the results. Short-term metrics are frame-based, which is a small fixed-length unit of time. Frame-based metrics facilitate a fine-grain analysis that resembles a continuous time analysis. By contrast, a block has no fixed-length and is defined as a continuous period of time of a sequence that has the same label. For example, a rumination block in the reference sequence is a rumination bout. Long-term metrics are block-based, which provide a different point of view, like a big picture of the recognition performance. This is particularly valuable to detect coarse-grain bias and to propose modifications in the recognizer.

The frame- and block-based error metrics were used to characterize each variant of the method. They are false negative rate (FNR_*), false discovery rate (FDR_*), recall (R_*), precision (P_*), fragmentation (F_*), merging (M_*), deletion (D_*), insertion (I_*), underfill (U_*), Overfill (O_*), and the standard F1-score ($F1_*$). All metrics were computed for each recording analyzed and then averaged for results presentation. For details about the computation of these metrics see Appendix A.

2.4. Experimental Setup

In this study, the following setup was considered for the proposed method. Computer experiments were performed considering that at time t the algorithm can use data available at time t and $t - \Delta t$ but no using data at $t + \Delta t$. This consideration is equivalent to online processing within the device. The configuration of CBIA was the same used in Chelotti et al. (2018). For the signal pre-processing stage, a Least Mean Square filter was used (Widrow et al., 1975). This adaptive filter has proven to be useful for removing trends at low computational cost. For detection of JM, the steps proposed in Chelotti et al. (2018) were implemented. For classification of JM, it was selected a one-hidden-layer multilayer perceptron.

An exploratory analysis on a subset of the training set was conducted for the segment buffering stage. Segments of 1.0, 2.5, 5.0, and 10.0 min in length were considered. The shortest segment considered (1.0 min) can capture at least a typical period of rumination. In addition, this segment length generally includes a number of JM that allows a suitable analysis. Segments longer than 10.0 min would result in poor temporal resolution. For the activity classification stage, two models were considered: i) a multilayer perceptron (MLP) with one hidden layer and a logistic activation function, and ii) a binary decision tree (DT) based in the Gini impurity measure. An hyper-parameter optimization was performed for both activity classifiers considering: the number of neurons in the hidden layer and learning rate for the MLP, and the pruning factor for the DT. This optimization was made following a 5-fold scheme with signals on the other subset of the training set and maximizing the accuracy measure¹. Finally, in the last stage, a smoothing process to avoid fragmentation in rumination and grazing bouts was applied: single segments were relabeled when they were surrounded by segments of the same activity.

For this study, 30 h of recordings containing rumination and grazing sessions were randomly selected to optimize the segment-length. Another set of 24 h of recordings were used to train an optimize parameters and hyper-parameters of the activity classifier and they were never used again. Classifiers were trained following a 5-fold scheme on the training set. Finally, the test results were obtained from a separate test set of 137 h of recordings, which were selected taking care that they correspond to a free-ranging environment. Those portions of the recordings captured inside the feeding barn were excluded from this study. The periods inside the feeding barn were identified acoustically by experts, guided by the environmental sound (machines, engines, and the reverberation inside the barn) and the distinctive sound of metal gates opening and closing, when the animals entered or left the barn. This selection has been guided by the labels (timestamps) provided by the experts and it is in agreement with the study that presents the RAFAR (Vanrell et al., 2018). The present work included a comparison with the RAFAR-MBBP variant.

A web demo of the method was developed with the tool (Stegmayer et al., 2016) and can be accessed at: <https://sinc.unl.edu.ar/web-demo/bufar/>.

¹This stage was implemented in python using the scikit-learn package.

Table 1: F1-score metrics on activity classification for different segment lengths using Bottom-Up Foraging Activity Recognizer - Multilayer Perceptron (BUFAR-MLP).

Segment-length	Grazing		Rumination	
	Frame-based	Block-based	Frame-based	Block-based
1.0 min	0.849(± 0.161)	0.693(± 0.355)	0.516(± 0.340)	0.500(± 0.173)
2.5 min	0.851 (± 0.165)	0.770(± 0.359)	0.631(± 0.311)	0.642(± 0.263)
5.0 min	0.812(± 0.181)	0.822 (± 0.196)	0.703 (± 0.274)	0.743 (± 0.318)
10.0 min	0.764(± 0.314)	0.811(± 0.244)	0.611(± 0.336)	0.567(± 0.279)

3. Results

3.1. Segment-length effect

Table 1 shows the effect of segment length in activity recognition using an MLP as the activity classifier (BUFAR-MLP). Frame- and block-based F1-scores provide measures of the recognition in a short and long timescale, respectively. The shortest segment considered (1.0 min) achieved good frame-based metrics on grazing but very poor metrics on rumination. The longest segment considered (10.0 min) achieved good block-based metrics on grazing and poor metrics on rumination. A comparison of block-based metrics on grazing between 2.5-min and 5-min segments showed a notable improvement in favor of 5-min segments. Regarding rumination, a comparison between 2.5-min and 5-min segments showed remarkable improvements in frame- and block-based metrics for 5-min segments. Similar results were obtained using a DT as the activity classifier. In an overall assessment, 5-min segments achieved a strong performance for both frame- and block-based F1-score on the studied activities.

3.2. Activity classification

Two variants of BUFAR were evaluated: i) one using a decision tree as the activity classifier (BUFAR-DT) and ii) one using a multilayer perceptron as the activity classifier (BUFAR-MLP). In a previous study (Vanrell et al., 2018), RAFAR showed notable performance when the entire sound recording was available (offline analysis). It is the only known method that estimates both grazing and rumination bouts from acoustic signals. For comparison purposes, the RAFAR-MBBP variant was considered in this study (in the following referred as RAFAR). For a fair comparison between RAFAR and

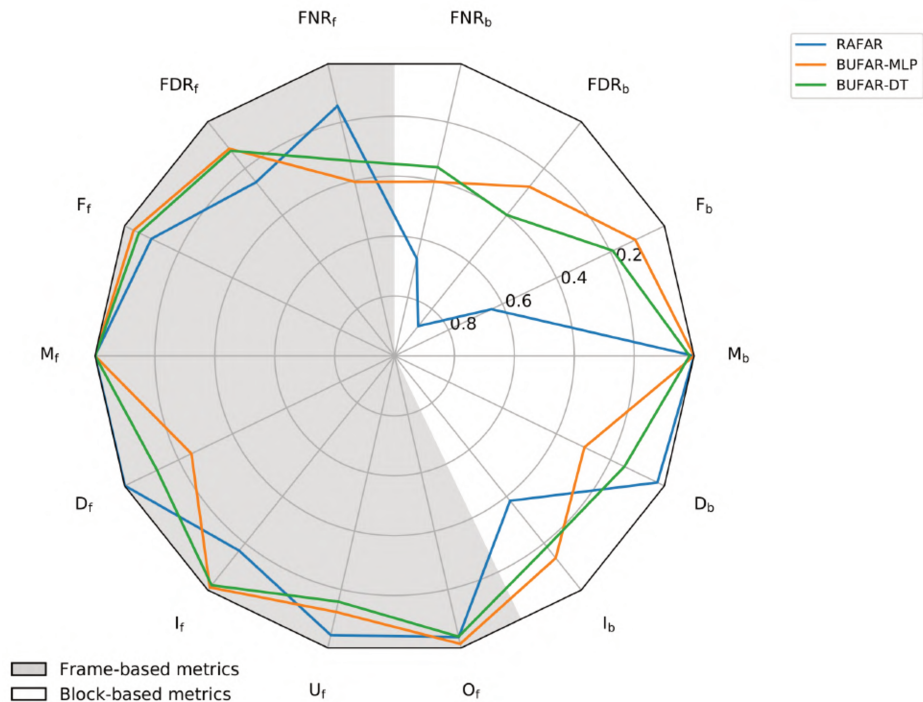


Figure 6: Spider plot of frame- and block-based metrics for grazing classification.

the proposed methods, the same limited data (5-min sound segments) was considered as the input.

A spider plot considering frame- and block-based metrics for grazing recognition is shown in Figure 6. A perfect recognizer would yield 0 for each error metric, which matches the boundary of the polygon. Frame-based metrics (gray side of the diagram) showed excellent FDR_f ($\sim 10\%$) and poor FNR_f ($< 40\%$) for both BUFAR variants. This means that most frames were correctly labeled as grazing, whereas some frames corresponding to grazing activity were not detected (false negatives). Deletions (D_f) and underfills (U_f) explain most of the undetected frames. The best FDR_f was achieved by BUFAR-MLP, while BUFAR-DT obtained a slightly lower FNR_f among variants. RAFAR presented the opposite situation, low FNR_f and high FDR_f . Regarding other metrics such as F_f , M_f , O_f , and I_f , the evaluated variants achieved excellent results ($< 5\%$), which indicates that hardly any frame is associated with fragmentation, merging, overfill, or insertion of grazing.

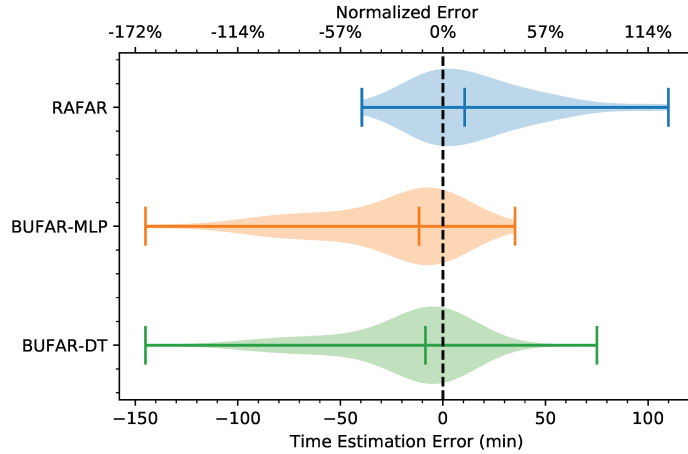


Figure 7: Violin plots of time estimation error for grazing classification. Normalized error considered the mean duration of grazing activity in the recordings (88.8 min).

Regarding the block-based analysis of grazing classification, BUFAR variants showed the lowest FDR_b and FNR_b and outperformed RAFAR on both metrics. BUFAR-MLP had slightly higher FNR_b but lower FDR_b than the BUFAR-DT. That is, BUFAR-MLP failed to detect some grazing block but added fewer extra grazing blocks (false positives) than BUFAR-DT. Both variants achieved low fragmentation ($F_b < 20\%$) and insertion ($I_b < 20\%$), and very low merging ($M_b < 5\%$). These are great improvements over RAFAR performance on the same metrics.

The errors on duration estimation of grazing activity are shown in the violin plots of Figure 7. The distribution of the errors is shown for both variants and the RAFAR, across all recordings analyzed. The medians of the distributions show a tendency to underestimation for BUFAR-MLP and BUFAR-DT, while RAFAR overestimated grazing. Between BUFAR variants, BUFAR-MLP achieved a lower dispersion. Absolute median errors were below 12.0 min.

Recognition results of rumination activity are shown in Figures 8 and 9. In the spider plot of Fig. 8, proposed variants achieved better FDR_f and FNR_f metrics than RAFAR. BUFAR-MLP showed the lowest FNR_f , which makes it the variant that detected most of the actual rumination bouts. Frames associated with fragmentation and merging of rumination bouts were very low and similar for both BUFAR variants. BUFAR-MLP achieved a

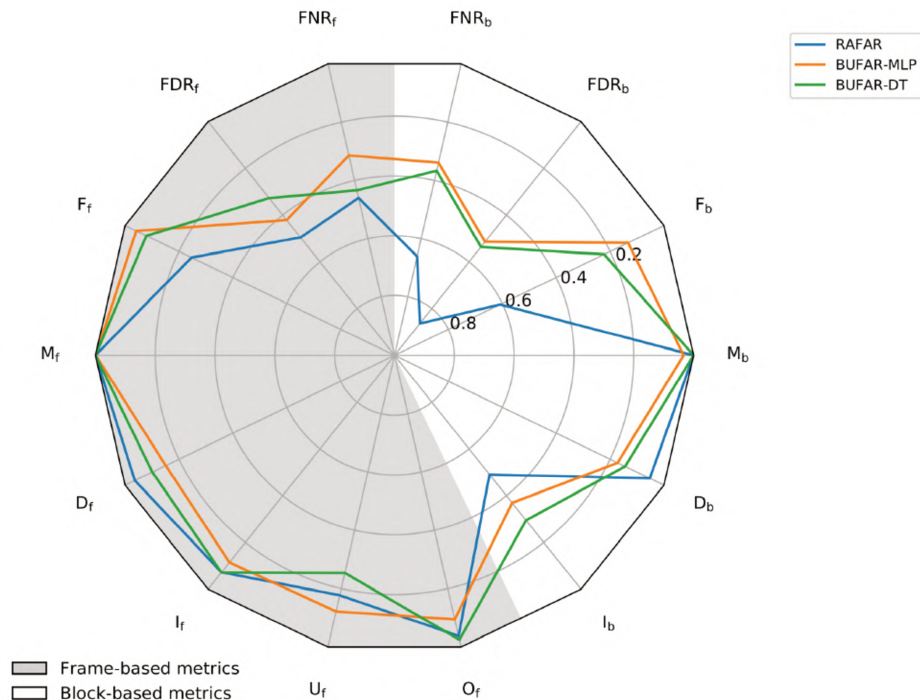


Figure 8: Spider plot of frame- and block-based metrics for rumination classification.

notable lower underfill error compared with BUFAR-DT.

Regarding block-based results (white side of the diagram), rumination recognition showed similar FNR_b and FDR_b for BUFAR variants. Even though, there was a small difference in favor of BUFAR-MLP. These performance metrics were much better compared to the results obtained with the RAFAR. Results indicate that rumination blocks were rarely fragmented or deleted by the proposed method. In addition, hardly any rumination block was merged.

Finally, the time estimation error on rumination activity is shown in Figure 9. The lowest median was achieved by BUFAR-MLP (0.3 min). Also, BUFAR-MLP showed a lower dispersion than BUFAR-DT.

3.3. Overall performance

A summary of the evaluated methods is shown in Table 2. As a general performance indicator, the F1-score was computed for the RAFAR and both BUFAR variants. For this global measure, BUFAR variants clearly

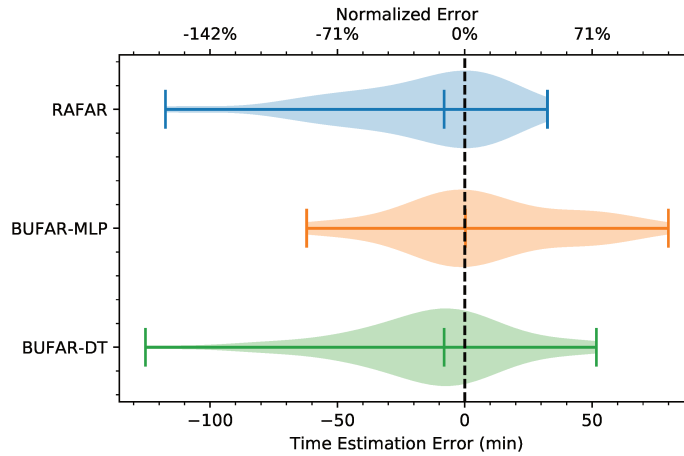


Figure 9: Time estimation error for rumination classification. Normalized error considered the mean duration of rumination activity in the recordings (69.9 min).

outperformed RAFAR for both grazing and rumination activities. This predominance is stronger on block-metrics, where 0.3 or higher improvements are seen. A comparison between the BUFAR variants showed similar results for grazing but a clear improvement for rumination in favor of BUFAR-MLP. Metrics differences between RAFAR and BUFAR variants has shown to be significant ($p < 0.05$) using a Wilcoxon signed-rank test (Wilcoxon, 1945). Thus, BUFAR-MLP achieved the best and most consistent results in recognition among studied activities.

In order to evaluate the feasibility of online execution, the computational cost of the proposed method was computed (Table 2). Also, the computational cost of the former method RAFAR was calculated for the sake of comparison. Notice that the cost of the two BUFAR variants is the same because the impact of the classifiers is negligible compared to the other stages. The computational cost of these variants is about 50 times lower than the corresponding to RAFAR. A detailed description of computations for RAFAR and BUFAR are provided in Appendices B and C, respectively.

4. Discussion

The proposed online method BUFAR showed a good performance in the estimation of grazing and rumination bouts. Frame- and block-based measures provide different points of view of the recognition. While frame-based

Table 2: F1-score metrics on activity classification and computational cost (operations per second) of analyzed methods.

	Grazing		Rumination		Computational cost (ops/s)
	Frame-based	Block-based	Frame-based	Block-based	
RAFAR	0.783 (± 0.180)	0.410 (± 0.267)	0.633 (± 0.240)	0.453 (± 0.327)	1,892,354
BUFAR-MLP	0.800 (± 0.236)	0.866 (± 0.165)	0.781 (± 0.230)	0.755 (± 0.289)	37,966
BUFAR-DT	0.795 (± 0.233)	0.819 (± 0.229)	0.661 (± 0.275)	0.734 (± 0.303)	37,966

metrics provide a fine-grain analysis which approximates to continuous time, block-based metrics provide information about the recognition of activities as blocks providing a big picture view of the recognition. In particular, BUFAR-MLP achieved frame- and block-based F1-scores higher than 0.75 (Table 2) This consistency among metrics and activities made it the preferred variant of the proposed method.

4.1. Comparison with a former method

The block-based metrics achieved by BUFAR were much higher than the corresponding ones to RAFAR. That is, more actual activity bouts were correctly recognized as activity blocks. Regarding time estimation of activities, the absolute errors were low for BUFAR variants (medians below 12 min) compared to the mean duration of activities (Figures 7 and 9). No significant differences were observed on time errors between the RAFAR and proposed variants. The time estimation error is a practical but ambiguous performance metric. False negatives frames could be compensated by false positives frames. Thus, in this study the estimation error has been complemented with the frame- and block-based metrics. These considerations support that the performance achieved by BUFAR is meaningful and makes auspicious its implementation on a portable device.

Foraging activity recognition throughout online processing of the acoustic signal is a main goal in this study. That is, the proposed method must process data within the device. As a consequence, only monitoring results need to be stored in the device until they can be transferred to a central server in a farm. BUFAR follows this approach by analyzing the sound signal in real-time. JM are identified in the moment and an activity segment is defined every 5 min. On the contrary, a method as the RAFAR is meant to perform

an offline processing, where an entire recording is required to obtain a proper result. The needs of massive volumes of data (several hours recordings) are not feasible for a limited device.

Another aspect to consider is the computational cost. Current microcontroller-based systems could operate at high frequency and perform heavy computations but at the expense of high power consumption. However, a method with low computational cost can be embedded in a microcontroller-based device working at low frequency and thus reducing the power consumption. This is essential for the development of a portable long-term monitoring device. The method proposed in this study requires 37,966 operations per second, which are much lesser than the operations required by RAFAR. Thus, BUFAR is truly suitable to perform online processing.

The use of fixed-length segments minimizes computational cost. A segment is classified into an activity by computing only a few operations every few minutes (segment length), when the segment buffer has been filled with the detected JM. Thus, computational cost is not increased and the device requirements are not modified. The use of this kind of segments is a design choice. Actual duration of foraging-activity bouts is expected to be similar to a multiple of segment length but not exactly the same. Duration mismatches exist, which affect the performance of the system. An alternative to the use of fixed-length segments would be dynamic segmentation, i.e, the length of each segment would be determined adaptively according to the features of the sound signal. However, it is expected that a dynamic segmentation approach would significantly increase the computational cost, which goes against the goal of this study. An intermediate approach is to consider a Markov process, where each segment is independent when given the previous one (Milone et al., 2012). Both approaches could be explored in order to improve the recognition performance, considering its corresponding computational cost and online implementation.

4.2. Comparison with a commercial system

A comparison of the rumination time estimation obtained by the Hi-Tag system and the BUFAR-MLP was performed. The Hi-Tag system summarizes the total time the animal spent ruminating during two-hour chunks (Schirmann et al., 2009). Raw data and timestamps of rumination bouts within a two-hour chunk are not available (Goldhawk et al., 2013). Therefore, the estimations with the BUFAR-MLP were aligned, and total duration

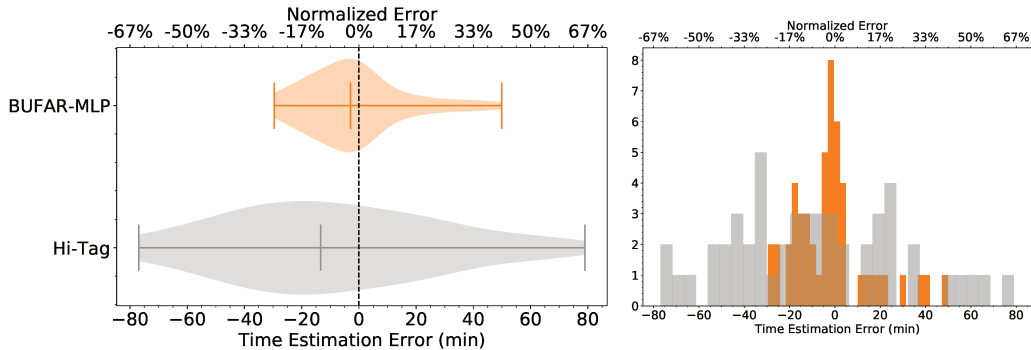


Figure 10: Time estimation error of rumination for BUFAR-MLP (orange) and Hi-Tag (gray). Top axis is normalized with the length of segments analyzed (2 hours).

of rumination was summarized to match the same two-hour chunks of the Hi-Tag system. The comparison was made with a total of 53 two-hour chunks from all the recordings analyzed as it was done in (Vanrell et al., 2018). Due that the Hi-Tag is a commercial system, its computational cost was not available to be considered in the analysis.

The results of time estimation error for rumination are shown in Fig.10. The medians of the distributions are -2.91 min and -13.55 min for BUFAR-MLP and the Hi-Tag system, respectively. Negative medians imply that both systems tend to underestimate the rumination time. The underestimation shown by the Hi-tag agree with previous studies that evaluate its performance (Burfeind et al., 2011; Goldhawk et al., 2013). BUFAR-MLP was more accurate and resulted in a narrower error distribution. While the error dispersion for BUFAR-MLP is in the range (-30, +50) min, the distribution corresponding to the Hi-Tag is wider and it is in the range (-80, +80) min. In practical terms, these errors are very high since they are in the same order of magnitude of the two-hour chunks analyzed. The wide dispersion shown by the Hi-tag has been seen in previous studies (Burfeind et al., 2011; Goldhawk et al., 2013).

5. Conclusions

In this study, an online method for recognition and estimation of foraging activity bouts from acoustic signals has been presented. The proposed method BUFAR follows a bottom-up approach, which goes from jaw movement recognition to foraging activity recognition. Sound signals are processed

and downsampled to operate at a lower frequency, aiming at the implementation of the method in a microcontroller-based system with limited resources. The recognition of grazing and rumination bouts was evaluated with specific metrics for activity recognition. Analyzing the results, the preferred variant of the proposed method is the BUFAR-MLP and medium-length segments. In addition, the BUFAR-MLP was superior in comparison with the former method RAFAR. Another important advantage is that the proposed method performs very few operations to recognize activity bouts. This ease the possibility of an online implementation for its execution on a low-cost embedded system. An additional comparison showed that the proposed method outperformed the Hi-Tag commercial system on rumination time estimation. Thus, the BUFAR good performance and simplicity achieved the stated goals. Future works could be focused on improving the recognition performance by including more complex features or processing techniques at the expense of an increased computational cost.

Acknowledgments

This study has been funded by Universidad Nacional del Litoral, PACT CAID 2011-525, Universidad Nacional de Rosario, projects 2013-AGR216, 2016-AGR266 and 80020180300053UR, Agencia Santafesina de Ciencia, Tecnología e Innovación (ASACTEI), project IO-2018-00082, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), project 2017-PUE-sinc(i). Also, this study was funded under the USDA-NIFA MICL0222 and MICL0406 projects, and direct support from AgBioResearch-MSU. Authors wish to thank the insightful help and dedication by the KBS Robotic Dairy Farm staff for their assistance and support during the completion of this study.

Appendix A. Definitions of frame- and block-based error metrics

The frame- and block-based error metrics are defined in Table A.1. Frame-based metrics are defined by considering the counts of true positives TP , false positives FP , false negatives FN , fragmented F , merged M , deleted D , and underfill U frames in the reference sequence, and by the count of inserted I , and overfill O frames in the recognized sequence, respectively. Frames of 1 s were considered as the smallest time unit for analysis. Block-based metrics are defined by considering the counts of total (B^{ref}), correctly detected

(C), fragmented (F), merged (M), and deleted (D) blocks in the reference sequence, and by the counts of total (B^{rec}) and inserted (I) blocks in the recognized sequence, respectively. In addition, the standard F1-score was computed for frames $F1_f = \frac{2R_fP_f}{R_f+P_f}$ and blocks $F1_b = \frac{2R_bP_b}{R_b+P_b}$ based on the corresponding precision and recall defined in Table A.1.

Table A.1: Definitions of frame- and block-based error metrics.

Error metric	Frame-based	Block-based
False negative rate	$FNR_f = 1 - \frac{TP}{TP+FN} = 1 - R_f$	$FNR_b = 1 - \frac{C}{B^{ref}} = 1 - R_b$
False discovery rate	$FDR_f = 1 - \frac{TP}{TP+FP} = 1 - P_f$	$FDR_b = 1 - \frac{C}{B^{rec}} = 1 - P_b$
Fragmentation	$F_f = \frac{F}{TP+FP}$	$F_b = \frac{F}{B^{ref}}$
Merging	$M_f = \frac{M}{TP+FP}$	$M_b = \frac{M}{B^{ref}}$
Deletion	$D_f = \frac{D}{TP+FP}$	$D_b = \frac{D}{B^{ref}}$
Insertion	$I_f = \frac{I}{TP+FN}$	$I_b = \frac{I}{B^{rec}}$
Overfill	$O_f = \frac{D}{TP+FP}$	–
Underfill	$U_f = \frac{I}{TP+FN}$	–

Appendix B. Computational cost of RAFAR

The computational cost of RAFAR-MBBP (Vanrell et al., 2018) depends on the sampling frequency (S_f) and duration (T) of the input signal. In order to get a straightforward comparison with other algorithms, a sampling frequency of $S_f = 2$ kHz and a duration of $T = 300$ s were selected to compute the computational cost. Worst-case scenarios were considered for each stage in order to get a theoretical upper bound.

The required number of operations per stage of computation for RAFAR-MBBP was:

1. Segmentation by regularity

- (a) *Envelope computation*: This task comprise signal rectification, signal filtering, and signal subsampling. First, signal rectification requires a comparison and a multiplication per sample. Second, a 3rd-order IIR low-pass filter is applied, which involves 7 multiplications and 6 additions per sample. Third, the envelope is sub-sampled at 1 kHz, which requires 1,000 comparisons/s.
- (b) *Regularity analysis*: The envelope is analyzed by frames of 30 s. The computation of autocorrelation requires $29.225 \cdot (S_f/2) \cdot 951$

multiplications and $[29.225 \cdot (S_f/2) - 1] \cdot 951$ additions for each 30 s frame. Then, a peak is searched, which requires 12,264 comparisons for each 30 s frame. Once a peak is found, the regularity rule is evaluated with two comparisons and the frame is labeled in one assignment.

- (c) *Smoothing filter*: A 5th-order median filter is implemented, which involves 10 comparisons for each 30 s frame.

The computational cost of the segmentation stage is 565,272,760 operations.

2. Classification of activity blocks

- (a) *Energy computation*: This task is performed using 1 s frames and requires $2 \cdot S_f + 4$ multiplications, $2 \cdot S_f + 2$ additions, and 3 assignments per frame.
- (b) *Sudden-drop detection*: Worst case scenario considers an 80 s sliding window with a 5 s step. The median of the energy is computed with 507 comparisons per window. A threshold is generated and compared requiring 1 multiplication and 1 comparison per window.
- (c) *Rules classification*: This task required 4 comparisons for each activity block.

The computational cost of this stage is 1,233,244 operations.

- 3. **Block partition**: Worst-case scenario for this stage is to consider the input signal as a single block. Computation of block duration requires 1 subtraction. A block is analyzed if the duration is greater than 10 min, which requires 1 comparison. A block is analyzed with 60 s frames. Energy is computed requiring $2 \cdot 60 \cdot S_f + 4$ multiplications, $2 \cdot 60 \cdot S_f + 2$ additions, and 3 assignments per frame. The detection of changes in the computed energy requires 1 multiplication and 1 comparison with a threshold, for each 60 s frames. If a block should be partitioned, 2 extra assignments are required. Therefore, the computational cost of this stage is 1,200,059 operations.
- 4. **Merging gaps**: The worst-case scenario for this stage is to consider that the entire input signal has the shortest activity blocks and the shortest inactivity gaps. A subtraction is required to compute the duration of the gap and it is compared with a threshold. If a gap should be merged, 3 extra assignments are required. Therefore, the computational cost of this stage is 9 operations.

The overall computational cost for the RAFAR-MBBP is: $565,272,760 + 1,233,244 + 1,200,059 + 9 = 567,706,072$ operations. The most computational-expensive stage is the segmentation, which requires 99.57% of the total operations. Specifically, the autocorrelation computation requires 97.91% of the total operations.

To compare the RAFAR-MBBP with an online method and considering the duration of the input signal (300 s), the computational cost can be estimated as 1,892,354 operations/s.

Appendix C. Computational cost of BUFAR

The computational cost of BUFAR depends on the sampling frequency (fixed in 2 kHz in this analysis) and the duration (fixed in $T = 300$ s in this analysis) of the input signal. A 5 min segment length and 2 jaw movements per second were selected in order to consider the worst-case scenario in the sense of computational cost. The required number of operations per second for the computation stages of BUFAR was:

1. **Signal pre-processing:** A least mean square filter (LMS) requires 5 operations per signal sample. Then, 10,000 operations/s are required.
2. **Jaw-movement detection:** 27,800 operations/s are required to detect jaw movements and to extract their features.
3. **Jaw-movement classification:** MLP requires 80 operations per jaw movement, thus, 160 operations/s are required.
4. **Segment buffering:** this stage requires 6 operations/s and 6 operations per segment to save the timestamp and to compute the segment features.
5. **Activity classification:** this stage was evaluated for MLP and DT. MLP requires 170 operations per segment. DT requires 6 operations per segment.
6. **Smoothing process:** to avoid fragmentation in rumination and grazing bouts, 2 comparisons per segment are required.

Hence, the overall computational cost is 37,966 operations/s + 178 operations/segment for BUFAR-MLP, and 37,966 operations/s + 14 operations/segment for BUFAR-DT. The costs of activity classification and smoothing process are negligible because the operations are performed just a few times in a long period of time (segment length).

References

- Andriamandroso, A., Bindelle, J., Mercatoris, B., Lebeau, F., 2016. A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnology, Agronomy, Society and Environment* 20.
- Andriamandroso, A.L.H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B., Dumortier, P., Blanchy, G., Blaise, Y., Bindelle, J., 2017. Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors. *Comput. Electron. Agric.* 139, 126–137.
- Arcidiacono, C., Porto, S.M.C., Mancino, M., Cascone, G., 2017. Development of a threshold-based classifier for real-time recognition of cow feeding and standing behavioural activities from accelerometer data. *Comput. Electron. Agric.* 134, 124–134.
- Beauchemin, K.A., 1991. Ingestion and mastication of feed by dairy cattle. *Vet. Clin. North Am. Food Anim. Pract.* 7, 439–463.
- Berkmans, D., 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. *Rev. Sci. Tech.* 33, 189–196.
- Bishop, C.M., 2006. *Pattern Recognition and Machine Learning*. Springer Verlag.
- Bristow, D.J., Holmes, D.S., 2007. Cortisol levels and anxiety-related behaviors in cattle. *Physiol. Behav.* 90, 626–628.
- Burfeind, O., Schirmann, K., von Keyserlingk, M., Veira, D., Weary, D., Heuwieser, W., 2011. Evaluation of a system for monitoring rumination in heifers and calves. *Journal of dairy science* 94, 426–430.
- Chelotti, J.O., Vanrell, S.R., Galli, J.R., Giovanini, L.L., Leonardo Rufiner, H., 2018. A pattern recognition approach for detecting and classifying jaw movements in grazing cattle. *Comput. Electron. Agric.* 145, 83–91.
- Chelotti, J.O., Vanrell, S.R., Milone, D.H., Utsumi, S.A., Galli, J.R., Rufiner, H.L., Giovanini, L.L., 2016. A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle. *Comput. Electron. Agric.* 127, 64–75.

- Clapham, W.M., Fedders, J.M., Beeman, K., Neel, J.P.S., 2011. Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle. *Comput. Electron. Agric.* 76, 96–104.
- Galli, J.R., Cangiano, C.A., Pece, M.A., Larripa, M.J., Milone, D.H., Utsumi, S.A., Laca, E.A., 2018. Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *Animal* 12, 973–982.
- Giovanetti, V., Decandia, M., Molle, G., Acciaro, M., Mameli, M., Cabiddu, A., Cossu, R., Serra, M.G., Manca, C., Rassu, S.P.G., Dimauro, C., 2017. Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer. *Livest. Sci.* 196, 42–48.
- Goldhawk, C., Schwartzkopf-Genswein, K., Beauchemin, K.A., 2013. Validation of rumination collars for beef cattle. *J. Anim. Sci.* 91, 2858–2862.
- González, L.A., Bishop-Hurley, G.J., Handcock, R.N., Crossman, C., 2015. Behavioral classification of data from collars containing motion sensors in grazing cattle. *Comput. Electron. Agric.* 110, 91–102.
- Greenwood, P.L., Paull, D.R., McNally, J., Kalinowski, T., Ebert, D., Little, B., Smith, D.V., Rahman, A., Valencia, P., Ingham, A.B., Bishop-Hurley, G.J., 2017. Use of sensor-determined behaviours to develop algorithms for pasture intake by individual grazing cattle. *Crop Pasture Sci.* .
- Hansen, S.S., Nørgaard, P., Pedersen, C., Jørgensen, R.J., Mellau, L.S.B., Enemark, J.D., 2003. The effect of subclinical hypocalcaemia induced by Na₂EDTA on the feed intake and chewing activity of dairy cows. *Vet. Res. Commun.* 27, 193–205.
- Herskin, M.S., Munksgaard, L., Ladewig, J., 2004. Effects of acute stressors on nociception, adrenocortical responses and behavior of dairy cows. *Physiol. Behav.* 83, 411–420.
- Hodgson, J., Illius, A.W., 1998. The ecology and management of grazing systems. Wallingford (United Kingdom) CAB International.
- Hodgson, J.G., 1990. *Grazing management: science into practice*. John Wiley & Sons Inc.

- Kilgour, R.J., 2012. In pursuit of “normal”: A review of the behaviour of cattle at pasture. *Appl. Anim. Behav. Sci.* 138, 1–11.
- Laca, E.A., Ungar, E.D., Seligman, N.G., Ramey, M.R., Demment, M.W., 1992. An integrated methodology for studying short-term grazing behaviour of cattle. *Grass Forage Sci.* 47, 81–90.
- Martiskainen, P., Järvinen, M., Skön, J.P., Tiirikainen, J., Kolehmainen, M., Mononen, J., 2009. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl. Anim. Behav. Sci.* 119, 32–38.
- Milone, D.H., Galli, J.R., Cangiano, C.A., Rufiner, H.L., Laca, E.A., 2012. Automatic recognition of ingestive sounds of cattle based on hidden markov models. *Comput. Electron. Agric.* 87, 51–55.
- Navon, S., Mizrach, A., Hetzroni, A., Ungar, E.D., 2013. Automatic recognition of jaw movements in free-ranging cattle, goats and sheep, using acoustic monitoring. *Biosystems Eng.* 114, 474–483.
- Nydegger, F., Gyga, L., Egli, W., 2010. Automatic measurement of rumination and feeding activity using a pressure sensor, in: *International Conference on Agricultural Engineering - AgEng 2010*.
- Paudyal, S., Maunsell, F.P., Richeson, J.T., Risco, C.A., Donovan, D.A., Pinedo, P.J., 2018. Rumination time and monitoring of health disorders during early lactation. *Animal* 12, 1484–1492.
- Phillips, C.J.C., 1993. *Cattle Behaviour*. Farming Press.
- Rutter, S.M., 2000. Graze: a program to analyze recordings of the jaw movements of ruminants. *Behav. Res. Methods Instrum. Comput.* 32, 86–92.
- Rutter, S.M., Champion, R.A., Penning, P.D., 1997. An automatic system to record foraging behaviour in free-ranging ruminants. *Appl. Anim. Behav. Sci.* 54, 185–195.
- Schirmann, K., von Keyserlingk, M.A.G., Weary, D.M., Veira, D.M., Heuwieser, W., 2009. Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.* 92, 6052–6055.

- Smith, D., Rahman, A., Bishop-Hurley, G.J., Hills, J., Shahriar, S., Henry, D., Rawnsley, R., 2016. Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Comput. Electron. Agric.* 131, 40–50.
- Stegmayer, G., Pividori, M., Milone, D.H., 2016. A very simple and fast way to access and validate algorithms in reproducible research. *Brief. Bioinform.* 17, 180–183.
- Ungar, E.D., Rutter, S.M., 2006. Classifying cattle jaw movements: Comparing IGER behaviour recorder and acoustic techniques. *Appl. Anim. Behav. Sci.* 98, 11–27.
- Vanrell, S.R., Chelotti, J.O., Galli, J.R., Utsumi, S.A., Giovanini, L.L., Rufiner, H.L., Milone, D.H., 2018. A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *Comput. Electron. Agric.* 151, 392–402.
- Ward, J.A., Lukowicz, P., Gellersen, H.W., 2011. Performance metrics for activity recognition. *ACM Trans. Intell. Syst. Technol.* 2, 6:1–6:23.
- Watt, L., Clark, C., Krebs, G., Petzel, C., Nielsen, S., Utsumi, S., 2015. Differential rumination, intake, and enteric methane production of dairy cows in a pasture-based automatic milking system. *Journal of dairy science* 98, 7248–7263.
- Welch, J.G., 1982. Rumination, particle size and passage from the rumen. *J. Anim. Sci.* 54, 885.
- Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., O’Brien, B., 2018. Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows. *J. Neurosci. Methods* 300, 138–146.
- Widrow, B., Glover, J.R., McCool, J.M., Kaunitz, J., Williams, C.S., Hearn, R.H., Zeidler, J.R., Dong, J.E., Goodlin, R.C., 1975. Adaptive noise cancelling: Principles and applications. *Proc. IEEE* 63, 1692–1716.
- Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bulletin* 1, 80.

Zehner, N., Umstätter, C., Niederhauser, J.J., Schick, M., 2017. System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *Comput. Electron. Agric.* 136, 31–41.

Anexo C

**Using jaw movement-agnostic features
to recognize foraging activities in
grazing cattle**

Using segment-based features of jaw movements to recognize foraging activities in grazing cattle

José O. Chelotti^a, Sebastián R. Vanrell^a, Luciano S. Martinez-Rau^a,
Julio R. Galli^{b,e}, Santiago A. Utsumi^c, Alejandra M. Planisich^b,
Suyai A. Almirón^b, Diego H. Milone^a, Leonardo L. Giovanini^a,
H. Leonardo Rufiner^{a,d}

^a*Instituto de Investigación en Señales, Sistemas e Inteligencia Computacional, sinc(i),
FICH-UNL/CONICET, Argentina*

^b*Facultad de Ciencias Agrarias, Universidad Nacional de Rosario, Argentina*

^c*W.K. Kellogg Biological Station and Department of Animal Science, Michigan State
University, United States*

^d*Facultad de Ingeniería, Universidad Nacional de Entre Ríos, Argentina*

^e*Instituto de Investigaciones en Ciencias Agrarias de Rosario, IICAR, UNR-CONICET,
Argentina*

Abstract

Precision livestock farming optimizes livestock production through the use of sensor information and communication technologies to support decision making, proactively and near real-time. Among available technologies to monitor foraging behavior, the acoustic method has been highly reliable and repeatable, but can be subject to further computational improvements to increase precision and specificity of recognition of foraging activities. In this study, an algorithm called Jaw Movement segment-based Foraging Activity Recognizer (JMFAR) is proposed. The method is based on the computation and analysis of temporal, statistical and spectral features of jaw movement sounds for detection of rumination and grazing bouts. They are called JM-

Email address: jchelotti@sinc.unl.edu.ar (José O. Chelotti)

segment features because they are extracted from a sound segment and expect to capture JM information of the whole segment rather than individual JMs. Two variants of the method are proposed and tested: (i) the temporal and statistical features only JMFAR-ns; and (ii) a feature selection process (JMFAR-sel). The JMFAR was tested on signals registered in a free grazing environment, achieving an average weighted F1-score greater than 95%. Then, it was compared with a state-of-the-art algorithm, showing improved performance for estimation of grazing bouts (+19%). The JMFAR-ns variant reduced the computational cost by 25.4%, but achieved a slightly lower performance than the JMFAR. The good performance and low computational cost of JMFAR-ns supports the feasibility of using this algorithm variant for real-time implementation in low-cost embedded systems.

Keywords: Acoustic monitoring, ruminant foraging behavior, precision livestock farming, pattern recognition, feature engineering, machine learning.

1. Introduction

Precision livestock farming offers the unprecedented opportunity to improve livestock management by switching the scale of observation from herds to individuals. The individualized monitoring with regard to animal behavior could provide rich and unique information on collective events related to the health, nutrition and welfare status of animals (Hodgson and Illius, 1998). The feasibility for deployment of animal sensing and information systems on farms (Andriamandroso et al., 2016) has been enhanced recently by new developments in animal-borne sensors, advanced communication and pro-

cessing technologies, enhanced data analytics and visualization techniques for easy inspection of animal behavior, including behavior events and activities, location and posture (Berckmans, 2014). However, the operational requirements for such monitoring systems, primarily device portability, robustness and power, along with the computational cost and complexity of required algorithms often emerges as major bottlenecks for further technological progresses (Stone, 2020).

For grazing cattle, feeding and rumination usually represent the largest fraction of the daily time budget and close inspection of their longitudinal changes could be used to accurately assess animal behavior to optimize animal and pasture management. For example, decreased daily grazing and/or rumination time could be interpreted as an indicator of severe environmental stress (Herskin et al., 2004), anxiety (Bristow and Holmes, 2007), or emerging diseases (Welch, 1982), whereas increasing daily activity and decreasing rumination time could flag the onset of estrus (Schirmann et al., 2009) or calving events (Schirmann et al., 2013). Similarly, increasing rumination time is associated with more saliva production and improved rumen and animal health (Beauchemin, 1991), whereas decreasing rumination would flag a lacking effective fiber in the diets of cows (Beauchemin, 2018). Grazing encloses behaviors related to feed search apprehension and consumption and is composed by three types of distinct jaw movements (JM): biting, when herbage is apprehended and severed; chewing, when herbage is comminuted; and chew-biting when a combination chewing and biting of grass occurs in a same JM (Galli et al., 2018; Laca et al., 1992; Ungar and Rutter, 2006). Rumination is characterized exclusively by feed chewing occurring within regular

feed swallowing-regurgitation cycles. One of the distinctive characteristics of grazing and rumination is the hierarchical array of foraging activities and duration and sequence of specific JM. Whereas chewing and biting in a single JM last about a second or a fraction of a second, the sequence of JM that defines a grazing or rumination bout usually expands from several minutes to hours (Andriamandroso et al., 2016).

There are several remote sensing methods to monitor foraging behavior. Among most popular methods are the accelerometers (Arcidiacono et al., 2017; Giovanetti et al., 2017; González et al., 2015; Martiskainen et al., 2009), inertial measurement units (Andriamandroso et al., 2017; Greenwood et al., 2017; Smith et al., 2016) and nose-band pressure sensors (Nydegger et al., 2010b; Werner et al., 2018b; Zehner et al., 2017a). The two methods are easy to deploy and provide easy straightforward information of whole body movements or movement of specific body parts, such as leg, neck or head movements. Head-placed microphones are usually preferred over nose-band pressure sensors and accelerometers (Nydegger et al., 2010a; Rutter, 2000; Rutter et al., 1997; Werner et al., 2018a; Zehner et al., 2017b) for precision monitoring of feeding and rumination behavior.

The use of sound to recognize bites, chews, and chew-bites and their individual characteristics allows to obtain precise ingestive information and also to accurately estimate dry matter intake and identify the forage species consumed (Galli et al., 2018, 2020). By contrast, pressure sensors and accelerometers present diminished performance on automatic monitoring of ingestive behavior (Li et al., 2021; Werner et al., 2019).

The acoustic method has obtained very good results in recognizing and

classifying JM and cattle grazing foraging activities. Vanrell et al. (2018) proposed the off-line acoustic method called regularity-based acoustic foraging activity recognizer (RAFAR) that relied on autocorrelation for recognition of grazing and rumination bouts. The main disadvantage of this algorithm is the requirement of the entire sound record to achieve satisfactory performance results. More recently, Chelotti et al. (2020) introduced BUFAR, an alternative algorithm to RAFAR to enable on-line processing of sound signals. Both, RAFAR and BUFAR achieved better performance compared to commercial rumination collars (Chelotti et al., 2020; Vanrell et al., 2018), yet the two also showed a noticeable misclassifications between grazing and rumination bouts.

The main objective of this work is to develop and test a new method for improved classification of grazing and rumination bouts in free grazing cattle. The new method will use broad sets of temporal, statistical and spectral features of the sound signal and it is therefore expected to minimize previous grazing and rumination misclassifications, without significantly incurring in additional computational costs. The new features are called JM-segment features because they are extracted from a sound segment to capture JM insight of the whole segment rather than individual JMs. The main benefits for the new method are: (i) the grazing and rumination classification is independent of the JM classification, thus avoiding the errors associated with this stage (Chelotti et al., 2018); and (ii) the ability to train models without the need of using previously labeled JM, which is a significant improvement that enable to train the method on large recording datasets. The paper is organized as follows: Section 2 introduces the JMFAR algorithm and the

new set of features. Databases used for model testing, the variants of the proposed method and the measures of performance to assess model suitability are described in Section 3. Section 4, describes the main study results with a focused discussion presented in Section 5. Finally, a brief conclusion with regard to the study objective and main results is presented in Section 6.

2. Material and methods

2.1. Background: Current acoustic method

In a previous work, Chelotti et al. (2018) introduced a JM recognition algorithm that estimates sequences of JM using the sounds produced by ruminants during foraging activities. It is comprised of three distinct elements: (i) a signal conditioning stage, which attenuates the effects of environmental noises and disturbances, (ii) a set of extracted acoustical features and (iii) a machine learning model that provides the algorithm with excellent discrimination capabilities. Based on this algorithm, a foraging activity recognizer called BUFAR was proposed by Chelotti et al. (2020). It analyses groups of JM to recognize grazing and rumination bouts. It is comprised of (i) a buffer segment, which comprises a sequence of estimated JM along a fixed period of time, (ii) a set of descriptive statistical features for JM in the buffer that includes JM rate and proportions for JM types, and (iii) a machine learning model to enable discrimination capabilities of the algorithm. Accordingly, rumination bouts were identified and determined by regular cycles of thirty to sixty pure chews, occurring at a near constant rate of one JM per second, and followed by brief sound pauses for feed cud swallowing and regurgitation (Galli et al., 2020). Conversely grazing bouts were determined by uneven

mixtures of bites, chews and chew-bites sequences that had no particular order and occurred at near constant rate of one JM per second (Chelotti et al., 2020).

2.2. Recognition system

A block diagram of the proposed system is shown in Figure 1. It consists of a JM recognizer (Chelotti et al., 2018) and an activity recognizer, which is inspired by BUFAR architecture. To improve the performance of the activity recognizer, a novel set of discriminative features were developed for the activity classifier. The system analyses the acoustic signal in non-overlapping segments of 5 min length to determine the type of activity performed in the segment. First, in a pre-processing stage the acoustic signal is conditioned to remove noises and disturbances, thereby improving the signal-to-noise ratio. Also the sound envelope is obtained after a decimation process is applied. Then, the candidate JM is detected (JM detection stage) in order to divide the sound envelope, isolate the JM, and extract the JM features. The signal envelope, the timestamp of the detected JM and the JM features are stored in the segment buffer, which is then processed to extract the novel set of activity features. These features are analyzed by the activity classifier to determine the activity performed by the animal within the time segment. The new features are called JM-segment features because they correspond to a sound segment and capture JM insight of the whole segment rather than individual JMs features. Their computation only requires the detection of events, regardless of their type (i.e. bite, chew or chew-bite). Unlike the BUFAR method, this approach has the advantage of being insensitive to JM classification errors. Although the JM type is not used for the system, it

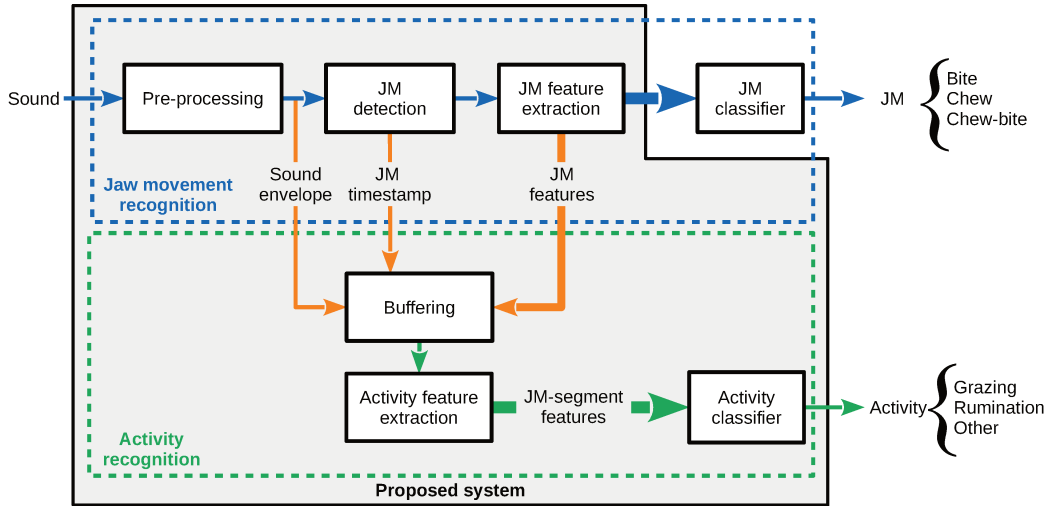


Figure 1: General diagram of the architecture. Blocks inside the gray box are part of the proposed system (JMFAR). The thickness of the arrows indicates the size of the feature vector.

might be obtained as an extra output from the JM classifier. Both classifiers (JM and activity) are implemented using a multilayer perceptron (Bishop, 2006). For both classifiers, the optimized hyperparameters were the learning rate and the number of neurons in the hidden layer. For the activity classifier, the MLP structure consisted of 21 neurons in the input layer, 20 in the hidden layer and 3 in the output layer.

2.3. Proposed activity features

In order to improve recognition ability, we propose a set of temporal, statistical and spectral features of JM for the foraging activity recognition called JM-segment features. Previous JM features included in BUFAR were

the proportion of chews (f1)¹, the proportion of bites (f2), the proportion of chew-bites (f3) and JM rate (f4), which may contain relatively low information and discriminatory power. The new set of JM-segment features is composed by:

- JM rate : is the proportion of detected events (f4) within a given sound signal segment.
- Statistics of JM amplitudes : includes mean (f5), standard deviation (f6), skewness (f7) and kurtosis (f8) for maximum amplitudes of JM that occur within a given sound signal segment.
- Statistics of JM durations : including mean (f9), standard deviation (f10), skewness (f11) and kurtosis (f12) for duration of JM that occur within a given sound signal segment.
- Spectral feature of the sound envelope : defined as energy in the 1.0-to-1.5 Hz band (f13) relative to the total energy of a given sound signal segment.
- JM jitter variants : variation of the period between two consecutive JM within a given signal segment. Alternative variants for this parameter include: absolute jitter (f14), relative jitter (f15), five-point period perturbation quotient (PPQ5) jitter (f16), and standard deviation of jitter (f17).

¹The number between parentheses indicates the ordinal number of the feature within the set.

- JM shimmer variants : variation of maximum amplitude between two adjacent JM within a given signal segment. Alternative variants for this parameter include: absolute shimmer (f18), relative shimmer (f19), five-point amplitude perturbation quotient (APQ5) shimmer (f20) and standard deviation of shimmer (f21).
- Temporal features of the JM tachogram : patterns of regular JM sequences are used for activity classification. This feature indicates the rate for JM intervals with lengths between 3 and 10 s (f22).
- Spectral features of the JM tachogram : is the energy in the 0.017-to-0.020 Hz (f23) and 0.00-to-0.02 Hz (f24) bands relative to the total energy of the tachogram signal and provides information about the swallowing-regurgitation events.

A qualitative analysis of the effect of proposed features on class separation was performed. Figure 2 shows the result of a dimensionality reduction analysis (Van der Maaten and Hinton, 2012) applied to the set of JM-segment features proposed in this work. Grazing and rumination are grouped in clear clusters without mixed data points between them. A few proportions of activities other than grazing or rumination are minimally mixed with rumination data points.

2.3.1. Statistical features for amplitude and duration

The foraging behavior of ruminants can be characterized by the temporal sequence of JM, with sounds that vary according to given animal factors, foraging events and biochemical and physical properties of the feed eaten.

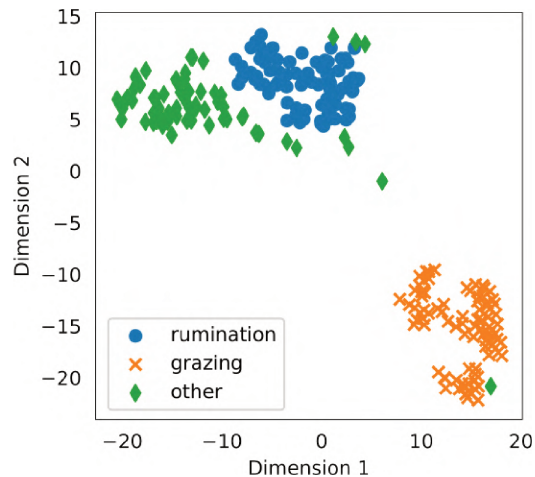


Figure 2: t-SNE analysis of JM-segment features.

For example, bites are characterized by a loud sound with a high amplitude peak, when forage is apprehended between incisors and the dental pad, and severed. Grazing chews produce loud sounds with lower intensity than bites. Composite chew-bites combine sound characteristics of grazing chews (first) and bites (later) during the same JM. Additionally, whereas exclusive bites and chews have a similar short duration, chew-bites have longer sound signals (Chelotti et al., 2018). Since grazing is composed by a mixed sequence of bites, chews and chew-bites, the amplitude frequency distribution peaks and spreads over a broader high amplitude region (Figure 3). The event duration distribution of grazing has a peak on the long duration region and it is distributed over the whole range of values. Conversely, ruminating chews are characterized by very low intensity sounds, associated probably to the wetness, fragmentation and partial degradation of feed being processed (rumen turnover is at least fifteen hours). Since rumination is composed of exclusive rumination chews, the amplitude frequency of rumination peaks in the

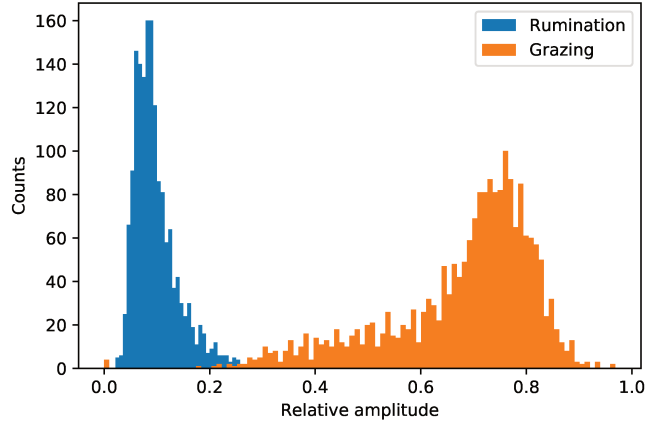


Figure 3: Histogram showing the amplitude distribution of JM associated with grazing and rumination.

very low amplitude region and is narrowly distributed around peak frequency (Figure 3). The duration histogram of rumination has similar characteristics to the amplitude histogram. Due to the asymmetric nature of amplitude and duration histograms, several moments (mean, standard deviation, skewness and kurtosis) were determined and considered as features.

2.3.2. Spectrum of the sound envelope

The spectrum of the sound envelope helps to discriminate among, grazing, rumination and other activities. This spectrum characterizes the temporal behavior of events according to the periodicity and variability of the type of activity an animal performs. The spectra of rumination sound segments show a very narrow frequency band determined by the low periodicity and low amplitude of rumination chew. On the other hand, the spectra of grazing sound segments show a broader frequency band determined by the irregular

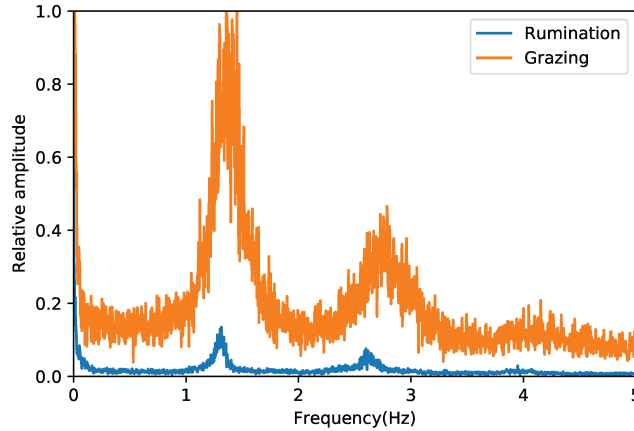


Figure 4: Spectra of grazing and rumination sound segments.

and mixed nature and higher amplitude of chews, bites and composite chew-bites during grazing.

Figure 4 shows the average spectra of typical sound segments corresponding to grazing (orange) and rumination (blue). The peaks of relative amplitude correspond to the fundamental frequency (and its harmonics) of JM, which could be useful to differentiate among grazing, rumination, and other activities. Further, the ratio between the energy in the band 1.0-to-1.5 Hz and the total sound segment energy was proposed as an additional discrimination feature. In this context, activities other than grazing and rumination are not expected to present this characteristic spectral pattern, as such activities may include periods of silence or different types of environmental noise.

2.3.3. Tachogram

A ruminant typically performs a bundle of JM per minute (between thirty to sixty JM) and the rate variability of JM depends on the activity being performed. The timestamps of detected events can be arranged into tachograms showing the inter-event periodicity time-series (Figure 5). Rumination is composed by groups of regular chews followed by periodic pauses of several seconds for feed bolus swallowing and regurgitation. Consequently, the rate variability of JM during rumination will show a great frequency of values around one second duration (the period between successive JMs) with periodic peaks of two second duration or greater for pauses of swallowing and regurgitation (Figure 5.b). On the other hand, grazing is composed by a sequence of regular JMs and less frequent and irregularly distributed long pauses associated with walking or food searching (Figure 5.a). Thus, the rate variability of JM for grazing shows a great frequency of inter-events periods around one second (similar inter-events between successive JMs) with occasional peaks of longer duration associated with exclusive feed searching (Bailey et al., 1996).

In order to obtain a frequency representation, Fast Fourier Transform (FFT) cannot be applied directly because it is an unevenly spaced signal, and it requires a prior transformation. Signal resampling by interpolation is the common strategy. Once the spectrum is estimated, it is separated into frequency bands, which are related to the JMs and swallowing-regurgitation pauses during rumination.

Figure 6 shows typical spectra of tachograms for grazing and rumination. The spectra show clear differences in the frequency band between 0 Hz and

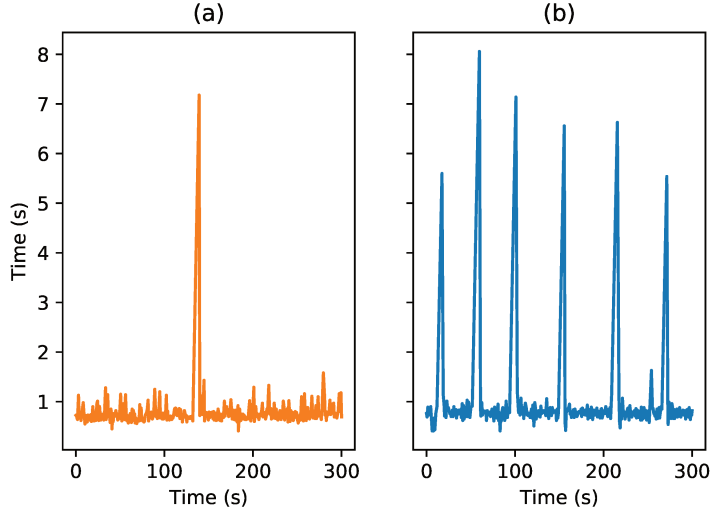


Figure 5: Tachograms of twenty minutes of (a) grazing and (b) rumination.

0.02 Hz. These differences are caused by the swallowing-regurgitation pauses, whose periodicity is revealed by the low frequency components close to zero. Both spectra show high intensity in very low frequencies, which is related to the offset close to 1 s shown in both tachograms (Figure 5). This contribution is not expected for activities other than grazing and rumination, where this kind of periodic events may not be present. For example, sip events are periodic but have a lower rate compared to chewing and biting JM (Galli, 2008).

2.3.4. Jitter variants

Jitter is the variation of time period or latency that occurs between consecutive JM events (Figure 7). During grazing bites, chews and chew-bites occur with an inter-event duration period that varies around 1 s; greater inter-event time interval variation indicates greater jitter. During rumina-

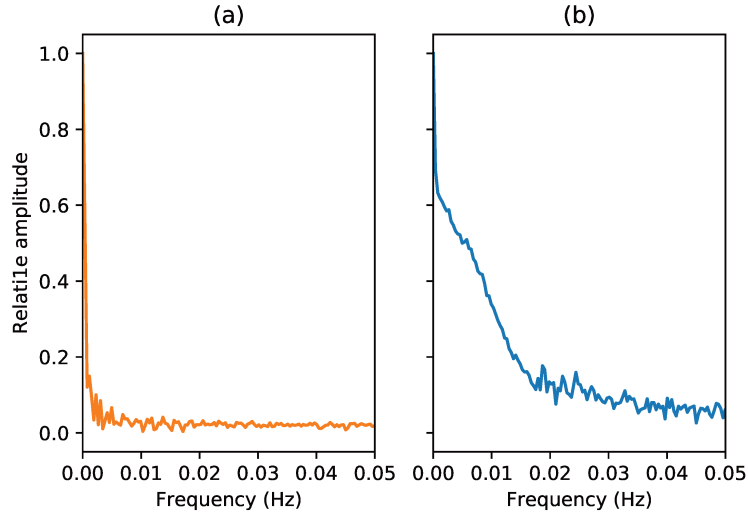


Figure 6: Spectrum of (a) grazing and (b) rumination tachograms.

tion, regular chews occur with an inter-event interval of lower variability compared to grazing, thereby with lower jitter. Other activities showing more irregular occurrence of detected events, will have greater jitter value compared to grazing and rumination.

There are alternative variants for the quantification of jitter:

- Absolute jitter ($jitter_{abs}$) - is the average absolute difference between

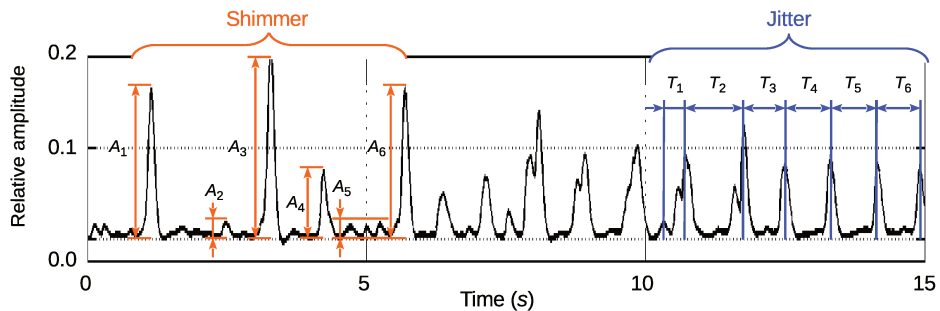


Figure 7: Representation of jitter and shimmer measures in audio envelope.

consecutive inter-event periods within the given sound segment

$$jitter_{abs} = \frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}| \quad (1)$$

where T_i is the i -th time period between two consecutive events and N is the number of inter-event periods in the sound signal segment being analyzed.

- Relative jitter ($jitter_{rel}$) - is the average absolute difference between consecutive inter-event periods divided by the average inter-event period of the sound segment

$$jitter_{rel} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |T_i - T_{i+1}|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (2)$$

- PPQ5 jitter ($jitter_{ppq5}$) - is the five-point period perturbation quotient computed as the average absolute difference between an inter-event period and the average of that period and the four closest neighbor periods, divided by the average period of the sound segment

$$jitter_{ppq5} = \frac{\frac{1}{N-4} \sum_{i=3}^{N-2} |T_i - \frac{1}{5}(T_{i-2} + T_{i-1} + T_i + T_{i+1} + T_{i+2})|}{\frac{1}{N} \sum_{i=1}^N T_i} \quad (3)$$

The absolute jitter quantifies the variability of the inter-event periods in a sound segment without taking into account the average period. Thus the absolute jitter loses sensitivity when the inter-event period changes. This problem is solved by the relative jitter, which scales the absolute jitter by the average period of the sound segment. It expresses the inter-event variability

as a fraction of the average period, providing a good sensitivity to inter-event period changes. However, these alternatives (equations 1 and 2) focus only on event variability, introducing a high variability in the jitter indexes. To incorporate information of event neighbors, the PPQ5 jitter is employed. This jitter variant considers the average of a period and the four nearest neighbor periods to compute the inter-event period information.

2.3.5. Shimmer variants

Shimmer is defined as the variation of peak to peak signal amplitude (dB) between events (Figure 7). During grazing, bites, chews and chew-bites occur along a wide range of signal amplitudes; thus the shimmer between events has a high value. During rumination, rumination chews exhibit lower variation in signal amplitude, or have a lower shimmer value compared to grazing. Other activities have no regularity in their amplitude, showing a shimmer value higher than grazing and rumination (not shown).

There are alternative variants for the calculation of shimmer:

- Absolute shimmer ($shimmer_{abs}$) - is the variability of the peak-to-peak amplitude in decibels, which is computed as the average of absolute base-10 logarithm of the ratio between the amplitudes of consecutive events, multiplied by 20,

$$shimmer_{abs} = \frac{1}{N-1} \sum_{i=1}^{N-1} |20 \log(A_{i+1}/A_i)|, \quad (4)$$

where A_i is the i -th extracted peak-to-peak amplitude and N is the number of detected events in the sound segment.

- Relative shimmer ($shimmer_{rel}$) - is the average absolute difference between amplitudes of consecutive events divided by the average amplitude of events in the sound segment,

$$shimmer_{rel} = \frac{\frac{1}{N-1} \sum_{i=1}^{N-1} |20\log(A_{i+1}/A_i)|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (5)$$

- APQ5 shimmer ($shimmer_{apq5}$) - is the five-point amplitude perturbation quotient, or the average absolute difference between the amplitude of an event and the average of the amplitudes of that event and the four closest neighbors events, divided by the average amplitude of events in the sound segment.

$$shimmer_{apq5} = \frac{\frac{1}{N-4} \sum_{i=3}^{N-2} |T_i - \frac{1}{5}(A_{i-2} + A_{i-1} + A_i + A_{i+1} + A_{i+2})|}{\frac{1}{N} \sum_{i=1}^N A_i} \quad (6)$$

Like jitter, the absolute shimmer quantifies the amplitude variability between consecutive events in a sound segment without taking into account the average amplitude, thus it loses sensitivity when the average amplitude changes. This problem is solved using the relative shimmer, which scales the absolute shimmer by the average amplitude of the sound segment. The relative variant expresses the inter-event amplitude variability as a fraction of the average amplitude, thereby providing a good sensitivity to inter-event amplitude changes. However, these shimmer variants (equations 4 and 5) focus only on event variability, introducing a high variability in the shimmer indexes. To incorporate information of event neighbors the APQ5 shimmer

is employed. This shimmer variant considers the average of amplitude of a given event and the four closest neighbors to compute this inter-event amplitude information.

3. Experimental setup

3.1. Database

The database was generated from measurements obtained from July 31 to August 18, 2014, at the Michigan State University’s Pasture Dairy Research Center facility at the W.K. Kellogg Biological Station, located in Hickory Corners, Michigan.

Sound tracks were recorded on five experimental, high-production multiparous Holstein cows (652.0 ± 5.4 kg of live weight; 2.3 ± 0.5 calvings; 102.0 ± 14.6 days in milk; 41.2 ± 3.9 kg of daily milk production) managed as part of a larger herd (~ 140 cows). The fieldwork was carried out following protocols for robotic milking routines (3.0 ± 1.0 daily milkings), using two Lely A3-Robotic milking machines (Lely Industries NV, Maassluis, The Netherlands). The cows had permanent access to pastures with a predominance of perennial ryegrass (*Lolium perenne*) and white clover (*Trifolium repens*), or with a predominance mixture of orchardgrass (*Dactylis glomerata*), tall fescue (*Lolium arundinacea*) and white clover. The pastures were grazed in two shifts, from 10:00 to 22:00 and from 22:00 to 10:00. In each shift cows freely accessed a new paddock with the similar level of daily allowance (15 kg of dry matter/cow). Consequently, the average daily pasture allowance was 30 kg of dry matter/cow. The forage quality of hand plucking samples was determined through five parameters: in vitro of dry matter digestibil-

ity (IVDMD, Daisy II, Ankom Technology Corp.) the average content of crude protein (CP, 4010 CN Combustion, Costech Analytical Technologies Inc., Valencia, CA), Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF, Fiber Analyzer, Ankom Technology Corp., Fairport, NY), and Acid Detergent Lignin (ADL). The values (mean \pm standard deviation) estimated for IVDMD, CP, NDF, ADF and ADL were $78.1 \pm 3.0\%$, $18.7 \pm 2.5\%$, $49.3 \pm 4.5\%$, $25.7 \pm 2.0\%$, $3.3 \pm 0.8\%$ respectively.

Twenty four hour continuous acoustic recordings obtained on 6 non-consecutive days were used. In each of the days the foraging behavior of the five cows was recorded. Five independent equipment devices (halter, microphone and recorder) were used, and they were rotated on the five cows throughout the six experimental days. Each equipment included a directional microphone pressed to the forehead of the animal connected to a digital recorder. The microphones were placed against the forehead of the cow, held and protected by an elastic band, to prevent microphone movement, isolate noise caused by wind and protect microphones from sources of friction and scratches. All recordings were performed with a 44.1 kHz sample rate and 16-bit resolution, providing a nominal recording bandwidth of 22 kHz and a dynamic range of 96 dB. The records were saved in Waveform Audio (WAV) file format.

Two experts with long experience in foraging behavior detection and digital analysis of acoustic signals identified the activities through direct observation and listening of recordings, using the Audacity² program. During

²<https://www.audacityteam.org/>

daylight hours, the activities (grazing, rumination, resting, drinking, etc.) of each of the cows were checked out every 60 minutes. This information was used to confirm or correct the grazing and rumination times obtained from Audacity sound records. Activity blocks were identified, classified and labeled, either as grazing, rumination, or other activities.

3.2. Variants proposed

Different combinations of the proposed algorithm were analyzed. Table 1 shows the features used by the former reference method (BUFAR), the proposed method called JM segment-based Foraging Activity Recognizer (JMFAR), and variants analyzed. Two aspects were taken into account: (i) the recognition performance and (ii) the computational cost. The variants analyzed were:

- JMFAR: This variant is the proposed method considering all new features.
- JMFAR-sel: This variant considers the resulting features after applying a feature selection method on the set of JM-segment features.
- JMFAR-ns: This variant does not consider the spectral information (f13, f23 and f24) in order to reduce the computational cost and memory requirements of the algorithm.

For all the proposed methods features related to JM classification (f1, f2 and f3) were excluded.

The feature selection was performed with a wrapper method using a genetic algorithm as a search method. As the internal classifier of the wrapper,

Table 1: Features used by BUFAR, the proposed method JMFAR and its variants.

Method	Features																							
	f1	f2	f3	f4	f5	f6	f7	f8	f9	f10	f11	f12	f13	f14	f15	f16	f17	f18	f19	f20	f21	f22	f23	f24
BUFAR	•	•	•	•																				
JMFAR				•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
JMFAR-sel							•	•	•			•	•	•			•	•	•	•			•	
JMFAR-ns				•	•	•	•	•	•	•	•			•	•	•	•	•	•	•	•	•	•	•

a multilayer perceptron similar to the external classifier was used. The parameters used were a mutation probability of 0.2, a crossover probability of 0.5, a population size of 50 individuals and a maximum of 30 generations (Kohavi and John, 1997). For this process, a 5-fold scheme on the training set was used. The accuracy was used as the scoring measure in the selection process. The features that were selected at less in two folds were included.

For this study, a set of 24 h of recordings were used to train and optimize parameters and hyper-parameters of the activity classifier and they were never used again. Classifiers were trained³ following a 5-fold cross validation on the training set. Finally, the test results were obtained from a separate test set of 137 h of recordings, which were selected taking care that they correspond to a free-ranging environment (i.e. outside of the feeding barn).

3.3. Performance assessment

The weighted average F1-score (Blair et al., 1979) was computed as a weighted metric for the precision and recall of methods. This metric summarizes the overall performance of methods and for their specificity for classification of grazing and rumination. In a previous work the authors presented

³This stage was implemented in python using the scikit-learn package.

results at the frame and at the block-level using spider-plots (Chelotti et al., 2020). However, in order to obtain each individual measure per signal, the proportion and duration of each activity were assumed to be equal. Although the information in the spider-plots is useful, in this work the actual length of time is taken into account and presented in a summarized and more accurate form. The F1-scores were computed using sequences of 1 s frames for the analysis. For each audio signal three F1-scores were computed, one for rumination, one for grazing, and one average score weighed by the number of frames for each activity. F1-scores range between 0.0 and 1.0 and values closer to 1.0 indicate better activity classification performance.

The analysis of each audio signal provided a list of blocks where the start, end, and type of activity are specified. These blocks were partitioned to 1 s non-overlapping frames sequences. Due to the fact that activity bouts last several minutes or even hours, using 1 s frames provides a very detailed resolution of results. Further, any modification of the limits of an activity block below 1 s can be discarded from the point of view of the recognition problem. This transformation was applied both to the algorithm output and ground truth of each signal. After this transformation there are corresponding pairs of sequences of labeled frames for the ground truth and each analyzed algorithm. Thus, conventional metrics for classification problems can be used.

4. Results

4.1. Comparative performance of BUFAR and JMFAR methods

The boxplots of Figure 8 show the weighted F1-score for BUFAR and JMFAR algorithms. Each data point considered for the graphic corresponds

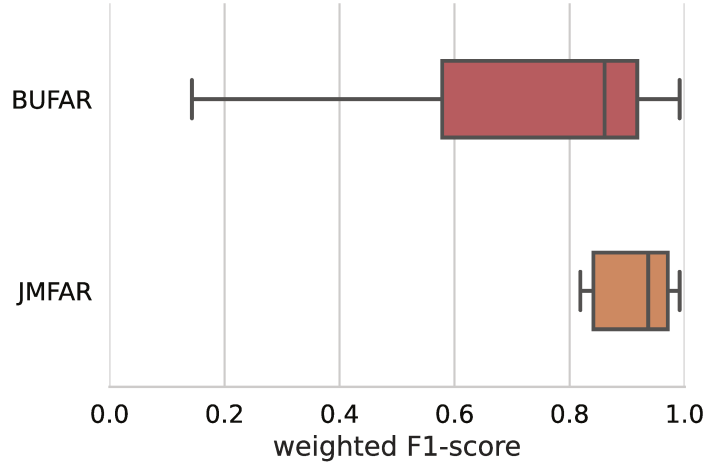


Figure 8: Weighted average of frame F1-score.

to an audio signal analyzed. The interquartile range for F1 scores of BUFAR spans from 0.55 to 0.91 and has a median of 0.86. By contrast, the JMFAR algorithm shows a narrower range and distribution of F1-scores. The interquartile range of JMFAR spans from 0.83 to 0.97 and has a median of 0.93. Score differences between BUFAR and JMFAR have shown to be significant ($p=0.012$) according to a Wilcoxon signed-rank test comparison (Wilcoxon, 1947).

Figure 9 shows the confusion matrices for the BUFAR and JMFAR. Rows represent frames and their actual activity label, whereas columns represent the frames and activity label as they were assigned by the algorithm. Each cell $[i, j]$ provides the proportions of frames from the activity class i that were categorized as activity j by the algorithm. The cells along the diagonal indicate the correct classifications. A classifier that provides correct labels for each frame would generate a unit matrix. Off diagonal cells repre-

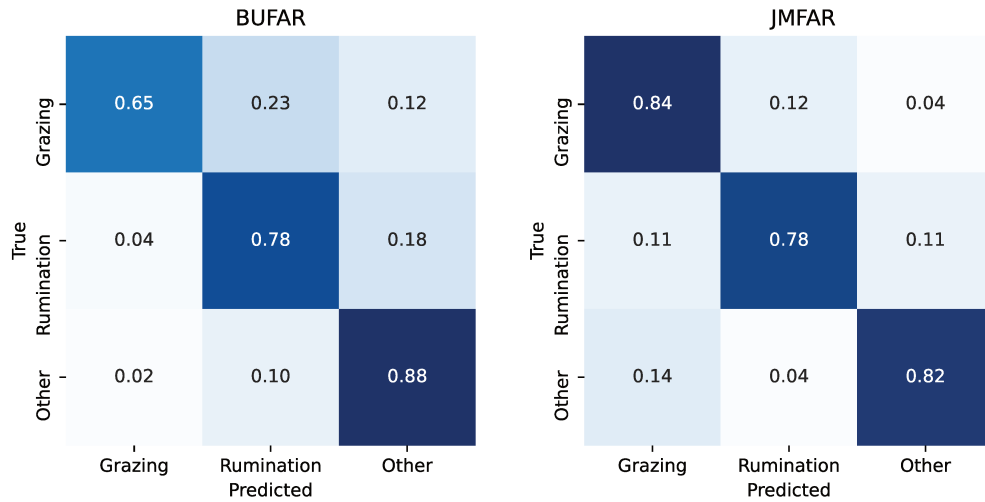


Figure 9: Confusion matrices for the BUFAR and JMFAR algorithms. The improvement in performance for grazing recognition is clear.

sent different kinds of errors for incorrectly classified frames. Matrix values range from 0.0 (white) to 1.0 (dark blue). The matrix for BUFAR shows that 65% of grazing and 78% of rumination frames were correctly classified. Further, 23% of grazing frames were incorrectly labeled as rumination. On undetected frames, the overestimation of grazing was 12% for grazing and 18% for rumination. JMFAR confusion matrix shows better results for the activities of interest. Recognition was correct for 84% of grazing and 78% for rumination. Confusions among rumination and grazing were similar (11 and 12%) and undetected frames were much lower, 4% for grazing, and 11% for rumination.

Previous results can be interpreted temporally as a function of sound signal length. For example, if a 180 min audio signal corresponds to 60 min of grazing, 60 min of rumination and 60 min to other activities, classification results would have been as follows. The BUFAR algorithm would have

predicted 42.6 min of grazing and 66.6 min of rumination, with 13.8 min of true grazing activity mislabeled as rumination and 2.4 min of true rumination mislabeled as grazing. Similarly, classifications by JMFAR would have improved to 65.4 min of grazing and 56.4 min of rumination, with 7.2 min of true grazing mislabeled as rumination and 6.6 min of true rumination mislabeled as grazing.

Additional spider plots for classifications of grazing and rumination by BUFAR and JMFAR algorithms are available in the supplementary material. Plots use the same metrics as in Chelotti et al. (2020), allowing for direct study comparisons of methods.

4.2. Variants of the proposed algorithm

Figure 10 shows the confusion matrices for JMFAR and variants JMFAR-sel and JMFAR-ns. The JMFAR-sel variant only uses 12 of the proposed JM-segment features, allowing savings on computational costs. However, benefits are quickly offset by weaker classification results. Correct classifications were 82% for grazing and only 53% for rumination, with a 47% of undetected rumination frames. The JMFAR-ns variant achieves satisfactory results for grazing and rumination (76% and 78%, respectively), with a lower misclassification between these activities (9% and 5% respectively). This variant has lower computational cost and buffer memory size, since spectral features were excluded from classification analyses.

A summary of the average performance and computational costs for JMFAR and variants are presented in Table 2. The JMFAR improves the previous recognition accuracy of BUFAR by 8.5% and without increasing significantly the associated computational costs. In general the use of fewer

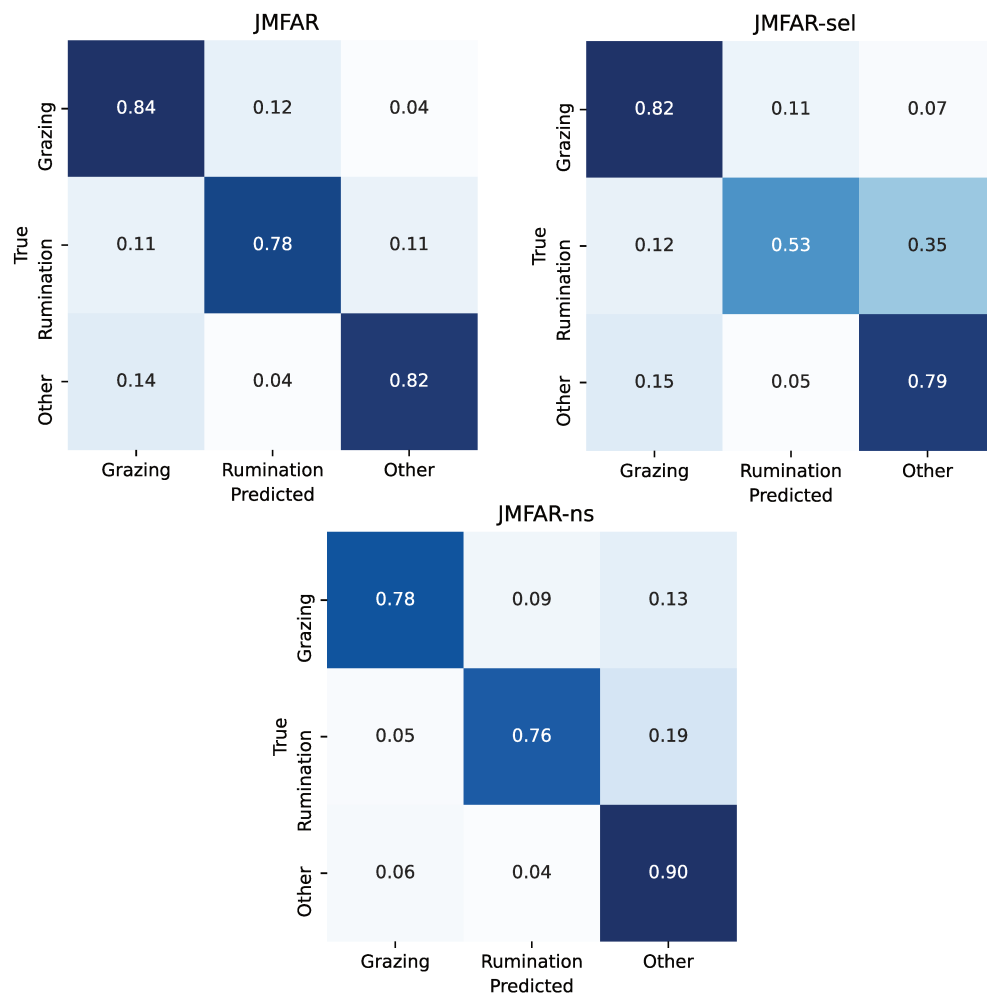


Figure 10: Confusion matrices of the algorithm variants analyzed.

Table 2: Overall performance and computational cost of BUFAR and JMFAR variants.

Method	Average accuracy (%)	Computational cost (operations/s)
BUFAR	71.5	37,966
JMFAR	81.0	50,445
JMFAR-sel	67.5	43,736
JMFAR-ns	77.0	37,645

classification features in JMFAR-sel and JMFAR-ns leads to a deterioration of the recognition accuracy compared to JMFAR. Remarkable is the deterioration of the recognition accuracy of rumination for JMFAR-sel (25%), despite the improvement in computational cost. The average decline of recognition accuracy for JMFAR-sel was 13.5% compared with JMFAR and 4% compared to BUFAR, which occurred without incurring major alteration of computational costs. The JMFAR-ns also shows a decline in recognition accuracy compared to JMFAR (this decline was 6% for grazing, 2% for rumination), and an improvement of recognition accuracy compared to BUFAR (the improvement was 13% for grazing, and 2% for rumination). The computational cost was numerically lower for JMFAR-ns than BUFAR.

5. Discussion

The individualized monitoring of grazing and rumination patterns could offer insights on sward–animal interactions affecting daily herbage intake, thereby providing unprecedented opportunities to optimize grazing management, enhance animal production or meet individualized requirements for animal nutrition, health, well-being and welfare (Beauchemin, 2018). This

study is limited to free grazing environment. Typical indoor scenarios, as a barn, could present additional challenges to the proposed method due to environmental noises. In addition, practical factors such as device placement, bumps and splashes could affect the performance of the system. These specific situations should be addressed in future studies.

Acoustic method has some advantages over pressure sensors and accelerometers. The sound quality obtained by placing a microphone on the forehead of the animal enabled algorithm to very precisely discriminate the sounds associated with diverse jaw movements during grazing and rumination. Moreover, this precise detection capability enables to further expand the method to estimate the dry matter intake, to recognize the ingested forage species and also to obtain information about the surrounding environment (Galli et al., 2020; Li et al., 2021).

In this study the reliability and accuracy of applying on-line acoustic methods to monitor grazing and rumination patterns were assessed and compared. The proposed system was built using signal processing and machine learning concepts. Repositories of acoustic data collected with lactating dairy cows on a robotic milking facility were used for the process. Overall, increasing the number of temporal, statistical and spectral features in the classifier improved the accuracy of algorithms, whereas the inclusion of greater number of features and classification steps slightly increased computational costs compared to previous algorithms for operation on-line (i.e. in portable embedded systems).

In this work sets of JM-segment features were tested for incorporation in a foraging activity recognizer called JMFAR. The proposed JMFAR algorithm

showed better overall performance compared to BUFAR (Chelotti et al., 2020), a former algorithm developed on a bottom-up approach for recognition of foraging activities in cattle (Figures 8 and 9). JMFAR achieved 84% and 78% for grazing and rumination recognition, respectively and a misclassification between grazing and rumination frames of 12% and 11%. Furthermore, the recognition of grazing was improved from 65% to 84% compared to the former algorithm BUFAR. This is an important achievement because BUFAR can misclassify about 35% of true grazing frames either as rumination (23%) or other non-grazing activity frames (12%). A less significant misclassification problem of BUFAR could be seen for rumination, where 18% of true rumination frames could be incorrectly assigned to other non-grazing activities. Hypothetically, the issue of misclassifications by BUFAR could be associated with a limited availability of statistical information on JM in the classifier. In BUFAR most of the activity classification lies on the accuracy of classification of JM. In this sense, the new algorithm JMFAR overcomes these limitations by using new JM-segment features of acoustic data, which makes the new activity classifier fully independent of the JM-classification. Indeed, JMFAR bridges completely the step for JM classification previously addressed by BUFAR. In JMFAR and the variants tested, new arrays and variants of JM-segment features of sound data such as the jitter, the shimmer and the sound spectra, are increasingly relevant. This sequence view rather than a simple group view of detected JM clearly makes the difference in the quality of activity classifications. The sequence patterns of acoustic features and their differences between grazing and rumination are the key distinctive elements between these two activities and for other non-grazing

and non-rumination activities.

Among variant algorithms tested the JMFAR-sel which relied on a supervised selection for JM-segment features, achieved the most modest results (Figure 10). This shows that in order to reduce the computational cost, retaining only the most discriminative features, relevant information is lost. The overall performance of JMFAR-ns, a variant that excluded spectral features in the classifier, showed a slightly lower accuracy compared to JMFAR (77% vs 81% from Table 2), although computational costs were reduced in JMFAR-ns. The exclusion of spectral information and reliance on features in the time domain, reduces the number of costly iterations such as fast Fourier transformations (see Appendix A for more details), thereby reducing the associated computational cost. Further, this JMFAR-ns variant achieved the lowest confusion of classification between grazing and rumination frames (which were 9% and 5%) yet it had a tendency to underestimate grazing and rumination by 13% and 19%, respectively. Hypothetically, this underestimation could be due to an exclusion of the 1.0-1.5 Hz energy band from the analysis, which is associated with regularly performing JM, increasingly significant in grazing and rumination chewing. Interesting results were observed when both JM-classification and JM-segment features were used together. In this case, the overall performance was 84% and 81% for grazing and rumination, respectively. This implies an overall improvement of 3% for rumination in comparison with BUFAR, which could be due to a lower confusion of classification between grazing and rumination. The addition of JM proportional information, in particular the addition of proportions of chews during rumination, improved the recognition performance of rumination and overall.

This findings support the hypothesis that greater classification accuracy is the result of distinct contributions aggregating across sets of discrimination features (i.e. statistical, temporal or spectral), thereby improving the overall activity recognition process.

In practice, for on-line analyses tallying a daily monitoring with 8.00 h of true grazing, BUFAR would have indicated the equivalent median of 9 h 28 min (interquartile range of 8 h 13 min - 12 h 14 min), while JMFAR would indicate the equivalent median of 7 h 59 min (interquartile range of 7 h 19 min - 8 h 35 min) of classified grazing. Assuming literature values of 1.5 kg/h of dry matter consumption (Galli et al., 2018), previous values correspond to an approximate consumption with a median of 14.2 kg (interquartile range 12.3 - 18.3 kg) and 12.0 kg (interquartile range 11.0 - 12.8 kg) for BUFAR and JMFAR respectively. Certainly, additional research would be necessary to evaluate the application of acoustic methods for estimation of herbage dry matter intake. As the pioneer work by Laca et al. (2000) has provided the proof of concept for the use of ingestive sounds for prediction of herbage dry matter intake of sheep (Galli et al., 2011), cattle (Galli et al., 2018), and potentially other herbivore species. For example chewing and biting sounds apparently enclose statistical, energy and possibly spectral information that could be used as metrics for accurate estimations of dry matter intake (CV 17%) of cows and even when those cows are confronted with forages that have varying physical and chemical composition yet comparable sound signal properties and metrics. Despite progresses, further work is still necessary to apply extractions of acoustic and information from JMFAR and variants for scalable predictions of herbage dry matter intake

across animals.

In this study, JMFAR and variants explored likely trade-offs between classification performance and the associated computational costs (Table 2). The former JMFAR, considering all new JM-segment features, requires 50,445 operations per second (ops/s). There are two variants that have attempted to minimize the computational cost. The JMFAR-sel, reduced the reference computational cost by 13.3% at expense of reducing the overall performance from 81.0% to 67.5%. The JMFAR-ns variant, required 37,645 ops/s, and yet it has reduced costs by 25.4% the performance was reduced from 81.0% to 77.0%, despite the fact that this variant had better performance and lower computational cost compared to the preceding algorithm BUFAR. In the JMFAR-ns variant the number of classification features was reduced to 17. Thus, fewer features included in this variant and avoidance for prior classification of JM would have been responsible for all computational cost improvements. A detailed analysis of the computational cost is included in the Appendix A. The requirements for random access memory (RAM) is another key feature for the online implementation of algorithms in portable embedded systems. As a reference, a simple 8-bit data representation with storage rate of 100 Hz, would have required 30 kB of RAM to store a segment of a given sound envelope. Furthermore, the implementation of JMFAR and JMFAR-sel variants would have required an addition of 30 kB of RAM to store the spectral representation of the sound envelope. Thus, after consideration of a proportional consumption of power, the JMFAR-ns variant is suggested as the most appropriate method for implementation in a low power microcontroller-based system.

This work was performed with the same number of animals as in previous studies (Chelotti et al., 2020; Vanrell et al., 2018), with a total of 250 hours analyzed. This is a first step, in which we test the performance of the system with a homogeneous group of animals. Future work will contemplate the expansion of these results to a much wider range, including a larger number of animals and incorporating different frames, sexes, breeds and age. The level of variation between cows in this experiment should not represent a problem for the application of the method, in fact but, in principle it could be extrapolated to groups of cows homogeneous in breed, size and age. The degree of variation between individuals is important because it will determine the generality of predictions based on chewing sounds. Chewing activity could present individual differences due to the influence of sex, body weight, physiological state and age, but this should not necessarily cause differences in ingestive or ruminant chewing. The cows used in this trial were similar in live weight, age and size, but could differ in the size and structure of their teeth, as well as the size and shape of their heads that produced different chewing sounds. In the worst case scenario, these uncontrolled sources of variation could suggest that a calibration of the recording equipment for each individual might be required when the animals are not uniform.

6. Conclusions

A novel acoustic method for the recognition and estimation of grazing and rumination bouts has been proposed. The proposed JMFAR is based on JM-segment features, i.e. that the computation of each feature is independent of the JM identification. The proposed features were: statistics

of JM amplitudes and JM durations, spectral features of the sound envelope, JM jitter and JM shimmer variants, temporal and spectral features of the JM tachogram. The algorithm has shown an improved performance in comparison with the BUFAR method based on JM-classification features. Variants relying on maximum numbers for JM-segment features achieved the best overall classification results with slight trade-offs on computational costs, and memory requirements for use in portable embedded systems. Among the variants analyzed, the one that ignores spectral information showed the best trade-off between recognition performance and computational cost.

Acknowledgments

This study has been funded by Universidad Nacional del Litoral, PACT CAID 2011–525, Universidad Nacional de Rosario, projects 2013-AGR216, 2016-AGR266 and 80020180300053UR, Agencia Santafesina de Ciencia, Tecnología e Innovación (ASACTEI), project IO-2018-00082, Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), project 2017-PUE-sinc(i). Also, this study was funded under the USDA-NIFA MICL0222 and MICL0406 projects, and direct support from AgBioResearch-MSU. Authors wish to thank the insightful help and dedication by the KBS Robotic Dairy Farm staff for their assistance and support during the completion of this study.

AUTHORSHIP STATEMENT

JOC, SRV, JRG, SAU, DHM, LLG and HLR participated in conceptualization; JOC, SRV, JRG, AMP and SA participated in the data cura-

tion; JOC, SRV and LSMR participated in the formal analysis; JOC, SRV, JRG, AMP, SAU, DHM, LLG and HLR participated in investigation and methodology stages; JOC and SRV participated in validation and visualization stages; SRV, JRG, SAU, DHM, LLG, HLR participated in the funding acquisition; JOC, JRG, DHM, LLG and HLR participated in project administration; All the authors contributed to the writing and reviewing of the original draft; All the authors reviewed and approved the manuscript.

Appendix A. Computational cost

The computation cost of the JMFAR variants depends on the sampling frequency, the considered segment length and the number of features used by the activity classifier. In this analysis a sampling frequency of 2 kHz, a segment length of 5 min and two JM per second were considered in order to get a simple comparison with other methods. Worst-case scenarios were considered for each stage in order to get a theoretical upper bound. The required number of operations per stage of computation for JMFAR was:

1. Signal pre-processing: A detrending filter requires 5 operations per signal sample. Then, 10,000 ops/s are required.
2. JM detection: 27,200 ops/s are required to compute the subsampled envelope, generate a time-varying threshold and detect the jaw movement.
3. JM feature extraction: This task involves 300 ops/s.
4. Buffering: 6 ops/s are required to store the timestamp, and the amplitude and duration features.

5. Activity feature extraction: the worst-case scenario for activity feature extraction considers 2 event/s, which is equivalent to 600 events per segment.
 - (a) JM rate (f4): this feature requires only 1 op/segment.
 - (b) JM statistics:
 - i. mean amplitude (f5) and mean duration (f9): each feature requires 600 ops/segment.
 - ii. standard deviation amplitude (f6) and standard deviation duration (f10): 1801 ops/segment per feature are required.
 - iii. skewness amplitude (f7), skewness duration (f11): compute each feature requires 2,401 ops/segment.
 - iv. kurtosis amplitude (f8) and kurtosis duration (f12): compute each feature requires 3,001 ops/segment.
 - (c) Spectral feature of the envelope signal (f13): compute the energy in the 1.0-1.5 Hz band requires 1,830,252 ops/segment.
 - (d) JM jitter and shimmer variants:
 - i. absolute variants: absolute jitter (f14) requires 2396 ops/segment and compute the absolute shimmer (f18) requires 2398 ops/segment.
 - ii. relative variants: absolute jitter (f14) and absolute shimmer (f18) are used to compute relative jitter (f15) and relative shimmer (f19), respectively. Each feature requires 601 extra ops/segment.
 - iii. perturbation quotient variants: partial calculations of feature (f15) and partial calculations of feature (f18) are used to calculate the PPQ5 jitter (f16) and the APQ5 shimmer (f20),

respectively. Additional 4770 ops/segment are required per feature computed.

iv. standard deviation variants: absolute jitter (f14) is used to calculate the standard deviation of jitter (f17). Additional 2,997 ops/segment are required. Similarly, the feature (f15) is used to compute the standard deviation of shimmer (f21) involving 4,195 ops/segment.

(e) Temporal feature of the JM tachogram (f22): worst-case scenario requires 1,201 ops/segment.

(f) Spectral features extraction of the JM tachogram requires both an interpolation, a frequency domain transform, and an energy computation involving 2,009,479 ops/segment. Additionally, the computation of energy in the 0.017-0.020 Hz (f23) and 0.00-0.02 Hz (f24) bands requires 5 and 53 ops/segment, respectively.

6. Activity classifier (MLP): considering all extracted features (21) and the largest number of neurons in the hidden layer (20), the activity classification task requires 2,279 ops/segment.

The overall computation cost of JMFAR is: $37,506 \text{ ops/s} + 3,881,604 \text{ ops/segment}$.

To compare the algorithm with an online method with the same block length, the computational cost can be estimated at 50,445 ops/s.

References

- Andriamandroso, A., Bindelle, J., Mercatoris, B., Lebeau, F., 2016. A review on the use of sensors to monitor cattle jaw movements and behavior when grazing. *Biotechnologie, Agronomie, Société et Environnement* 20.
- Andriamandroso, A.L.H., Lebeau, F., Beckers, Y., Froidmont, E., Dufrasne, I., Heinesch, B., Dumortier, P., Blanchy, G., Blaise, Y., Bindelle, J., 2017. Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors. *Comput. Electron. Agric.* 139, 126–137.
- Arcidiacono, C., Porto, S.M.C., Mancino, M., Cascone, G., 2017. Development of a threshold-based classifier for real-time recognition of cow feeding and standing behavioural activities from accelerometer data. *Comput. Electron. Agric.* 134, 124–134.
- Bailey, D.W., Gross, J.E., Laca, E.A., Rittenhouse, L.R., Coughenour, M.B., Swift, D.M., Sims, P.L., 1996. Mechanisms that result in large herbivore grazing distribution patterns. *Rangeland Ecology & Management/Journal of Range Management Archives* 49, 386–400.
- Beauchemin, K., 2018. Invited review: Current perspectives on eating and rumination activity in dairy cows. *Journal of dairy science* 101, 4762–4784.
- Beauchemin, K.A., 1991. Ingestion and mastication of feed by dairy cattle. *Vet. Clin. North Am. Food Anim. Pract.* 7, 439–463.
- Berckmans, D., 2014. Precision livestock farming technologies for welfare management in intensive livestock systems. *Rev. Sci. Tech.* 33, 189–196.

- Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer Verlag.
- Blair, D.C., et al., 1979. Information retrieval, cj van rijsbergen. london: Butterworths; 1979: 208 pp. Journal of the American Society for Information Science 30, 374–375.
- Bristow, D.J., Holmes, D.S., 2007. Cortisol levels and anxiety-related behaviors in cattle. *Physiol. Behav.* 90, 626–628.
- Chelotti, J.O., Vanrell, S.R., Galli, J.R., Giovanini, L.L., Rufiner, H.L., 2018. A pattern recognition approach for detecting and classifying jaw movements in grazing cattle. *Computers and Electronics in Agriculture* 145, 83–91.
- Chelotti, J.O., Vanrell, S.R., Rau, L.S.M., Galli, J.R., Planisich, A.M., Utsumi, S.A., Milone, D.H., Giovanini, L.L., Rufiner, H.L., 2020. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Computers and Electronics in Agriculture* 173, 105443.
- Galli, J.R., 2008. Medición acústica del comportamiento ingestivo y del consumo de rumiantes en pastoreo. Tesis Doctorado en Ciencias Agrarias – Orientación Producción Animal. Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, Balcarce, Argentina .
- Galli, J.R., Cangiano, C.A., Milone, D.H., Laca, E.A., 2011. Acoustic monitoring of short-term ingestive behavior and intake in grazing sheep. *Livestock Science* 140, 32–41.

- Galli, J.R., Cangiano, C.A., Pece, M.A., Larripa, M.J., Milone, D.H., Utsumi, S.A., Laca, E.A., 2018. Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *Animal* 12, 973–982.
- Galli, J.R., Milone, D.H., Cangiano, C.A., Martínez, C.E., Laca, E.A., Chelotti, J.O., Rufiner, H.L., 2020. Discriminative power of acoustic features for jaw movement classification in cattle and sheep. *Bioacoustics* 29, 602–616.
- Giovanetti, V., Decandia, M., Molle, G., Acciaro, M., Mameli, M., Cabiddu, A., Cossu, R., Serra, M.G., Manca, C., Rassu, S.P.G., Dimauro, C., 2017. Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer. *Livest. Sci.* 196, 42–48.
- González, L.A., Bishop-Hurley, G.J., Handcock, R.N., Crossman, C., 2015. Behavioral classification of data from collars containing motion sensors in grazing cattle. *Comput. Electron. Agric.* 110, 91–102.
- Greenwood, P.L., Paull, D.R., McNally, J., Kalinowski, T., Ebert, D., Little, B., Smith, D.V., Rahman, A., Valencia, P., Ingham, A.B., Bishop-Hurley, G.J., 2017. Use of sensor-determined behaviours to develop algorithms for pasture intake by individual grazing cattle. *Crop Pasture Sci.* .
- Herskin, M.S., Munksgaard, L., Ladewig, J., 2004. Effects of acute stressors on nociception, adrenocortical responses and behavior of dairy cows. *Physiol. Behav.* 83, 411–420.

- Hodgson, J.G., Illius, A.W., 1998. *The Ecology and Management of Grazing Systems*. C A B International.
- Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. *Artificial intelligence* 97, 273–324.
- Laca, E.A., Ungar, E.D., Seligman, N.G., Ramey, M.R., Demment, M.W., 1992. An integrated methodology for studying short-term grazing behaviour of cattle. *Grass Forage Sci.* 47, 81–90.
- Laca, E.A., WallisDeVries, M.F., et al., 2000. Acoustic measurement of intake and grazing behaviour of cattle. *Grass and Forage Science* 55, 97–104.
- Li, G., Xiong, Y., Du, Q., Shi, Z., Gates, R.S., 2021. Classifying ingestive behavior of dairy cows via automatic sound recognition. *Sensors* 21. URL: <https://www.mdpi.com/1424-8220/21/15/5231>, doi:10.3390/s21155231.
- Van der Maaten, L., Hinton, G., 2012. Visualizing non-metric similarities in multiple maps. *Machine learning* 87, 33–55.
- Martiskainen, P., Järvinen, M., Skön, J.P., Tiirikainen, J., Kolehmainen, M., Mononen, J., 2009. Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines. *Appl. Anim. Behav. Sci.* 119, 32–38.
- Nydegger, F., Gyga, L., Egli, W., 2010a. Automatic measurement of rumination and feeding activity using a pressure sensor, in: *International Conference on Agricultural Engineering - AgEng 2010*, pp. 1–8.

- Nydegger, F., Gygax, L., Egli, W., 2010b. Automatic measurement of rumination and feeding activity using a pressure sensor, in: International Conference on Agricultural Engineering-AgEng 2010: towards environmental technologies, 6-8 September 2010, Clermont-Ferrand, France, pp. 1–8.
- Rutter, S.M., 2000. Graze: a program to analyze recordings of the jaw movements of ruminants. *Behav. Res. Methods Instrum. Comput.* 32, 86–92.
- Rutter, S.M., Champion, R.A., Penning, P.D., 1997. An automatic system to record foraging behaviour in free-ranging ruminants. *Appl. Anim. Behav. Sci.* 54, 185–195.
- Schirmann, K., Chapinal, N., Weary, D.M., Vickers, L., von Keyserlingk, M.A.G., 2013. Short communication: Rumination and feeding behavior before and after calving in dairy cows. *J. Dairy Sci.* 96, 7088–7092.
- Schirmann, K., von Keyserlingk, M.A.G., Weary, D.M., Veira, D.M., Heuwieser, W., 2009. Technical note: Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.* 92, 6052–6055.
- Smith, D., Rahman, A., Bishop-Hurley, G.J., Hills, J., Shahriar, S., Henry, D., Rawnsley, R., 2016. Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems. *Comput. Electron. Agric.* 131, 40–50.
- Stone, A.E., 2020. Symposium review: The most important factors affecting adoption of precision dairy monitoring technologies. *J. Dairy Sci.* 103, 5740–5745.

- Ungar, E.D., Rutter, S.M., 2006. Classifying cattle jaw movements: Comparing IGER behaviour recorder and acoustic techniques. *Appl. Anim. Behav. Sci.* 98, 11–27.
- Vanrell, S.R., Chelotti, J.O., Galli, J.R., Utsumi, S.A., Giovanini, L.L., Rufiner, H.L., Milone, D.H., 2018. A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *Comput. Electron. Agric.* 151, 392–402.
- Welch, J., 1982. Rumination, particle size and passage from the rumen. *Journal of animal science* 54, 885–894.
- Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., O’Brien, B., 2018a. Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows. *J. Neurosci. Methods* 300, 138–146.
- Werner, J., Leso, L., Umstatter, C., Niederhauser, J., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., O’Brien, B., 2018b. Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows. *Journal of Neuroscience Methods* 300, 138–146.
- Werner, J., Umstatter, C., Leso, L., Kennedy, E., Geoghegan, A., Shalloo, L., Schick, M., O’Brien, B., 2019. Evaluation and application potential of an accelerometer-based collar device for measuring grazing behavior of dairy cows. *animal* 13, 2070–2079. doi:10.1017/S1751731118003658.
- Wilcoxon, F., 1947. Probability tables for individual comparisons by ranking methods. *Biometrics* 3, 119–122.

Zehner, N., Umstätter, C., Niederhauser, J.J., Schick, M., 2017a. System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *Computers and Electronics in Agriculture* 136, 31–41.

Zehner, N., Umstätter, C., Niederhauser, J.J., Schick, M., 2017b. System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows. *Comput. Electron. Agric.* 136, 31–41.

Anexo D

Acoustic Real-time Sensor for Ingestive Behaviour of Grazing Cattle

Acoustic Real-time Sensor for Ingestive Behaviour of Grazing Cattle

Luciano S. Martínez Rau*, Nestor N. Deniz*, José O. Chelotti*,
Leonardo L. Giovanini* and Pablo A. Kler†

*Research Institute for Signals, Systems and Computational Intelligence, sinc(i), FICH-UNL/CONICET
Email: {lmrau, ndeniz, jchelotti, lgiovanini}@sinc.edu.unl.ar

†Research Center for Computational Methods (CIMEC, UNL-CONICET)
Email: kler@cimec.unl.edu.ar

Abstract

The increment of food market requirements and competitiveness makes that industry needs more accurate sensing and monitoring tools. Due to the large temporal dynamics necessary to monitor some animal behaviours, the autonomy and storage capacity of the devices involved becomes critical. In this work, we introduce a device specifically developed for on-line monitoring, quantifying and recording the feeding patterns of dairy cows. The prototype consists of an embedded circuit that records and analyzes the sounds produced by the animal to detect, classify and quantify the events related with the feeding behaviour. It implements an algorithm recently developed, which was adapted for its execution in a microcontroller based embedded system. A microcontroller with power management technology, combined with high efficiency harvesting power supply and power management firmware, enable long-term operation. The technology presented within this publication is protected under international patent application PCT/IB2016/057627.

Index Terms

Acoustic monitoring, automated detection and classification, embedded system, precision livestock farming.

Acoustic Real-time Sensor for Ingestive Behaviour of Grazing Cattle Preprint submitted to "IX Congreso de Microelectrónica Aplicada (UEA2018)".

I. INTRODUCTION

Despite the large number of successful applications, the livestock sector shows a lack of Wireless sensor solutions for monitoring the feeding behaviour. This situation can be understood from the stringent operational requirements that devices must satisfy to fulfill this task. In this sense, a device for monitoring the feeding behaviour of cattle must be: *i*) **Portable** in order not to interfere with the natural behaviour of the animal; *ii*) **Real-time processing** to allow continuous monitoring and reduce the communication and memory requirements; *iii*) **Autonomous operation** and **long operating time** to allow the recording and analysis of long term behaviours in the grassland (pasture); and *vi*) **Robustness** to operate in the environmental conditions imposed by outdoor activities.

A major requirements is that the sensors can operate automatically and autonomously for long period of time. However, to obtain a sensor with such characteristics it should be used a scheme of energy harvesting and a wireless technology to transmit the data frames to a personal computer (PC). In fact, in spite of the recent sensors on recording data for free-grazing cattle (see [1], [2]), up to this date to the best of our knowledge, there is no sensor capable of analyzing in real-time the acoustic emission produced by ruminants during their feeding activities.

One of the most accepted ways to perform monitoring of ruminant feeding is through the detection, classification and quantification of the three most common events of grazing activity: chew, bite, and composed chew-bite [3]. The importance of being able to determine these events is related to the quality and quantity of the food ingested by the animal [4]. However, currently available sensors are only able to detect jaw movement, behaviour detection (eating, ruminating, other) or to distinguish between chew and bite grazing event, and is unable to recognize the chew-bite event [5]. The first attempts for developing acoustic emission based algorithms for the detection and classification of these events are recent [6], [7].

The novel contribution of this paper includes a real-time sensor for ingestive behaviour monitoring of grazing cattle by means of acoustic emission analysis. Moreover, this paper introduces an acoustic emission sensing technique based on the algorithm developed by Chelotti et al. [6], embedded in the sensor for detection and classification of ingestive events. This is the first time acoustic emission sensing and processing, wireless communication, and energy harvesting, converge in a unique sensor solution for feeding behaviour analysis.

II. SENSING PRINCIPLE

The proposed on-site real-time data processing and evaluation algorithm aims to detect the feeding activities of ruminants. When a ruminant is feeding, it moves the jaw to perform two activities: **biting (B)**, when forage is apprehended and cut, and **chewing (C)**, when forage is crushed to reduce the size of the particles and increase the surface/volume ratio. The combined motion, **chewbite (CB)**, is the combination of the both movement. Then, the feeding behaviour can be categorized in terms of a temporal sequence of **B**, **C**, and **CB** [7].

Based on the shape, intensity and duration of the time-domain envelope of the these three characteristic signatures (**C**, **B**, and **CB**), Chelotti et al. [6] developed and evaluated an algorithm able to detect and classify the ingestive events in real-time. Fig. 1 shows a diagram of the algorithm. The algorithm was

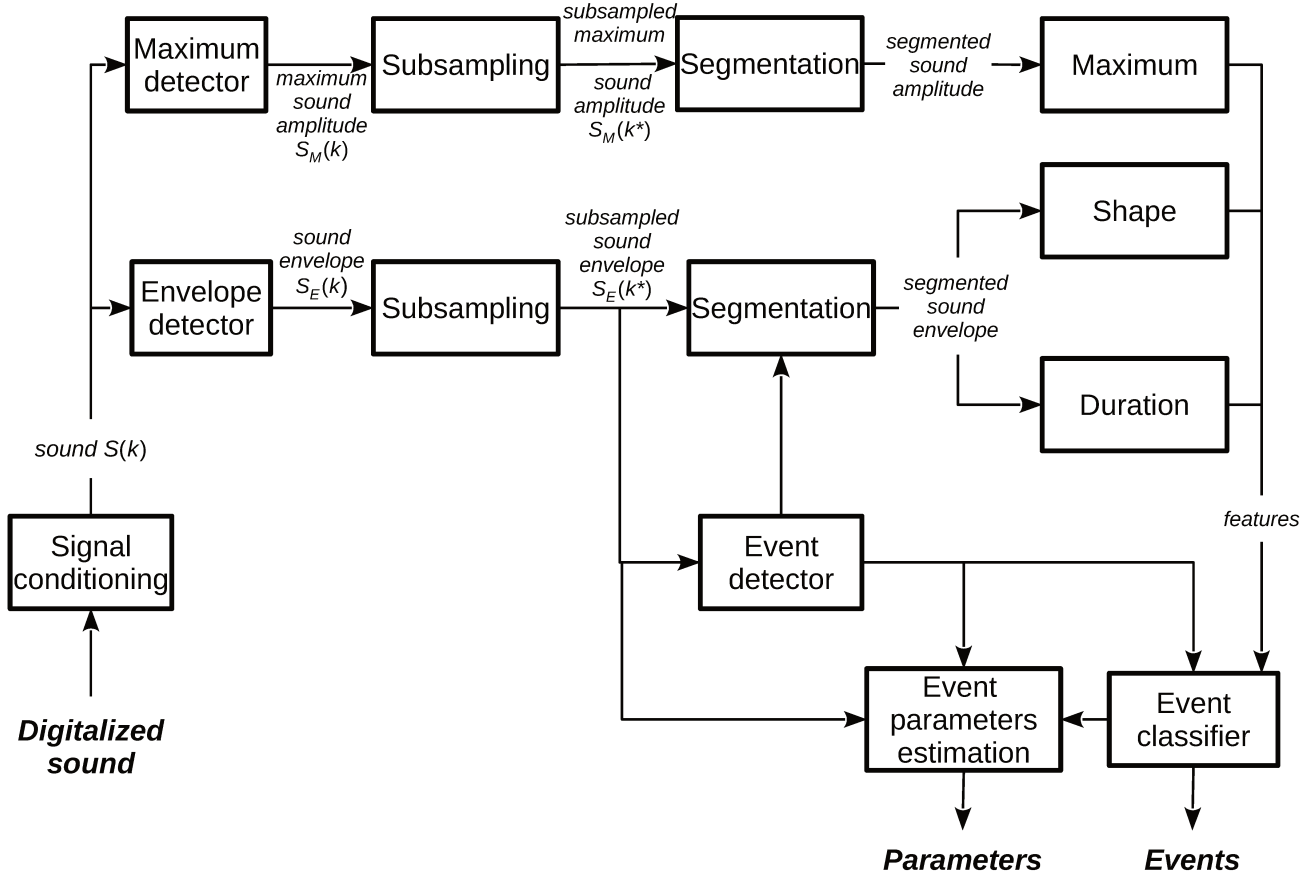


Fig. 1. Diagram of the algorithm for ingestive event detection and classification.

divided into two sequential tasks: **detection and segmentation** and **classification**. These tasks can be thought as a set of successive stages [Fig. 1].

After digitalizing the acoustic monitoring, the resulting signal is conditioned by removing the tendencies present in the signal. This task is performed using an adaptive least-mean square filter. Once the trend and noises have been removed, two signals are computed from $S(k)$: (i) the signal envelope $S_E(k)$ which is computed by filtering the absolute value of $S(k)$; and (ii) the peak amplitude $S_M(k)$ which is determined as the maximum value in a sliding window of 100 samples. $S_E(k)$ and $S_M(k)$ are then subsampled to 100 S/s, generating $S_E(k^*)$ and $S_M(k^*)$ in order to decrease the computational and memory requirements.

The next steps are the detection and classification of the events. This is performed by comparing $S_E(k^*)$ with a time-varying threshold $T(k^*)$, which is generated online. If an event is detected (i.e. $S_E(k^*) > T(k^*)$), the algorithm proceeds to compute the acoustic monitoring time-domain features (shape, duration, and maximum amplitude) in a window centered at the sample in which the event was detected. Then, the classifier proceeds to label the event using the extracted features. Finally, once the event is classified, the system estimates the characteristic parameters of the event (period, maximum amplitude, energy, among others). The period and energy content of the event are computed from $S_E(k^*)$, while the maximum amplitude $S_M(k^*)$ is computed from the sound $S(k)$. For further details about the algorithm, see [6]. Once a potential event is detected, the algorithm computes the features of the envelope and then the event classifier labels the event. The features are gathered with other relevant information in an information package (see section III-C).

III. EMBEDDED SYSTEM ARCHITECTURE AND MANAGEMENT

The design of a battery-powered embedded system requires a detailed analysis of each subsystem in order to minimize size, cost, and, principally power consumption. The Sensor Unit Device (SUD) has been designed as a trade-off between minimizing power consumption and parameter estimation accuracy. The SUD comprises different interconnected modules as shown in Fig. 2. All the modules have been designed to efficiently use the power available in the device while providing the best performance. The final resulting packaged SUD is shown in Fig. 3. The SUD is located on the neck of the animal, just behind the head. The embedded device was built around an MCF51JM128 (NXP Semiconductors, Eindhoven, Netherlands) microcontroller unit (MCU).

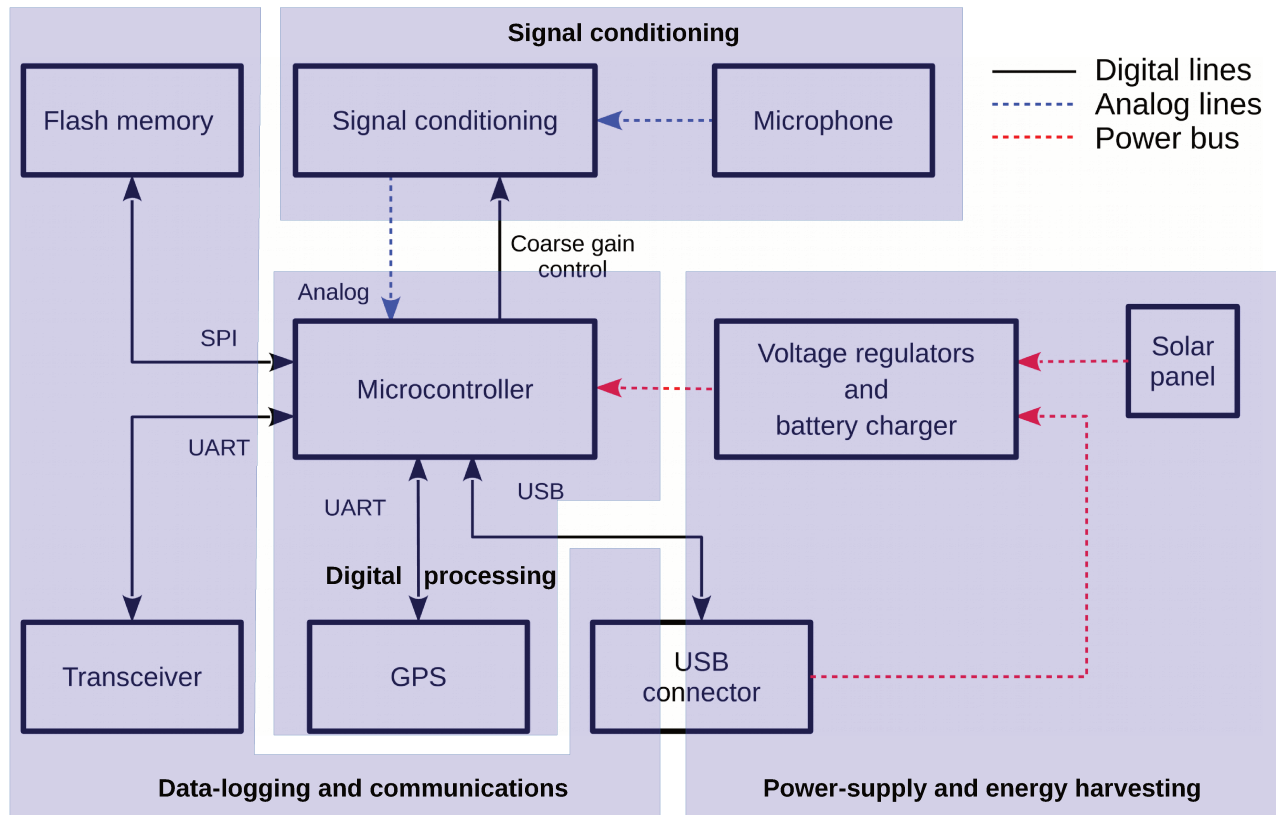


Fig. 2. Block diagram of the embedded system.

A. Signal conditioning

The sound produced by the animal is sensed with an electret microphone facing inward on his forehead. The electrical signal generated by the microphone is conditioned by an analog circuit that consists of three stages (Fig. 4).

In order to limit the signal bandwidth to the range of interest (below the cut-off frequency $f_c=2$ kHz) and to minimize the effect of quantization noise on the electrical signal, a low pass filter (low-pass filter 1 in Fig. 4) was designed and implemented. In order to obtain an output voltage lower than the corresponding to 1 least significant bit (LSB) at a $f=f_c$. It was required a minimum attenuation of 48.16 dB in the stopband. An eighth-order Butterworth low pass filter has been designed in a Sallen-Key configuration. The integrated circuit op-amp used for the filter implementation was a TLV2784 (Texas Instruments, Dallas, US).

An automatic gain control (AGC) amplifier MAX9814 (Maxim Integrated, San José, US) is used in order to maximize the signal-to-noise ratio (SNR), avoiding signal distortions, and clipping produced by

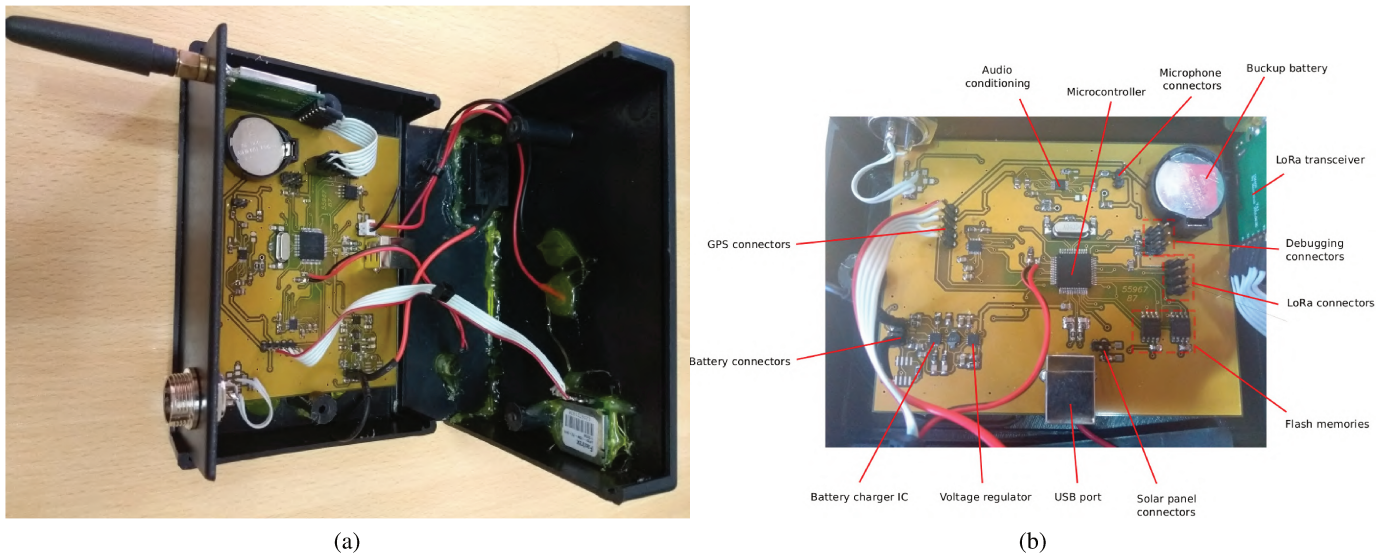


Fig. 3. Printed circuit board design of the embedded system.

high gain amplifications. The AGC output is connected to an analog input channel of the A/D converter in the MCU (*Analog input 0*), configured with 8 bits resolution. The number resulting from the conversion includes information of the sign of the signal. The AGC applied gain is available at an output pin, for supplying the MCU with this critical information through another input channel of the A/D converter (*Analog input 2*).

The AGC amplifier output is also connected to a second order Sallen-Key low pass filter (low-pass filter 2 in Fig. 4) with $f_c=0.1\text{Hz}$ and unity gain. The op-amp used was an TLV2781 (Texas Instruments, Dallas, US). This filter computes the acoustic monitoring root mean square value, that it is both connected to a third input channel of then A/D converter in the MCU (*Analog input 1*) to detect when there is acoustic monitoring present and is compared with a defined threshold reference voltage using an TLV2781 in an op-amp voltage comparator configuration. The output compared signal is employed to wake-up the MCU (IRQ input) from the sleep mode.

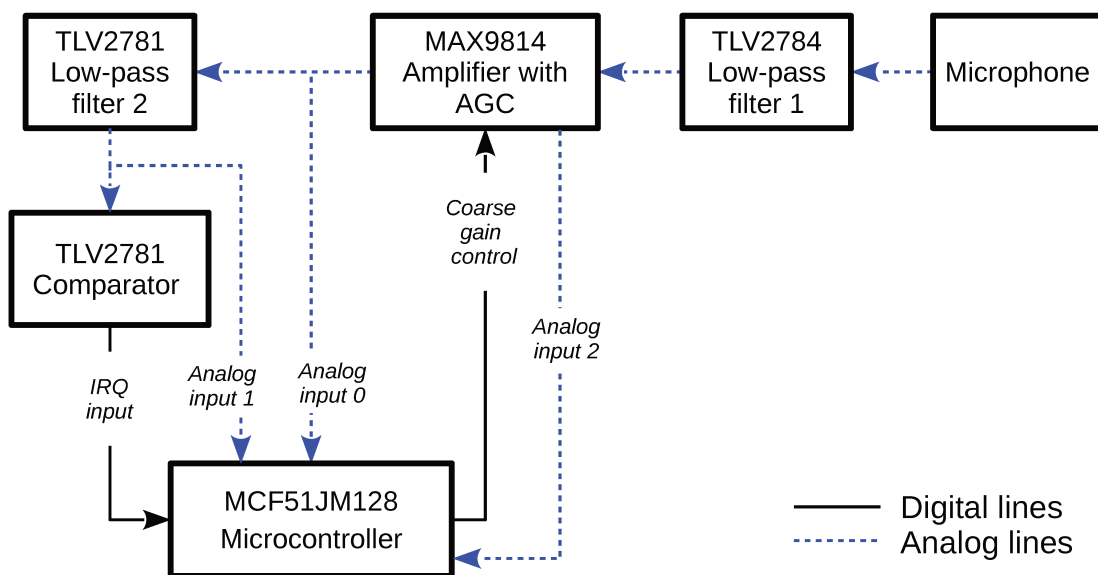


Fig. 4. Block diagram of the signal conditioning circuit.

B. Digital Processing

The embedded software is organized in four tasks: *i*) signal conditioning and preprocessing, event detection, and classification, *ii*) data logging, *iii*) internal and external communication, and *iv*) device configuration. The software architecture is driven by interrupts that switch between active and sleep modes between interrupt events. The use of interrupt driven programming allows to the power consumption of the device to be optimized by using low power sleep modes ensuring a precise data sampling and processing intervals. Each of these four main interrupts wake up the MCU into active power mode and a routine is used to determine which function caused the interrupt event.

When the operational comparator detect an acoustic monitoring, the algorithm start to work for 5 min. Every 500 μ s time interval, the MCU goes to the active mode by the finished A/D conversion and then it execute the algorithm to detect and classify the ingestive events. Then, the software extract the information that characterized the ingestive activities and accumulates the partial results, and the MCU is set into sleep mode until the next finished A/D conversion. Finally, after 5 minutes since the last acoustic monitoring detection, the device is hibernated.

For every 15 min interval a data frame is built (Section III-C), where the data frame is integrated by the information related to the feeding activities (i.e. quantity, average time, and average energy of each potential event), the time, and position at that moment. The time and position is obtained from a global positioning system (GPS) that normally operates in power save mode and it is switched to maximum performance mode to acquire the position. Finally, the data frame is stored as a text file in the flash memory, and the MCU is set into sleep mode.

Every 15 min the PC starts and control the wireless communication with the SUD, generating an interrupt in the MCU. It is achieved through the use of request and response packages protocol. The reception module of the transceiver is on continuously, waiting for a request package from the PC. When a request package is received, the MCU turns on the transmission module, transmits the requested data frames, waits for the reception acknowledgment of the PC and then it turns off the transmission module and the MCU.

The final interrupt routine is triggered by an Universal Serial Bus (USB) port data event. This is caused by messages being sent from the PC to the embedded system when an USB communication is established. When a download command is sent to the SUD, the MCU transfers the data stored in the flash memory to the PC through the USB port. A message containing a new configuration settings can also be sent to the SUD.

a) **##,D,T,Lat,Long,NC,DC,AC,EC,NB,DB,AB,EB,NCB,DCB,ACB,ECB,TNE,##**

Fig. 5. Data file format used to store information

C. Data-logging and communications

Internal communications in the SUD are managed by the MCU, which collect data to create a data frame structure. The accurate time and position are obtained from a NEO-6M GPS receiver (u-blox,Thalwil, Suiza), using NMEA 0183 standard to communicate with the microprocessor. A 3 V and 20mm CR2032 coin cell battery is incorporated as a backup power supply to ensure accurate time and position, as well as a fast startup of the GPS in the event of primary battery failure, since satellite ephemeris data and configuration settings are stored in the backup memory.

The data frame structure is created in the MCU, as shown in Fig. 5(a). It starts and ends with two numerals characters (#) and the data fields are separated by commas (.). This structure reduces the parsing time and avoids data loss or misleading. It has been categorized in two parts: *i*) the first four fields are

the date (D), time (T), latitude (Lat), and longitude (Long) in NMEA 0183 standard format where the data was collected. *ii*) The remaining fields corresponds to the parameters of grazing activities. It stores the number of events (**C**, **B**, and **CB**), the average duration of time (DC, DB, and DCB), the average amplitude (AC, AB, and ACB), the average energy (EC, EB, and ECB) of the grazing events, respectively, and the the total number of events (TNE) . Each data field of this section is stored in 32 bits integer format. This information is codified by using a low-density parity-check (LDPC) code to guarantee the integrity of the data from errors.

Data logging, time, and position stamping are carried out through the host MCU interfacing with two W25Q128FV serial flash-NOR memory chips (Winbond Electronics, Taichung, Taiwan).

The USB module provides an On-The-Go (OTG) dual-role controller. All these features simplify the hardware and software required for communication. The additional wireless communication mode operating at 433/470 MHz has been implemented using a UM402 SX1278 LoRa module (ManThink Co., Beijing, China) with wired antenna that enables a transfer rate up to 37.5 kbps for a length range up to 6 km.

D. Power supply and energy harvesting

The design of a battery powered embedded system requires a detailed analysis of each subsystem and software to minimize power consumption, considering both energy provision and consumption simultaneously. In this way, the device has been designed with the goal of minimizing power consumption whereas it is able of harvest all the energy needed for its operation. Therefore, three complementary approaches were used to develop the energy management scheme: *i*) A combined duty-cycling and data-driven operating scheme, driven by data, to operate only when relevant information is available. It is implemented through the embedded software in the microcontroller combined with an interrupt scheme to allow the digital parts of the SUD to alternate between active and sleep modes, when there is acoustic monitoring present or not, respectively. *ii*) An energy harvesting scheme, by using a 1W solar panel (5.5 V open circuit voltage and 170 mA short circuit current) able to recharge the batteries when the SUD is operating outdoors; *iii*) A USB port when the device is connect to a PC or an energy source.

The duty-cycling power management scheme is implemented through the embedded software combined with the aforementioned interrupt scheme.

A high efficiency battery charge with protected overvoltage input and overcurrent input is provided between each harvester and the battery pack. These features enables the SUD to run with a defective or missing battery pack. The battery power is supplied by two polymer batteries (Li-Ion 3.7 V 2500 mAh). The MCU monitors the true battery voltage through an analog input channel of the A/D converter (*Analog Input 3* in Fig. 6) and in case of errors, turn off the SUD to avoid problems in the detection and classification of events.

A primary regulated 3.3 V power supply is provided to the SUD through a high efficiency buck/boost charge pump regulator TPS63001 (Texas Instruments, Dallas, US). This integrated circuit allow to achieve a low ripple, stable high quality voltage, and long autonomy of the device.

IV. LABORATORY TESTS AND DISCUSSION

To evaluate the performance of the proposed SUD, two experimental analysis have been performed. The first analysis is focused on the recognition and classification capabilities, whereas the second one lies on the power consumption and energy harvesting capabilities.

A. Experimental Setup and Performance measurements

The validation of the system was done via laboratory tests. The soundtracks employed were obtained by a field experiment conducted at the Campo Experimental Villarino (Universidad Nacional de Rosario, Zavalla, Argentina) dairy facility, during october of 2014. The foraging behaviour of five Holstein lactating

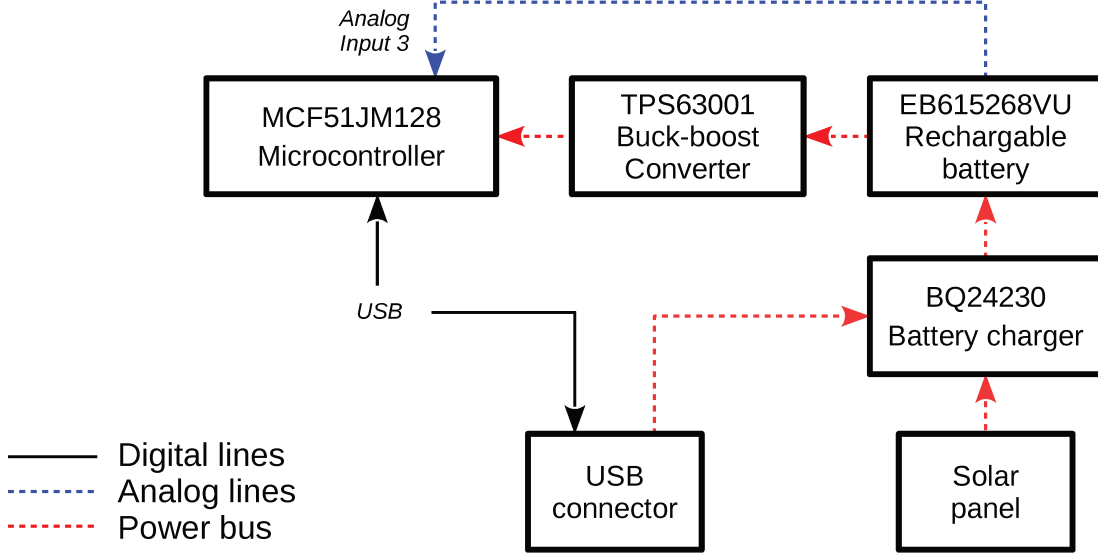


Fig. 6. Block diagram of power supply circuit.

cows were continuously monitored during six non consecutive days. The evaluation was performed with 24 hours continuous sound recordings at 44.1 kHz sampling frequency using 16-bit resolution and WAV format. The acoustic monitoring were recorded using a directional microphone mounted onto the forehead of the animal and covered by a elastic band fastened to a halter, where a commercial recorder was attached. The microphone/recorder devices were randomly assigned to the cows and rotated over the six days.

For experiments, five representative segments of 10-min from a total of 24h of continuous audio recording were aurally segmented and labeled event by event within rumination and grazing activities by two experts in animal behaviour. In order to consider different types of foraging behaviours, each segment consisted of 5-min of rumination (contain only chew events) and 5-min of grazing (contain the three types of events), which were randomly selected.

One important issue for the comparison between events recognized and classified by the algorithm and the corresponding reference labels made by the experts is the synchronization time of events in both sequences. To solve this, the HTK performance analysis tool HResults was used. The comparison is based on a Dynamic Programming-based string alignment procedure [8]. The outputs of this tool were: (i) the number of deletions (D), which are considered as false negatives, (ii) the number of substitutions (S), which are considered as misclassified events, (iii) the number of insertions (I), which are considered as false positives, and the total number of events (T) in the reference transcription files provided by the experts. Using this information, performance measures can be established. Regarding the **detection task**, the percentage of correctly detected events is defined as

$$D\% = (T - D)/T \ 100\% \quad (1)$$

where the number of substitutions was not considered, because it only matters if an event has occurred or not, regardless of the type of event. Regarding the **classification task**, the percentage number of labels correctly recognized **CL%** and the accuracy **A%** is given by

$$CL\% = (T - D - S)/T \ 100\% \quad (2)$$

$$A\% = (T - D - S - I)/T \ 100\% \quad (3)$$

where both substitutions S and insertions I are considered.

B. Recognition results

The main parameters used by Chelotti et. al (i.e. shape=2, intensity=0.3, duration=0.3s) suffer small adaptations due to the integer arithmetic used to achieve a real-time operation on the MCU. The recognition rates obtained from the analysis are shown in Table I.

The results show that the embedded system developed was able to correctly detect and classify 78% of the feeding events. The best recognized event was the C event, in agreement with the results obtained by Chelotti et al. (2016) [6]. It should be noted the similarity between segment and for the average between the values of correct and accuracy, which evidences a low number of insertions (false positives) for the system, which represents a good stability of the proposed methodology. These results reveal that the analysis of the time-domain acoustic monitoring features shown properly the corresponding jaw motion. In addition it was shown that processing was executed in real-time.

C. Autonomy analysis

The estimation of the time in which the SUD will be in active mode is one of the most important parameters to know about the autonomy, because it is when the greatest energy consumption will occur. This time is the sum of the rumination and grazing activities, due to the fact that these are the periods in which the target masticatory events are present.

The circadian rhythm of the cattle feeding behaviour will be crucial to determine the SUD's autonomy. According to [9] during daylight hours, the average time of both activities (rumination and grazing) is about 40 min per hour, whereas during night this value drops to around 30 min. This indicates that the SUD will be in active mode around 60% of the daily time.

According to the values in Table II, the calculated SUD current consumption during the active power mode is $I_A = 34.9mA$. The measured current consumption is 31-34 mA and it is close to the measured value. Every 5 min, due to the data frame interruption, the GPS is waked up and there is a peak consumption of 44 mA. When the SUD is in active mode, there are many current peaks of 34 mA each 500 μs , which correspond to the sampling frequency of the acoustic monitoring signal and the execution of the detection and classification algorithm.

TABLE I
CONFUSION MATRIX, AVERAGE RECOGNITION RATES AND ACCURACY.

		Bite	Chew	Chew-bite	Correct	Accuracy
S_1	B	82	9	9	78	77
	C	1	98	1		
	CB	10	9	81		
S_2	B	77	8	15	83	81
	C	1	97	2		
	CB	5	6	89		
S_3	B	81	11	8	75	74
	C	1	97	2		
	CB	6	11	83		
S_4	B	95	5	0	70	67
	C	4	84	12		
	CB	3	11	86		
S_5	B	84	8	8	85	83
	C	1	96	3		
	CB	3	13	84		
Average					78.2	76.4

TABLE II
ESTIMATED POWER CONSUMPTION AND PERCENTAGE DURATION OF EACH MODULE.

Module	Power Consumption [mA]	Time
MCU	11.5	100 %
Signal Conditioning	2.5	50 %
GPS	11.5	0.33 %
Communication RX/TX	13/110	5.00%/0.22%
Storage	4	0.33 %
Power Supply	20	100 %

Taking into account the current consumption give by I_A , the charge consumption per hour during active mode will be 34.9 mAh. Considering a battery capacity of 5000 mAh (two batteries of 2500 mAh each) and in absence of the solar panel, the SUD would have an autonomy of $T_A = 143.27h$.

On the other hand, the instant measured current provided by the solar panel is 120 mA during daylight hours. If we suppose it is maintained during 6 h (25% of the daytime), the average charge provided per hour is $Q_{SP} = 30.0mAh$. According to the circadian cycle of the cattle feeding behaviour, the average charge consumed per hour for a whole day is given by $Q_{CR} = 20.6mAh$. Thus, the net charge balance is $Q_{SP} - Q_{CR} = 9.4mAh$. This shows that the harvested power is sufficient to energize the SUD.

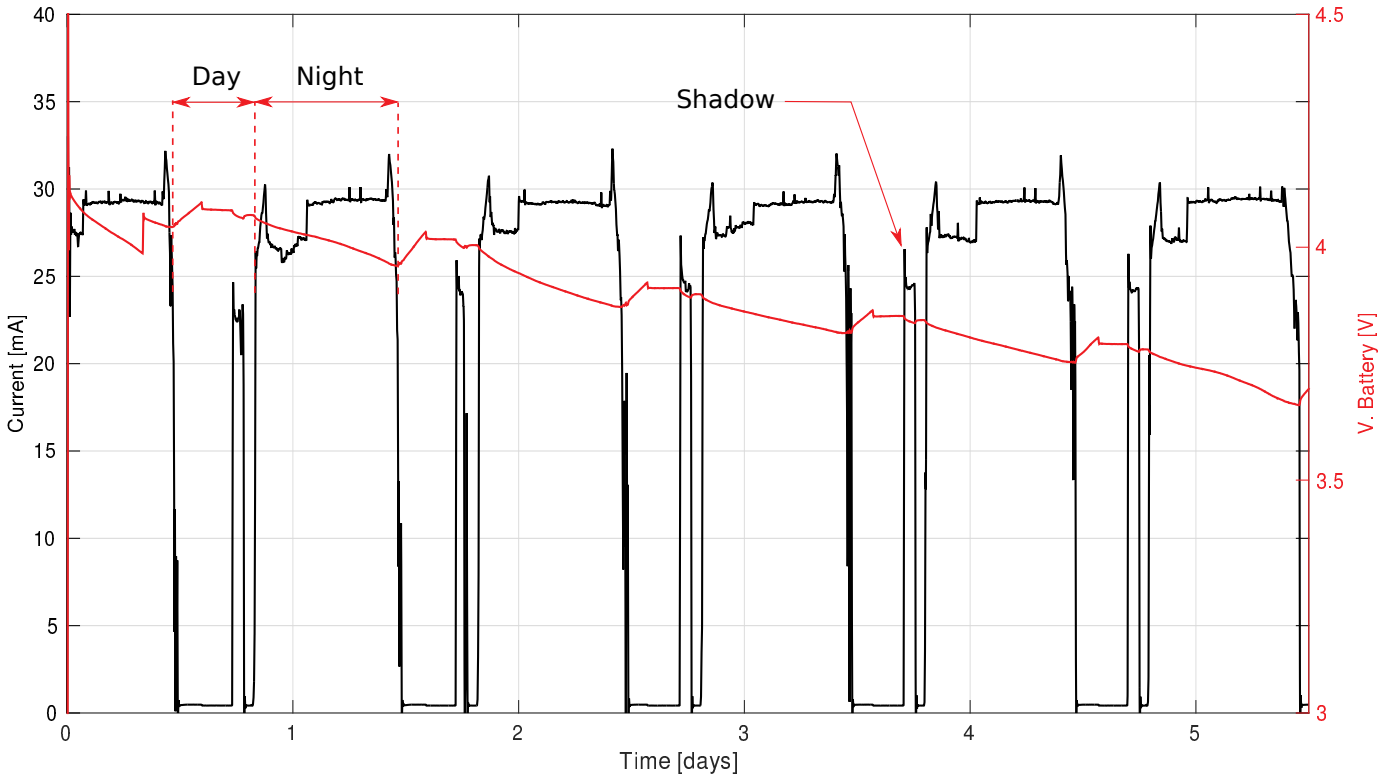


Fig. 7. Current consumption (black line) and voltage of the battery pack (red line).

In order to determine the profile of voltage and current variations in the battery pack to estimate the autonomy of the SUD, it was tested in active power mode all the time, i.e., the worst-case conditions, during a period of 5 days outdoors with the batteries fully charged initially. In Fig. 7 the evolution of voltage and current of the batteries are shown. During the day, the current provided by the batteries decrease to zero, when the SUD is powered entirely by the solar panel and, at the same time, the batteries are partially recharged. However it is not enough to accomplish a full charge of the batteries. As expected, during the night, the batteries supply the embedded system and its voltage decreases considerably. During each daylight, there is a small peak in the current consumption, which corresponds to the projection

of a shadow above the solar panel making that much of the current consumed by the SUD has to be provided by the batteries during this time. This projected shadow emulates the situation in which there is no incident light on the solar panel during daylight, either by clouds covering the sun, by a motion of the cow's head, or because the cattle is located under a tree, among others.

At the end of the test, the SUD still worked properly and the voltage tends to stabilize close to 3.7 V, which is the nominal batteries voltage.

V. CONCLUSION

This paper introduces a novel embedded system for real-time monitoring the feeding behavior of ruminants by means of acoustic monitoring analysis. The motivation of this work was to provide a tool to enhance the understanding of feeding behavior of dairy cows by developing sensor technology that allows the continuous monitoring of animal feeding activities under different environmental conditions.

The proposed solution integrates different key technologies in a unique device, i.e. acoustic monitoring signal conditioning, wired/wireless communications, data logging, power management, energy harvesting, and embedded detection and classification algorithm. The proposed methodology was implemented in a microcontroller and it is based on the analysis of certain characteristics of the acoustic monitoring signal, since the different jaw action of the animal have different shape, intensity, and duration.

The proposed embedded algorithm has been validated over a test bench by showing that it was able to correctly detect and classify 78% of the feeding events identified by human experts, which is similar to the results obtained by Chelotti et al. (2016) [6]. Regarding the power consumption, the sensor solution can work autonomously following the circadian rhythm of the cattle feeding behaviour, as well as, it can work for more than 5 days in the worst case, in which the animal feeds 24 hours a day, without recharging the batteries externally. When the solar panel is illuminated, it energizes the sensor and recharges the batteries.

Since the device has not yet been tested on animals, in future studies, the device will be tested on a cow under field operational conditions for continuous operation over time, where the cow motion and the weather could play important roles in the power supply system.

REFERENCES

- [1] A. Z. Abbasi, N. Islam, Z. A. Shaikh *et al.*, "A review of wireless sensors and networks' applications in agriculture," *Computer Standards & Interfaces*, vol. 36, no. 2, pp. 263–270, 2014.
- [2] B. Panckhurst, P. Brown, K. Payne, and T. Molteno, "Solar-powered sensor for continuous monitoring of livestock position," in *Sensors Applications Symposium (SAS), 2015 IEEE*. IEEE, 2015, pp. 1–6.
- [3] E. D. Ungar and S. M. Rutter, "Classifying cattle jaw movements: comparing iger behaviour recorder and acoustic techniques," *Applied animal behaviour science*, vol. 98, no. 1, pp. 11–27, 2006.
- [4] J. R. Galli, C. A. Cangiano, D. H. Milone, and E. A. Laca, "Acoustic monitoring of short-term ingestive behavior and intake in grazing sheep," *Livestock Science*, vol. 140, no. 1, pp. 32–41, 2011.
- [5] A. L. H. Andriamandroso, J. Bindelle, B. Mercatoris, F. Lebeau *et al.*, "A review on the use of sensors to monitor cattle jaw movements and behavior when grazing," *BASE*, 2016.
- [6] J. O. Chelotti, S. R. Vanrell, D. H. Milone, S. A. Utsumi, J. R. Galli, H. L. Rufiner, and L. L. Giovanini, "A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle," *Computers and Electronics in Agriculture*, vol. 127, pp. 64–75, 2016.
- [7] D. H. Milone, J. R. Galli, C. A. Cangiano, H. L. Rufiner, and E. A. Laca, "Automatic recognition of ingestive sounds of cattle based on hidden markov models," *Computers and electronics in agriculture*, vol. 87, pp. 51–55, 2012.
- [8] S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey *et al.*, "The htk book, vol. 2," *Entropic Cambridge Research Laboratory Cambridge*, vol. 4, 1997.
- [9] L. Balocchi, F. Pulido, V. Fernández *et al.*, "Comportamiento de vacas lecheras en pastoreo con y sin suplementación con concentrado," *Agricultura Técnica*, vol. 62, no. 1, pp. 87–98, 2002.

Anexo E

Developments on real-time monitoring of grazing cattle feeding behavior using sound

Developments on real-time monitoring of grazing cattle feeding behavior using sound

Luciano S. Martínez Rau*, Nestor N. Deniz*, José O. Chelotti*,
 Leonardo L. Giovanini* and Pablo A. Kler† *Research Institute for Signals, Systems and
 Computational Intelligence, sinc(i), FICH-UNL/CONICET
 Email: {lmrau, ndeniz, jchelotti, lgiovanini}@sinc.edu.unl.ar †Research Center for
 Computational Methods (CIMEC, UNL-CONICET)
 Email: kler@cimec.unl.edu.ar

Abstract

Estimating forage intake and monitoring the foraging behavior of grazing livestock are difficult tasks. Detection and classification of jaw movements are very useful to obtain that information. In a similar way, the monitoring and analysis of long-term activities such as rumination and grazing provide useful insight. Several works have demonstrated that acoustic monitoring is an adequate way to analyze ruminant feeding behavior. In this work, we present a complete system for monitoring ruminant foraging behavior. As components of such a system, a review about two own methods based on the analysis of acoustic signals is included: i) a short-term analysis system that automatically detects and classifies jaw movements, and ii) a long-term analysis system for the recognition of grazing and rumination activities. Both systems use simple concepts and tools derived from signal processing and pattern recognition areas. A description of an ad-hoc electronic platform is also included.

Index Terms

feeding behavior, signal processing, pattern recognition, embedded system

I. INTRODUCTION

The world dairy industry has undergone profound changes over recent decades. A trend exists in dairy farming toward the automation of processes to reduce labor and labor costs [1]. This development is partly driven by the economic reality of increasing labor costs relative to capital costs. Automated systems enable dairy farmers to manage larger herds with lower labor requirements, which means that the application of automated systems fits with the trend of increasing herd sizes.

Animal behavior is a clear indicator of its physiological and physical state [2]. Eating, ruminating, and resting are the main daily activities of ruminant livestock. Monitoring these activities is key to many important management decisions in free-grazing systems [3]. Such information enables farmers to check the living conditions of cattle in the pasture and make effective decisions about food supplement and pasture management. Therefore, accurate monitoring of the feeding behavior of free-grazing cattle is necessary to ensure the welfare and health of these animals, which will improve quantity and quality of livestock products.

Monitoring of foraging behavior is key to ensure the fulfillment of the basic health and welfare requirements of grazing cattle and to improve the efficiency of pasture-based production systems [3]. Thus, the continuous monitoring of such behavior can help retrieve individual status information for each animal [4], [5], build a log, detect emerging diseases [6] or the onset of estrus, and optimize pasture and animal management.

Cattle foraging behavior is mainly composed of grazing and rumination bouts. Grazing can cover from 25% to 50% of the day and rumination, from 15% to 40% [7]. Grazing involves searching, apprehending,

chewing, and swallowing herbage. Rumination includes bolus regurgitation, chewing, and deglutition. While grazing, the animal moves its jaw with no predefined sequence of jaw movements, a typical rumination involves chewing for 40–60 s followed by a 3-to-5 s interruption due to bolus deglutition and regurgitation [3], [8], [9]. During both activities, jaw movements (JM) are performed rhythmically with a frequency that ranges from 0.75 to 1.20 events per second. JMs are: biting, when herbage is apprehended and severed; chewing, when herbage is comminuted; and a compound movement called chew-bite, when herbage is severed and comminuted in the same JM [10]–[12]. JMs length is around 1 s, whereas activity bouts can last from minutes to hours. Thus, foraging behavior is characterized by JMs (short timescale) and activities (longer timescale).

An approach to measure feeding behavior is acoustic monitoring [13]. Laca et. al. [10] instrumented an inward-facing microphone on the forehead of steers to register stronger and readily distinguishable sounds of bites, chews, and chew-bites. Consequently, acoustic monitoring proved to be a more effective methodology to discriminate sensitive differences in feeding and rumination than previous jaw recorders or visual observation methods [11], and since then it has been increasingly applied as a research tool to study different aspects of grazing behavior in sheep and cattle [14].

In this work, we review previously published algorithms based on the acoustic method: i) one related to the recognition of JMs [15] and the other ii) related to the recognition of grazing and rumination activities [16]. The concept and results of each method are presented. Also, the design and evaluation of an ad-hoc embedded system are presented.

II. SENSOR SYSTEM FOR FEEDING BEHAVIOR

A. Jaw movement recognition

Since the 1980s, a lot of work has been put into developing sensors that measure parameters of individual cows. A sensor system consists of devices plus the software that processes the data (see Fig. 1):

- **Transduce and record** the signal of interest;
- **Analyze** the data to explain the changes to produce information about the state of the cow;
- **Integrate** the information provided by sensor with supplementary data to improve the quality and accuracy of animal information; and
- **Make a decision** using the information obtained by the system to advise the farmer.

The stages defined here describe the abstraction level of information provided by the sensor system. The sensor itself is only the first stage. The second stage is to process the sensor data with algorithms that provide information about the state of each individual cow. In this stage, it is possible to combine sensor data with data about cow history. The algorithm produces information about the cow's state by determining changes in the sensor data. The third stage uses this information in a decision support model that uses economic information to produce advice about how to act upon the detected events. Finally, the fourth stage is the decision regarding the change in the health status of the cow, as detected by the sensor.

In this work, we focused on the initial stages of the system for the case of real-time monitoring of the feeding behavior of grazing cattle (left side in Fig. 1). The proposed system is based on acoustic sensors and algorithms capable of achieving good performance in the detection and classification of feeding activities with a low computational cost, which allows its real-time execution. Therefore, a group of measurable properties should be found to characterize the sounds produced by JMs.

III. THE ALGORITHMS

A. Jaw Movement Recognition

A pattern recognition system is an automatic process that aims at classifying input data into a set of specific classes [17]. This system can be described by a series of generic stages that allow: (i) the input signal description, which facilitates the extraction of distinctive features, and (ii) its classification, which

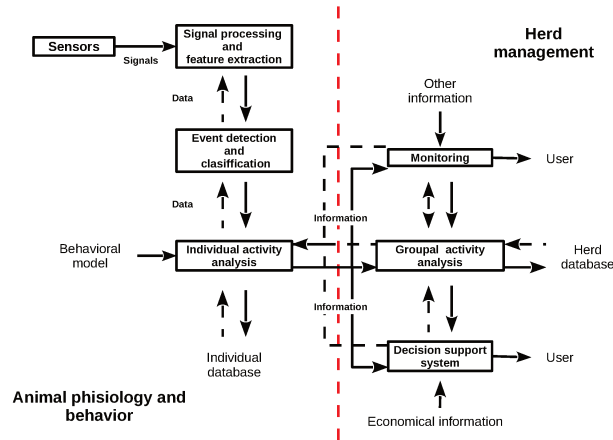


Fig. 1. Structure of a sensor system for animal monitoring

enables identification of patterns. A block diagram of the jaw movement recognition algorithm (called CBIA in the original work) is shown in Fig. 2.

Fig. 2 shows the relationship between a typical pattern recognition system and the different stages of the algorithm: signal conditioning, preprocessing, event detection, feature extraction, and event classification. The input of the system is the digitized sound, whose numerical representation is normalized and its range is matched with the range of the computer in the signal conditioning stage. Sound signals sometimes show slow time-varying noises added to the target signal, especially in barn environments. Therefore, CBIA uses a detrending technique that removes the non-stationary noises at the signal conditioning stage. Within the pre-processing stage, the sound signal follows two paths:

- A maximum detector computes the maximum amplitude of the sound signal over a sliding window whose length is half of the duration of a typical chew-bite event.
- An envelope detector computes the sound envelope using synchronous demodulation and a low-pass filter.

Since the sound envelope only has low-frequency components, the signals computed by both detectors are down-sampled to reduce the amount of data processed by the remaining stages. The events are detected by comparing the sound envelope with a time-varying threshold [18]. Then, the sound envelope

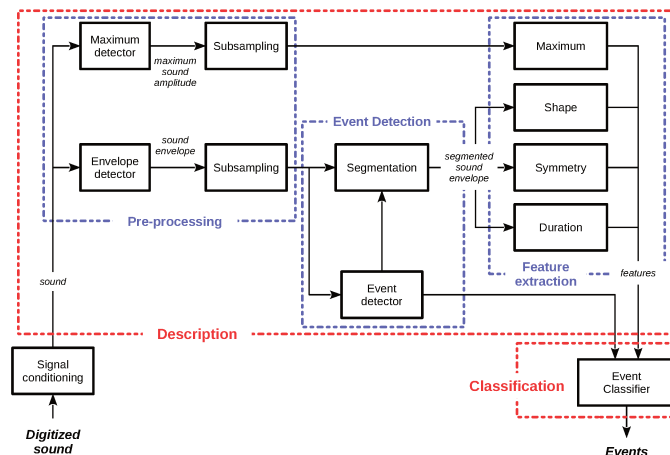


Fig. 2. Block diagram of the JM recognition system (CBIA).

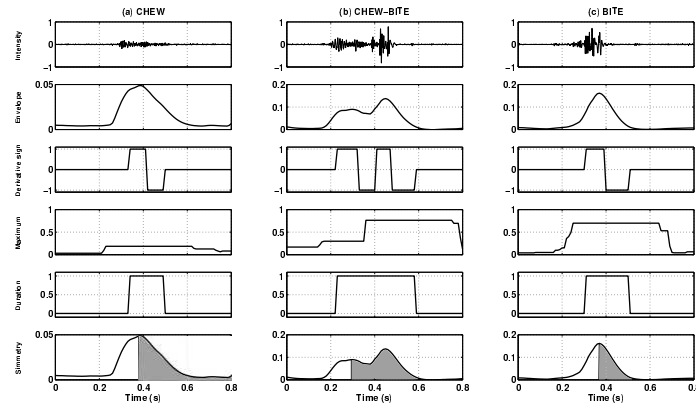


Fig. 3. Acoustic events, jaw movements, and derived signals.

is segmented and it is used to compute the features. Once the candidate JMs are detected, their features are extracted over a time window centered at the sample where the event was detected. Four temporal features that are low-cost and with discriminative power for this problem are extracted:

- **Shape index:** is computed as the number of zero-crossings in the sign of the derivative signal obtained from the envelope signal (third row in Fig. 3). This calculation is performed only if the envelope amplitude exceeds a noise threshold. This feature provides useful information to differentiate simple JMs (chews and bites) from combined JMs (chew-bites).
- **Maximum intensity:** provides information to differentiate low-amplitude JMs (chews) from high-amplitude ones (bites and chew-bites). This feature is extracted directly from the sound signal over a sliding window with length equal to the period of a typical chew-bite event (fourth row in Fig. 3).
- **Duration:** is calculated as the time in which the envelope amplitude is greater than a given threshold. In general, the duration of compound events (chew-bites) is larger than simple events (chews or bites), which are similar (fifth row in Fig. 3).
- **Symmetry:** is computed as the ratio between the left area and the total area of the event. Left and right event areas are divided at the first peak of the event (last row in Fig. 3). It can provide discriminative information because events have different symmetries.

B. Activity recognition

Grazing and rumination are activities with quasi-periodic characteristics. In addition, each activity has a different proportion of JMs. The proposed activity recognition algorithm aims to use this discriminative information to provide grazing and rumination bouts [16]. To achieve a low computational cost, tasks within each stage have been simplified whenever it was possible. The input of the system is the sound signal produced during foraging activities. Three activities are considered: rumination, grazing, and other activities. The latter category includes any activity other than rumination or grazing (i.e. from silence to different noises).

Detection and classification of JMs are performed with the algorithm presented in Section III-A. Then, the feeding activity is recognized by analyzing fixed-length segments of the acoustic signal. JMs that are detected and classified within a segment are stored in a segment buffer. The rate of JMs in a segment and the proportions of their types are computed to feed the last processing stage. At this point, activity classification could be seen as a simple task, but an exploratory data analysis on the training set has shown a complex underlying distribution of the segment features (rate, %c, %b, %cb). The rate of recognized JMs during rumination and grazing is expected to be in the range from 0.75 to 1.40 Hz (Fig. 5). By contrast, the rate of JMs identified during other activities presents a lower frequency. The overlapping among rate distributions of activities is part of the problem.

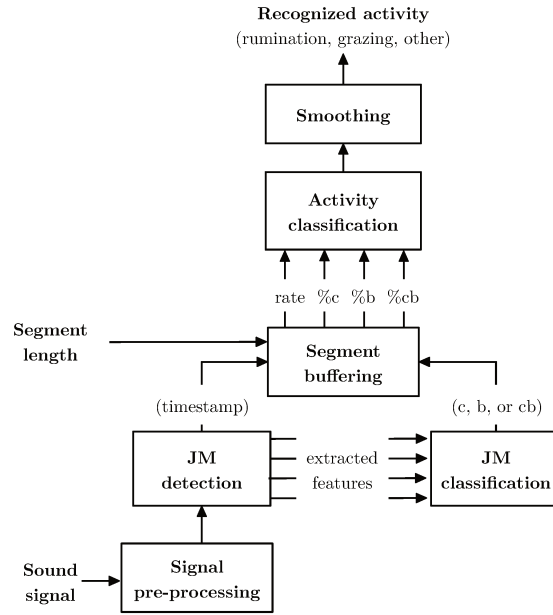


Fig. 4. General diagram of the activity classification algorithm.

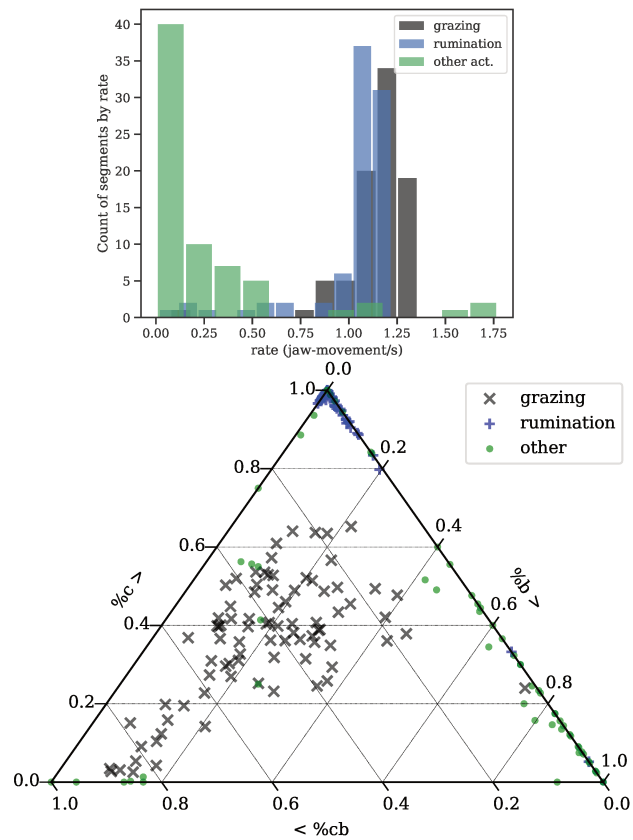


Fig. 5. Distribution and proportions of jaw movements.

The triangle plot in Fig. 5 shows the proportions of the identified JMs for several segments of the training set. Proportions of a single segment always sum to 1.0. The top corner corresponds to 100% of chews, the bottom left corner corresponds to 100% of chew-bites, and the bottom right corner corresponds to 100% of bites. Points inside the triangle correspond to segments composed by more than one type of JMs. For example, while rumination is mainly composed of chews, grazing has a diversity of JMs compositions. During other activities, bites are the most assigned type of JMs.

Distributions of segment features show that the recognition of JMs within grazing and rumination activities is not perfect. For example, CBIA detects a few bites during rumination, which is not actually true. Thus, the problem of distinguishing between activities requires a powerful method to handle these errors. In this study, the use of a simple method of machine learning is proposed. Activity classification is performed by a trainable model, such as a multilayer perceptron or a decision tree, which assigns an activity label to the segment. In this way, at the end of the processing stages, each segment of the input signal has a label that indicates if it corresponds to rumination, grazing, or other activity. Finally, a smoothing process is applied over the sequence of labeled segments to remove short gaps and thus reduce fragmentation of activity bouts. Thus, long recognized bouts are encouraged, which mimics the typical length of activity bouts.

IV. THE EMBEDDED SYSTEM

The design of a battery-powered embedded system requires a detailed analysis of each subsystem to minimize size, cost, and, principally power consumption. The sensor unit device has been designed as a trade-off between minimizing power consumption and parameter estimation accuracy. It is located on the neck of the animal, just behind the head. The embedded device was built around a microcontroller (MCU) and it comprises four interconnected modules: i) the signal conditioning, ii) the data-logging and communications, iii) the power supply and energy harvesting, and iv) the digital processing, as shown in Fig. 6.

The MCU for this application (MCF51JM128, NXP Semiconductors) was chosen based on its availability in the local market, power consumption, computational power, analog ports, communication resources (SPI, USART and USB), internal clocking resources, and a real-time clock module.

A. Signal conditioning

The signal conditioning module acquires and conditions the animal feeding sound. It senses the sound with an electret microphone facing inward on his forehead. The signal bandwidth is limited to 2 kHz to minimize the quantization noise with an eighth-order Butterworth low pass filter. An automatic gain control (AGC) amplifier is used to maximize the signal-to-noise ratio. The AGC output is connected to one of the analog input channels of an 8 bits A/D converter in the MCU. The AGC applied gain level is delivered to the MCU through another analog input channel. The AGC output is also connected to a low pass filter to detect when an acoustic signal is present. It is compared with a defined threshold reference voltage to detect when there is feeding activity and to wake-up the MCU (IRQ input).

B. Digital processing and data-logging

The embedded firmware is organized into four tasks: i) signal conditioning and preprocessing, event detection, and classification, ii) data logging, iii) internal and external communication, and iv) device configuration. The software architecture is driven by four possible interrupts to wake-up the MCU from sleep mode and execute one of these tasks.

When the voltage comparator detects a sound level above a threshold, an A/D conversion begins. Every 500 μ s a finished A/D conversion wake-up the MCU and execute the algorithm to detect and classify JMs. The software extracts the information that characterized the ingestive activities and accumulates the partial results, and the MCU is set into sleep mode until the next A/D conversion is completed. Finally, after 5 minutes since the last sound level activation, the device is hibernated.

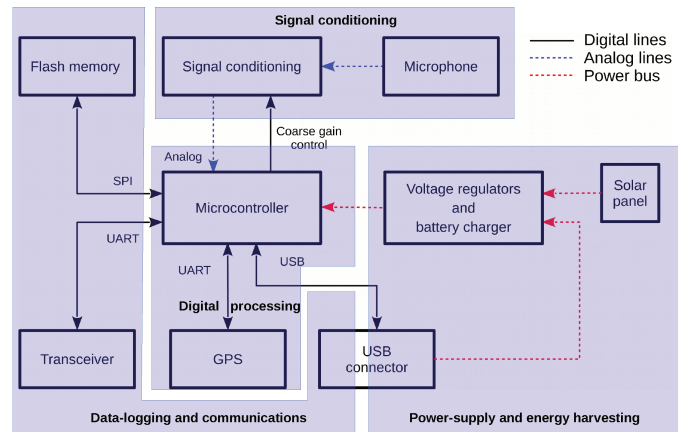


Fig. 6. Block diagram of the embedded system.

For every 30 min interval, a data frame is built gathering instant information from the GPS (time and position) and information related to the feeding activities (i.e. quantity, average time, and average energy of each potential event). Finally, the data frame is stored as a text file in one of the two flash memories, and the MCU is set into sleep mode.

Every 15 min the PC starts and controls the wireless communication with the sensor device, generating an interrupt in the MCU. The reception module is on continuously. When a request package is received from the PC, the MCU turns on the transmission module, transmits the requested data frames, waits for the reception acknowledgment and then it turns off the transmission module and the MCU is set into sleep mode. The final interrupt is triggered when a USB communication is established to transfers the data stored in the flash memories to the PC or to set a new sensor configuration.

C. Communications

Internal sensor communications are managed by the MCU. A 3 V coin cell battery is incorporated as a backup power supply for the GPS. The USB module provides an On-The-Go (OTG) dual-role controller. The additional wireless communication mode operates at 433/470 MHz with a transfer rate of up to 37.5 kbps. All these features simplify the hardware and software required for communication.

D. Power supply

The device has been designed to minimize power consumption whereas it is able of harvesting all the energy needed for its operation. Therefore, three complementary approaches were used to develop the energy management scheme: i) A combined duty-cycling and data-driven operating scheme, driven by data, to operate only when relevant information is available. It is implemented through the firmware in the MCU. ii) An energy harvesting scheme, by using a 1W solar panel able to recharge the two Li-Ion batteries when the sensor device is operating outdoors; iii) A USB port when the device is connected to a PC or an energy source.

E. Autonomy analysis

To determine the autonomy of the sensor device it is mandatory to know the time it will be on the active mode, due to the sum of the times of the rumination and grazing activities. According to the circadian rhythm of the cattle feeding behavior described in the sensor device will be on the active mode around 60% of the daily time.

The measured charge consumption per hour during active mode is 34.9 mAh. Considering the battery capacity of 5000 mAh and in the absence of the solar panel, the sensor device would have an autonomy

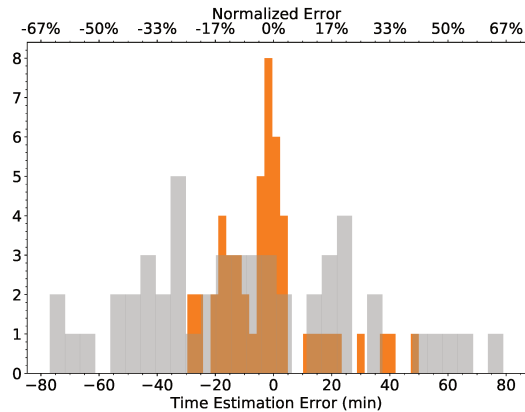


Fig. 7. Time estimation error of rumination for proposed activity recognition system (orange) and Hi-Tag (gray). Top axis is normalized with the length of segments analyzed (2 hours).

of $T_A = 143.27$ h. Otherwise, the instant current provided by the solar panel is 120 mA during daylight hours. If we suppose it is maintained during 6h, the average charge provided is $Q_{SP} = 30.0$ mAh.

According to the circadian cycle of the cattle feeding behavior, the average charge consumed per hour for a whole day is given by $Q_{CR} = 20.6$ mAh. Thus, the net charge balance is $Q_{SP} - Q_{CR} = 9.4$ mAh. This shows that the harvested power is sufficient to energize the sensor device.

To determine the worst-case conditions, the sensor device was tested in active mode during 5 days outdoors with the batteries fully charged initially. During the day, the sensor device is powered entirely by the solar panel and, at the same time, the batteries are partially recharged. However, it is not enough to accomplish a full charge of the batteries. As expected, during the night, the batteries supply the embedded system and its voltage decreases considerably. At the end of the test, the sensor device still worked properly and the voltage tends to stabilize close to 3.7 V, which is the nominal battery voltage.

V. RESULTS

A comparison of the rumination time estimation obtained by a commercial system (Hi-Tag) and the proposed system was performed. The Hi-Tag system summarizes the total time the animal spent ruminating during two-hour chunks [19]. Raw data and timestamps of rumination bouts within a two-hour chunk are not available [20]. Therefore, the estimations with the proposed system were aligned, and the total duration of rumination was summarized to match the same two-hour chunks of the Hi-Tag system. The comparison was made with a total of 53 two-hour chunks from all the recordings analyzed as it was done in [21]. Since the Hi-Tag is a commercial system, its computational cost was not available to be considered in the analysis.

The results of time estimation error for rumination are shown in Fig. 7. The medians of the distributions are -2.91 min and -13.55 min for the proposed system and the Hi-Tag system, respectively. Negative medians imply that both systems tend to underestimate the rumination time. The proposed system was more accurate and resulted in a narrower error distribution. While the error dispersion for the proposed system is in the range $(-30, +50)$ min, the distribution corresponding to the Hi-Tag is wider and it is in the range $(-80, +80)$ min. In practical terms, these errors are very high since they are in the same order of magnitude of the two-hour chunks analyzed [16].

Fig. 8 shows the evolution of voltage and current of the batteries. During the day, the current provided by the batteries decreases to zero, when the device is powered entirely by the solar panel and, at the same time, the batteries are partially recharged. However, it is not enough to accomplish a full charge of batteries. As expected, during the night and periods when there is no solar light incident on the solar panel, the batteries supply the embedded system and its voltage decreases considerably. During each daylight,

there is a small peak in the current consumption, which corresponds to the projection of a shadow above the solar panel making that much of the current consumed by the device has to be provided by the batteries during this time. At the end of the test, the device still worked properly and the voltage tends to stabilize close to 3.7 V, which is the nominal battery voltage.

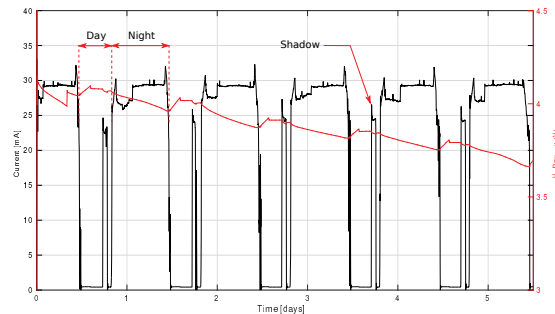


Fig. 8. Current consumption (black line) and voltage of the battery pack (red line).

VI. CONCLUSIONS

In the tests performed, the acoustic methods presented have achieved good performance rates. In addition, they provide comprehensive information (short- and long-term) of the foraging behavior of the ruminant. The low computational cost of the proposed methods allows its real-time execution in a simple embedded system. The activity recognition method showed better performance than a commercial system, under certain conditions. Robust recognition of rumination and grazing activities is another challenge to be addressed in future studies. The developed embedded system has shown suitable communication, processing and autonomy characteristics.

ACKNOWLEDGMENT

The authors would like to thank Julio Galli, Marcelo Larripa, Alejandra Planisich, and Martín Quinteros from Facultad de Ciencias Agrarias, Universidad Nacional de Rosario for their assistance in animal management and gathering data. This study has been funded by Universidad Nacional del Litoral, PACT CAID 2011-525, Universidad Nacional de Rosario, projects 2013-AGR216 and 2016-AGR266, Agencia Santafesina de Ciencia, Tecnología e Innovación (ASACTEI), project 2010-216-16 (2016), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), project 2017-PUE-sinc(i).

REFERENCES

- [1] K M Svennersten-Sjaunja and G Pettersson. Pros and cons of automatic milking in europe. *J. Anim. Sci.*, 86(13 Suppl):37–46, March 2008.
- [2] A R Frost, C P Schofield, S A Beulah, T T Mottram, J A Lines, and C M Wathes. A review of livestock monitoring and the need for integrated systems, 1997.
- [3] J (ed) Hodgson and A W (ed) Illius. The ecology and management of grazing systems. 1998.
- [4] Mette S Herskin, Lene Munksgaard, and Jan Ladewig. Effects of acute stressors on nociception, adrenocortical responses and behavior of dairy cows. *Physiol. Behav.*, 83(3):411–420, December 2004.
- [5] Daniel J Bristow and David S Holmes. Cortisol levels and anxiety-related behaviors in cattle. *Physiol. Behav.*, 90(4):626–628, March 2007.
- [6] J G Welch. Rumination, particle size and passage from the rumen, 1982.
- [7] Robert J Kilgour. In pursuit of “normal”: A review of the behaviour of cattle at pasture. *Appl. Anim. Behav. Sci.*, 138(1):1–11, 2012.
- [8] Thatianne Coutheux Trindade. Influência do controle glicêmico no potencial de crescimento fetal em pacientes com diabetes melito gestacional.
- [9] M A Benvenuti, D R Pavetti, D P Poppi, I J Gordon, and C A Cangiano. Defoliation patterns and their implications for the management of vegetative tropical pastures to control intake and diet quality by cattle, 2016.
- [10] E A Laca, E D Ungar, N G Seligman, M R Ramey, and M W Demment. An integrated methodology for studying short-term grazing behaviour of cattle, 1992.

- [11] Eugene David Ungar and Steven Mark Rutter. Classifying cattle jaw movements: Comparing IGER behaviour recorder and acoustic techniques, 2006.
- [12] J R Galli, C A Cangiano, M A Pece, M J Larripa, D H Milone, S A Utsumi, and E A Laca. Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle. *Animal*, 12(5):973–982, May 2018.
- [13] Philip U Alkon, Yosef Cohen, and Peter A Jordan. Towards an acoustic biotelemetry system for animal behavior studies, 1989.
- [14] Julio R Galli, Carlos A Cangiano, Diego H Milone, and Emilio A Laca. Acoustic monitoring of short-term ingestive behavior and intake in grazing sheep, 2011.
- [15] José O Chelotti, Sebastián R Vanrell, Julio R Galli, Leonardo L Giovanini, and H Leonardo Rufiner. A pattern recognition approach for detecting and classifying jaw movements in grazing cattle, 2018.
- [16] José O Chelotti, Sebastián R Vanrell, Luciano Martinez-Rau, Diego H Milone, Santiago A Utsumi, Julio R Galli, H Leonardo Rufiner, and Leonardo L Giovanini. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Comput. Electron. Agric.* (Under review since May 2019).
- [17] Richard O Duda, Peter E Hart, and David G Stork. *Pattern Classification*. John Wiley & Sons, November 2012.
- [18] José O Chelotti, Sebastián R Vanrell, Diego H Milone, Santiago A Utsumi, Julio R Galli, H Leonardo Rufiner, and Leonardo L Giovanini. A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle. *Comput. Electron. Agric.*, 127:64–75, 2016.
- [19] K Schirmann, M A G von Keyserlingk, D M Weary, D M Veira, and W Heuwieser. Technical note: Validation of a system for monitoring rumination in dairy cows. *J. Dairy Sci.*, 92(12):6052–6055, December 2009.
- [20] C Goldhawk, K Schwartzkopf-Genswein, and K A Beauchemin. Technical note: validation of rumination collars for beef cattle. *J. Anim. Sci.*, 91(6):2858–2862, June 2013.
- [21] Sebastián R Vanrell, José O Chelotti, Julio R Galli, Santiago A Utsumi, Leonardo L Giovanini, H Leonardo Rufiner, and Diego H Milone. A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle. *Comput. Electron. Agric.*, 151:392–402, August 2018.

Anexo F

**Sistema para el seguimiento y
evaluación de actividades alimentarias
en producción animal basado en la
energía del sonido**

Sistema para el seguimiento y evaluación de actividades alimentarias en producción animal basado en la energía del sonido

Campo de la invención

La presente invención proporciona un método y un dispositivo para el seguimiento y evaluación de la actividad alimentaria de individuos en rodeos de bovinos en tiempo real, mediante la recolección, procesamiento y análisis de la información característica relacionada con la actividad alimentaria manifestada en los sonidos producidos por los animales al ingerir y rumiar el forraje. Esta información se procesa de forma tal que reduce los impactos negativos del entorno (ruidos y perturbaciones) donde opera y hábitos y situaciones anómalas de los animales que afectan los resultados obtenidos, proveyendo resultados más precisos y robustos que los métodos disponibles hasta el momento.

Estado de la Técnica

Los productores ganaderos, tanto de leche como de carne vacuna (cría e invernada), se enfrentan a diferentes dificultades cuando tratan de mejorar la producción y mantener la rentabilidad de sus establecimientos. Estas dificultades se originan en la falta de información precisa sobre el manejo de la alimentación y el estado de salud de los animales. Esto conlleva a una baja eficiencia de conversión del alimento en producto (leche y/o carne) debido a una gran pérdida de recursos en el proceso de alimentación.

Varios métodos y dispositivos se han desarrollado, en un intento de resolver este problema de los productores, permitiendo la supervisión automatizada de la salud animal, brindando información sobre el estado nutricional de los animales de manera de mejorar la eficiencia de la alimentación en especial y del sistema de producción en general. Al final de esta memoria descriptiva se incluyen las referencias bibliográficas más relevantes que describen los intentos previos de investigadores, empresas e incluso de este mismo grupo de trabajo.

Se describen a continuación algunos documentos de patentes de invención que definen intentos previos para abordar el campo de la técnica descrito.

La patente de Estados Unidos N° 5.901.660 describe un sistema de comunicaciones que permite a los animales llevar transpondedores auriculares, que pueden ser accedidos selectivamente de forma remota cuando un animal particular está próximo a una antena simple. La comunicación entre el transpondedor y el receptor se realiza a través de un acoplamiento de campo próximo, tal como un acoplamiento de campo eléctrico. Cuando se detecta la dirección de un animal, el transpondedor del animal responde con un flujo de datos modulado en frecuencia. La antena del receptor puede estar posicionada en un recipiente de alimentación/agua para detectar cuando un animal en particular se alimenta o alrededor de una puerta para detectar cuando

un animal pasa a través de la misma.

La patente de Estados Unidos N° 5.984.875 describe un dispositivo ingerible (bolo) por los animales para el seguimiento de los parámetros fisiológicos. El bolo incluye los circuitos necesarios para almacenar un código de identificación, detectar un parámetro fisiológico y transmitir una señal de datos que incluye toda esta información. El sistema también incluye un receptor para recibir señales de los datos transmitidos desde el bolo. Un parámetro utilizado frecuentemente es la temperatura.

Otro bolo ingerible está descrito en la patente de Estados Unidos N° 6.059.733, que es utilizado para determinar la temperatura corporal interna de un animal rumiante. El bolo incluye un sensor de temperatura y un transmisor. Las patentes de Estados Unidos N° 6.285.897 y N° 6.689.056 describen un sistema ambulatorio para detectar, registrar y analizar los parámetros fisiológicos tales como el pH, temperatura y presión, dentro del esófago u otros lúmenes del cuerpo. El sistema incluye un sensor implantable y un transmisor de radiofrecuencia, un receptor y una grabadora externa, y un paquete de software de análisis.

La patente de Estados Unidos N° 6.535.131 B1 describe un aparato y método para identificar automáticamente los patrones de sonidos emitidos por los animales. El aparato identifica automáticamente cuando un animal está en peligro a partir de reconocer un patrón determinado en el sonido producido por el animal seleccionado. Este reconocimiento se realiza a partir de la comparación de los sonidos registrados con patrones de audio pre-almacenados correspondientes a los respectivos sonidos que se espera produzca el animal cuando se encuentra en diferentes situaciones de interés. Cuando se encuentra una coincidencia lo suficientemente buena, se envía una señal de forma automática a una unidad de comunicación remota.

La patente de Estados Unidos N° 8.152.734 B2 describe un método y sistema para el diagnóstico de las enfermedades respiratorias de la especie bovina utilizando técnicas de auscultación. Los espectrogramas de las señales acústicas registradas se comparan con los datos existentes, permitiendo el diagnóstico de un animal enfermo. Los sonidos pulmonares se obtienen mediante el uso de un estetoscopio electrónico y se almacenan como datos digitales. La señal registrada es procesada y acondicionada para organizar los datos en el formato deseado y para eliminar el ruido indeseable asociado con los sonidos grabados. Luego un algoritmo de análisis se aplica a los datos para evaluar el estado del pulmón a partir del cálculo de un índice. Estos índices se clasifican en diferentes niveles de acuerdo con la patología percibida, en base a los datos de referencia que clasifican los índices. De los índices obtenidos, un reconocedor puede asociar el índice con un diagnóstico, pronóstico y un tratamiento recomendado. El software de análisis genera los índices de pulmón a partir de los sonidos grabados y también puede proporcionar una presentación visual de los posibles diagnósticos, así como los tratamientos recomendados.

La patente de Estados Unidos N° 7.350.481 B2 describe un método y sistema para el monitoreo de las condiciones fisiológicas y/o la idoneidad de la alimentación para

animales rumiantes a partir de la detección de las acciones del animal que indican la actividad de rumia y acumula el tiempo de las actividades de los rumiantes durante un período de tiempo predeterminado para proporcionar indicadores de la condición fisiológica del animal y cambios en su alimentación. Una de las formas de realización de esta idea se basa en la detección del ritmo de masticación del animal durante la alimentación a partir del registro de los sonidos mediante un micrófono colocado contra el cuello del animal. De acuerdo al ritmo detectado se puede identificar si el animal se está alimentando o rumiando. Una segunda forma de realización de las ideas descritas se basa en el sensado de las regurgitaciones desde el rumen a la boca del bolo alimenticio a través de un collar con dos interruptores mecánicos alrededor del cuello del animal. Una de las limitaciones de esta patente es la necesidad de dos sensores (tipo microswitch) que detectan el paso del bolo alimenticio y la dirección del mismo para discriminar si el animal se está alimentando o rumiando. Un problema técnico que no resuelve el dispositivo descrito en ese documento es que el ruido ambiente no puede ser filtrado, por lo que la señal que se procesa incorpora una gran cantidad de fuentes de error. Otra limitación es que este dispositivo posee parámetros que deben ser ajustados por el usuario para cada animal.

Un enfoque alternativo al problema de monitoreo de actividades alimentarias se basa en el reconocimiento de los movimientos de la mandíbula de los animales (eventos masticatorios - EM -) a partir de los sonidos producidos. Con esta información no solo es posible determinar las actividades alimentarias realizadas por el animal, sino también cuantificarlas. En esta línea de trabajo, Milone y col. (2009) utilizaron conceptos del campo de reconocimiento de voz automático para desarrollar un algoritmo basado en modelos ocultos de Markov (HMM) que identifica y clasifica las masticaciones, los arranques y los movimientos compuestos masticación-arranque en el ganado ovino. El algoritmo logra tasas medias del 89%, 58% y 56% para reconocer las masticaciones, arranques y masticaciones-arranques, respectivamente.

Luego, Clapham y col. (2011) desarrollaron un sistema de detección automática de masticaciones para el pastoreo de ganado en pasto perenne mixto. El algoritmo solo detecta arranques de ganado bovino para diferentes tipos de pastos y alcanza una tasa de detección de alrededor del 95%. El sistema propuesto utiliza un programa comercial de análisis de sonido para el sistema operativo Windows (desarrollado por la empresa Engineering Design) para la detección de los eventos masticatorios. Las señales utilizadas deben grabarse en alta calidad (frecuencia de muestreo de 44,1 kHz y una resolución de 16 bits), lo cual implica cargas computacionales elevadas. El algoritmo utiliza un filtro paso alto con una frecuencia de corte en 600 Hz para atenuar las perturbaciones como el viento, y los eventos se detectan analizando la señal en la banda desde 17 kHz a 22 kHz. Sin embargo, se encontraron algunas dificultades al tratar de escalar los experimentos de monitoreo para varios días, incluyendo la capacidad de los dispositivos de almacenamiento y fuentes de energía.

El método requiere de una calibración cuidadosa antes de utilizarlo en cada condición experimental.

Posteriormente, Milone y col. (2012) desarrollaron un algoritmo que extiende el uso de modelos ocultos de Markov para reconocer los tres tipos de sonidos producidos por el ganado durante su alimentación (masticación, arranque y masticación-arranque) en alfalfa (alto y bajo) y festuca (alto y bajo). El algoritmo logra tasas medias de reconocimiento del 80% para las diferentes pasturas y alturas, con un máximo del 94% para la festuca corta y un mínimo de 65% para la alfalfa corta.

Navon y col. (2012) propusieron un nuevo algoritmo basado en el análisis de las características de dominio de tiempo (forma, intensidad, duración y secuencia) del sonido registrado para la detección de los movimientos de la mandíbula. El algoritmo utiliza técnicas de aprendizaje automático para identificar los ruidos de fondo presentes en los sonidos grabados. Este algoritmo no requiere calibración e identifica los movimientos de la mandíbula de acuerdo con las características temporales de la señal que se definen en términos relativos. El desempeño del algoritmo en un entorno de bajo ruido utilizando un micrófono de solapa colocado en la frente, logró resultados similares a los trabajos anteriores.

Luego, Tani y col. (2013) desarrollaron un sistema de monitoreo mediante el uso de un acelerómetro de un solo eje en lugar de los micrófonos, para registrar los movimientos y vibraciones de la mandíbula. El algoritmo propuesto utiliza técnicas de procesamiento de señal para eliminar los ruidos de fondo de las señales grabadas y luego se extraen los patrones característicos de la alimentación. El algoritmo sólo detecta y clasifica actividades de ingesta y rumia, pero no identifica los eventos individuales. El sistema desarrollado tiene dos problemas relacionados con los sensores y los algoritmos utilizados:

- La similitud espectral de los eventos de rumia y alimentación, lo que podría ser un problema el análisis de señales crudas con mala relación señal / ruido, y
- La propiedad no estacionaria de ruido (ruido de fondo o ruidos inesperados que podrían ocurrir), lo que lo hace cambiar sus propiedades estadísticas en el tiempo.

Además, la estimación espectral de señales muestreadas a alta frecuencia tiene un alto coste computacional que hace difícil su implementación en un sistema embebido. La idea de utilizar un acelerómetro para monitorear eventos masticatorios es interesante, pero el sensor tiene errores intrínsecos que deben ser analizados para una grabación más fiable, así como un análisis cuidadoso de los movimientos y posiciones del animal cuando rumia o pastorea. Los autores realizaron estos experimentos solo en establos.

La patente de Argentina AR100013B1 describe un método y sistema para el monitoreo y cuantificación de las actividades alimentarias de animales rumiantes a

partir de un análisis estadístico simple de la señal sonora. Utilizando un estadístico de la señal sonora calculado sobre ventanas temporales largas (una hora), a partir del cual el algoritmo determina si el animal realiza actividad alimentaria, o no, y acumula el tiempo de las actividades realizadas durante el tiempo de la ventana de análisis para proporcionar indicadores de cambios en su alimentación. Una de las formas de realización de esta idea se basa en la detección del ritmo de masticación del animal durante la alimentación a partir de la adquisición y registro de los sonidos mediante un micrófono colocado en la cabeza del animal. De acuerdo al ritmo detectado se puede identificar si el animal se está alimentando o rumiando. Un análisis de la complejidad de los algoritmos implementados en esta patente demuestra que los requerimientos de almacenamiento, comunicaciones y cómputo necesarios para su implementación hacen que no sea posible el monitoreo en tiempo real de animales individuales, limitando su uso a muestreos del comportamiento de los animales de un rodeo.

Posteriormente, la patente de Argentina AR103025A1 describe un método y un sistema para el monitoreo y cuantificación de las actividades alimentarias de animales rumiantes a partir de la detección y clasificación de los EM que el animal realiza durante su alimentación. Este proceso se realiza analizando la posición de la cabeza y las características temporales de la envolvente del sonido producido por el animal durante su alimentación. El sistema propuesto consiste en una plataforma de hardware que implementa un algoritmo que procesa la señal sonora para acondicionarla, luego calcula su envolvente, extrae las características que permiten detectar y clasificar un EM (duración, amplitud máxima, forma de evento, simetría) y finalmente aplica un conjunto de reglas para detectar la ocurrencia de un EM; y si se esto ocurre clasificarlo a partir de un conjunto de reglas que relaciona las características con los EM. Luego, el algoritmo genera información estadística de los EM ocurridos durante una ventana de tiempo de duración fija, para clasificar la ventana en alguna de las posibles actividades: rumia, pastoreo o nada. La clasificación se realiza aplicando un conjunto de reglas que relaciona la información estadística de la ventana con la actividad alimentaria. Un análisis de la complejidad de los algoritmos implementados en esta patente demuestra que los requerimientos de almacenamiento, comunicaciones y cómputo necesarios para su implementación hacen factible el monitoreo del comportamiento de los animales de un rodeo en tiempo real.

Chelotti y col. (2016) propuso un algoritmo de detección y clasificación de EM basado en el análisis de características de las señales de sonido en el dominio del tiempo, logrando una tasa de éxito de reconocimiento de eventos similar al algoritmo presentado por Millone y col. (2012) pero con un costo computacional cien veces menor y tiempos de procesamiento mil veces más rápidos, permitiendo su implementación en sistemas embebidos basados en microcontroladores (Deniz y col. 2017). El algoritmo logra tasas medias de reconocimiento del 77% para las diferentes

pasturas y alturas, con un máximo del 79% para la alfalfa alta y un mínimo de 74% para la alfalfa corta. Luego, Chelotti y col. (2018) modificaron el algoritmo anterior a partir de utilizar mejores algoritmos de procesamiento de señales e incorporar un modelo de clasificación más sofisticado, mejorando el desempeño y la robustez del algoritmo sin incrementar el costo computacional. El algoritmo logra tasas medias de reconocimiento del 87% para las diferentes pasturas y alturas, con un máximo del 90% para la festuca corta y un mínimo de 80% para la alfalfa corta.

Mientras que los transductores acústicos son el método predilecto para abordar el reconocimiento de EM, los acelerómetros son el transductor más utilizado para el reconocimiento de actividades alimentarias de largo plazo. Esto se debe a que muchos algoritmos existentes, además de reconocer actividades alimentarias, también reconocen otras actividades relacionadas a posición o movimiento del animal. Martiskainen y col. (2009) fueron capaces de identificar dentro de un galpón las actividades alimentarias de rumia e ingesta de silaje, y los movimientos acostarse, ponerse de pie, estar de pie, estar acostado y caminar normal y cojo. Diosdado y col. (2015) utilizaron un acelerómetro y un podómetro para identificar si el animal localizado dentro de un establo se encuentra acostado, de pie o ingiriendo silaje. Rayas-Amor y col. (2017) utilizó un equipo comercial HOBO (Onset Computer Corporation, Bourne, MA, Estados Unidos) que almacena los datos obtenidos de un sensor de aceleración y de inclinación para cálculos tiempos de rumia y pastoreo. Sin embargo, la clasificación está basada en las componentes estáticas de los ejes del acelerómetro, lo cual no es práctico en situaciones reales donde un ligero movimiento del animal podría cambiar el valor de las componentes estáticas dentro del mismo comportamiento (Benaissa y col. 2019.a). Benaissa y col. (2019.b) desarrolló un método utilizando un acelerómetro para reconocer las actividades de rumia y alimentación, estando esta última conformada por la actividad de pastoreo y bebida de agua. Finalmente, Shen y col. (2020) obtuvo mejores resultados que Benaissa y col. (2019.b) a un menor costo computacional posibilitando implementar dicho método en un dispositivo embebido.

Aunque algunos autores indican que el uso de acelerómetros presenta mayor inmunidad al ruido que el uso de micrófonos para reconocer actividades alimentarias, ninguno de los métodos anteriores basados en acelerómetros evalúa los efectos de los cambios de la posición de los acelerómetros. En este sentido, deberían realizarse experimentos de mayor duración para ver este efecto. Además aún no se ha desarrollado un algoritmo reconocedor de EM utilizando acelerómetros como fuente de información. Lo cual es necesario para calcular la materia seca ingerida por el animal.

Un algoritmo reconocedor de actividades alimentarias utilizando señales acústicas, a partir de analizar sus regularidades fue propuesto por Vanrell y col. (2018), obteniendo resultados superiores a los del equipo comercial Hi-Tag (SCR Engineers Ltd., Netanya, Israel). Este algoritmo reconoce las actividades de rumia, pastoreo y

otras actividades, y estima sus respectivas duraciones. Sin embargo, este algoritmo requiere tener toda la señal registrada previo a realizar el análisis, lo cual imposibilita ser ejecutado en tiempo real y ser implementado en un dispositivo electrónico de bajo consumo con capacidad limitada de almacenamiento.

Chelotti y col. (2020) propuso un algoritmo para la clasificación de actividades alimentarias (rumia, pastoreo y otro). Este algoritmo analiza los EM reconocidos a partir del sonido, utilizando el algoritmo de Chelotti y col. (2018), en grupos obtenidos en ventanas temporales fijas para reconocer las actividades de rumia y pastoreo. El algoritmo calcula estadísticas de los eventos identificados en un periodo de tiempo y luego utiliza un modelo computacional (perceptrón multicapa y árbol de decisión, entre otros) para determinar el tipo de actividad alimentaria realizada (rumia, pastoreo y otros). Este algoritmo se evaluó y comparó con los resultados obtenidos por Vanrell y col. (2018) . El algoritmo propuesto obtuvo resultados superiores en la estimación de los tiempos de pastoreo y rumia. El buen desempeño y el bajo costo computacional hacen factible su implementación y ejecución en tiempo real en un sistema embebido de bajo costo, lo que permite el desarrollo de un dispositivo portátil para el seguimiento en tiempo real del comportamiento de alimentación de los rumiantes.

Li y col. (2021) propuso un algoritmo de detección y clasificación de EM basado en el uso de modelos computacionales de aprendizaje profundo (deep learning) aplicados a las señales de sonido, logrando una tasa de éxito de reconocimiento de eventos similar al algoritmo presentado por Chelotti y col. (2018). El algoritmo propuesto combina las señales de sonido con características frecuenciales de las mismas, por lo que su costo computacional y los requerimientos de almacenamiento son muy superiores a los del sistema propuesto por Millone y col. (2012) y Chelotti y col. (2018). Los autores no reportaron los tiempos de procesamiento utilizados durante los estudios, porque no fueron realizados en tiempo real.

Esencialmente, la presente invención resuelve el problema de adquirir, procesar, analizar y comunicar en tiempo real la información necesaria para caracterizar y cuantificar las actividades alimentarias de los animales en rodeos en condiciones de campo o pastoreo, aunque también es aplicable a rodeos confinados en corrales o establos. A diferencia de otros equipos disponibles en el mercado, la presente invención solo utiliza los sonidos generados por los animales durante las actividades alimentarias (pastoreo y rumia) para determinar el tipo de actividad alimentaria realizada por el animal, sus parámetros característicos y la cantidad de alimento ingerido. A diferencia de propuestas anteriores, la presente invención provee resultados más robustos ya que utiliza la envolvente y la energía del sonido en lugar de solo la envolvente, que es más sensible a los ruidos y perturbaciones. Además, es capaz de diferenciar entre masticaciones realizadas durante la rumia de las realizadas durante el pastoreo, proveyendo cuatro clases de eventos lo que conduce

a resultados más robustos y precisos de las estimaciones de actividades alimentarias y la cantidad de materia seca ingerida.

Resumen de la invención

Breve descripción de las figuras

La presente invención se entenderá y apreciará mejor mediante una descripción detallada en las siguientes figuras:

Figura 1: diagrama de bloques que ilustra una posible realización del dispositivo de acuerdo con la invención mostrando sus bloques;

Figura 2: diagrama de bloques que ilustra una posible realización del bloque de adquisición y procesamiento de las señales;

Figura 3: ilustración esquemática de la ubicación de los sensores y el dispositivo en la cabeza de un animal;

Figura 4: un ejemplo de la salida de la unidad de sensor de sonido para a) pastoreo durante 200 s y b) rumia durante 200 s, con detalles de eventos masticatorios asociados y los porcentajes promedio por actividad;

Figura 5: un ejemplo de la salida de la unidad de sensor de sonido para a) una masticación de rumia, b) una masticación de pastoreo, c) un arranque, y d) una masticación-arranque y las características asociadas;

Figura 6: diagrama de bloques que ilustra un posible realización del bloque de detección y clasificación de eventos;

Figura 7: diagrama de bloques que ilustra un posible realización del a) cálculo de energía por frame y b) el detector de eventos;

Figura 8: diagrama de la organización de los datos almacenados en a) el buffer de envolvente y b) el buffer de energía, así como los parámetros asociados;

Figura 9: diagrama de bloques que ilustra un posible realización del bloque de análisis de actividades alimentarias basados en eventos masticatorios identificados;

Figura 10: diagrama de flujo que ilustra el funcionamiento del algoritmo de reconocimiento de las actividades alimentarias basados en eventos masticatorios identificados;

Figura 11: diagrama de bloques que ilustra un posible realización del bloque de análisis de actividades alimentarias basados en características del sonido;

Figura 12: diagrama de bloques que ilustra la realización del bloque de almacenamiento y comunicaciones;

Figura 13: diagrama de bloques que ilustra la realización del bloque de alimentación.

Figura 14: diagrama de bloques que ilustra la implementación a través de dos dispositivos utilizando microcontroladores y circuitos integrados asociados.

Figura 15: diagrama de bloques que ilustra la implementación del dispositivo utilizando un microcontrolador y circuitos integrados asociados;

Figura 16: matrices de confusión de los eventos masticación pastoreo (GC), masticación-arranque (CB), arranque (B) y masticación rumia (RC) en las combinaciones de las bases de datos;

Figura 17: matrices de confusión de los métodos a) propuesto usando sólo eventos identificados; a) propuesto usando sólo características del sonido; c) patente AR103025A1 y d) patente AR100013B1.

Figura 18: Tasa de reconocimiento para distintos niveles de ruido de los métodos a) propuesto usando sólo eventos identificados (BUFAR-CBEBA); b) propuesto usando sólo características del sonido (JMFAR) y c) patente AR103025A1 (BUFAR-CBIA).

Se apreciará que por simplicidad y claridad de la ilustración, los elementos mostrados en las figuras no han sido necesariamente dibujados a escala. Por ejemplo, para mayor claridad las dimensiones de algunos de los elementos pueden estar exageradas con relación a otros. Además, cuando se considera apropiado, los números de referencia pueden ser repetidos entre las figuras para indicar elementos correspondientes o análogos.

Descripción detallada de la invención

La presente invención proporciona un método y un dispositivo para el seguimiento y evaluación robusta de la actividad alimentaria de individuos en rodeos de rumiantes, mediante la recolección, procesamiento y análisis de la información característica relacionada con la ingestión y rumia del forraje, que se traduce en sonidos particulares producidos por los animales. Esta información procesada y analizada de forma adecuada reduce los errores que podrían producirse debido a hábitos del animal (lamido, olfateo, rascarse, rumiar caminando y acostada, entre otros) y situaciones anómalas (enfermedades y estrés térmico, entre otros) que modifican el comportamiento de los animales y condiciones ambientales que afectan las mediciones, lo cual deteriora los resultados obtenidos.

Definiciones.

De acuerdo a la presente invención, el término “**sonido**” es equivalente a señal sonora, vibraciones, señal auditiva, de manera que se los usa a lo largo del texto de manera intercambiable.

Mediante el término “**adquirir**”, en la presente invención se refiere a obtener, captar o capturar, a través de un dispositivo los sonidos producidos por el animal. Esta acción se puede realizar mediante dispositivos conocidos en el estado del arte, como son los micrófonos, o cualquier otro medio para captar sonidos.

Mediante el término “**registrar**”, en la presente invención se refiere al acto de dejar asentado en un medio electrónico el sonido captado. Esto se puede realizar en una memoria del dispositivo o en una memoria de un servidor del sistema.

Mediante el término “**determinar**”, en la presente invención se refiere al acto de señalar o indicar con claridad.

Mediante el término “**identificar**”, en la presente invención se refiere al acto de reconocer si una cosa es la que se busca.

Mediante el término “**detectar**”, en la presente invención se refiere al acto de descubrir la existencia de algo que se busca.

Mediante el término “**monitorear**”, en la presente invención se refiere al acto de observar la evolución de parámetros o actividades.

Mediante el término “**clasificar**”, en la presente invención se refiere al acto de ordenar o disponer por clases.

Mediante el término “**cuantificar**”, en la presente invención se refiere al acto de expresar numéricamente un parámetro.

Mediante el término “**procesar**”, en la presente invención se refiere al acto de someter datos a una serie de operaciones programadas.

Esencialmente, la presente invención resuelve el problema de adquirir, procesar y analizar en tiempo real y de manera robusta y precisa la información necesaria para caracterizar y cuantificar las actividades alimentarias de los animales en rodeos en condiciones de campo o pastoreo, aunque también es aplicable a rodeos confinados en corrales o establos. A diferencia de otros equipos disponibles en el mercado, este dispositivo sólo utiliza los sonidos generados por los animales durante las actividades alimentarias (pastoreo y rumia) para detectar y clasificar cuatro clases de eventos masticatorios (EM) realizados durante la alimentación, y a partir de ellos determinar el tipo de actividad alimentaria realizada, sus parámetros característicos y la cantidad de alimento ingerido.

Durante su alimentación el animal mueve la mandíbula, generando eventos masticatorios (EM) relacionados con las actividades alimentarias. Las características de los sonidos producidos durante la alimentación están determinadas por el alimento ingerido y el EM realizado. De esta manera se puede **establecer una relación unívoca entre los movimientos que produce la mandíbula del animal mientras se alimenta y los sonidos que emite**. Entonces, se pueden estudiar las características temporales de la señal sonora producida por las mandíbulas del animal al alimentarse y relacionarlas con el movimiento mandibular para poder detectar, clasificar y cuantificar los eventos masticatorios. Mientras el animal **pastorea** ingiere forraje y mueve sus mandíbulas para: i) *Arrancar el forraje (arranque)* a partir de aprehenderlo y cortarlo con los dientes incisivos inferiores y el rodete dentario superior; ii) *masticar el forraje (masticación pastoreo)* arrancado para reducir el tamaño de las partículas con los dientes premolares y molares (triturar), insalivar y producir el bolo alimenticio antes de deglutirlo hacia el rumen; y iii) un movimiento que combina las dos acciones anteriores en donde *arranca-mastica*

el forraje (**arranque-masticación**). Por otra parte, mientras el animal **rumia** reprocesa el forraje ya ingerido, mueve sus mandíbulas para solo triturar el forraje contenido en el bolo (**masticación rumia**).

Las masticaciones realizadas durante el pastoreo (**masticación pastoreo**) tienen características muy diferentes a la realizada durante la rumia (**masticación rumia**): i) la masticación de pastoreo producen un sonido con un mayor contenido de energía y amplitud que el de la masticación de rumia, ya que las fibras del forraje aprehendido están enteras y solo tiene el contenido de humedad natural de la planta mientras que el material procesado durante la rumia incorpora agua extra debido a su estancia en el rumen y las fibras ya está parcialmente trituradas; ii) las masticación de rumia es regular, con un periodo aproximado de un segundo entre masticación, debido a que el material a procesar está disponible en la boca y solo depende de la biomecánica del animal mientras que la masticación de pastoreo es irregular porque el forraje a se va juntando a partir de intercalar arranques y arranques-masticaciones con las masticaciones; y iii) la eventos de masticación de rumia ocurren en grupos de 45 a 60 eventos (45 y 74 segundos) seguidos de una pausa de deglución/regurgitación (3 a 5 segundos).

Los eventos de **arranque** y **arranque-masticación** (que se realizan sólo durante la ingesta) tienen características únicas y diferentes de cualquiera de las masticaciones: el sonido producido por un **arranque** se caracteriza por una gran amplitud y una corta duración (asociado al arranque y cortado), mientras que la **arranque-masticación** es la combinación de un **arranque** y una **masticación pastoreo**, resultando en un sonido de gran amplitud y duración. Finalmente, cuando el animal no realiza actividad alimentaria se pueden registrar silencios o ruidos producidos por el ambiente o el animal (respiración, caminar, bebida, lamer, etc.). De modo que, utilizando las características adecuadas de la señal sonora, es posible distinguir entre los diferentes eventos masticatorios relacionados con las actividades alimentarias del animal.

El comportamiento ingestivo se define en términos de una secuencia temporal de EM. Cada una de estas acciones producen sonidos cuyo contenido frecuencial estará dado por las características del material ingerido (contenido de fibra, características de la fibra, contenido de humedad, etc.), mientras que sus características temporales (duración, simetría, energía, etc.) estarán determinados por la cantidad de material y los movimientos de la mandíbula. Entonces, los EMs que ocurren durante la actividad de alimentación pueden identificarse a partir de analizar las características temporales más relevantes de la energía de los sonidos (energía instantánea y acumuladas de la envolvente, duración, periodicidad, secuencia y simetría, entre otros) que produce un animal cuando se alimenta. Un clasificador basado en herramientas de inteligencia computacional (redes neuronales, árboles de reglas, etc.) procesa las características temporales del sonido para identificar el evento y su grado de certidumbre. Los eventos identificados por el dispositivo son: **masticación**

de pastoreo (cuando mastica el alimento recién aprehendido), **masticación de rumia** (cuando mastica el alimento ya procesado en el rumen), **masticación-arranque** (cuando el animal mastica y arranca el alimento recién ingerido en el mismo movimiento mandibular), **arranque** (cuando el animal aprehende y corta el alimento para ingerirlo), ruido (cualquier sonido no reconocido como un evento) y silencio.

Una vez identificados los EMs se los agrupa hasta formar una **ventana de actividad** de duración fija (entre 3 y 15 minutos). Una vez armada la ventana se la clasifica según la actividad que denotan: **pastoreo**, **rumia**, **otra-actividad** y **sin-actividad**. En una ventana temporal fija, el animal realiza una proporción diferente de EMs en cada una de estas actividades. De esta manera, se puede establecer una **relación unívoca** entre los porcentajes de movimientos de la mandíbula reconocidos en el sonido con las actividades alimentarias durante su alimentación. La información de los eventos (etiqueta, grado de certidumbre y parámetros) es utilizada por un modelo estadístico basados en herramientas de inteligencia computacional (perceptron multicapas, árboles de reglas, etc) para identificar las actividades realizadas por el animal. Por ejemplo, si se detecta una secuencia de **masticaciones de rumia**, con algunas pocas **masticaciones de pastoreo** (o **arranques**) aisladas, es altamente probable que el animal estuviera **rumiando**. Por el contrario, si se detecta una secuencia de **arranques** y **masticaciones-arranques**, con algunas **masticaciones de rumia** es altamente probable que el animal esté **pastoreando**.

En una **manera alternativa** de implementar la presente invención, el comportamiento ingestivo puede definirse, a partir de un conjunto de características temporales, estadísticas y espectrales de la envolvente del sonido producido por el animal durante la alimentación. Estas características no contienen **información explícita** sobre el tipo de evento masticatorio realizado por el animal en cada instante de tiempo. Las características que utilizamos para identificar los comportamientos alimentarios son:

- i. Parámetros estadísticos de las amplitudes máximas, las duraciones y periodos entre los eventos detectados en un segmento de señal de duración fijo. Los parámetros estadísticos utilizados para estos datos son la media, la varianza, el sesgo y la curtosis
- ii. Parámetros estadísticos de la variabilidad de las amplitudes máximas (shimmer) y los periodos entre eventos detectados (jitter) en un segmento de señal de duración fijo. Los parámetros estadísticos utilizados para estos datos son la media y la varianza.

Estos parámetros estadísticos, y cualquier otra información útil que no involucre la clasificación de los eventos masticatorios (espectros de las señales, parámetros obtenidos a partir de herramientas de sistemas dinámicos (mapas de poincaré, etc.) y procesamiento de información (clustering, aprendizaje no supervisado, etc.)), es utilizada por un modelo estadístico basados en herramientas de inteligencia

computacional (perceptron multicapas, árboles de reglas, etc) para clasificar las actividades realizadas por el animal. Por ejemplo, si se detecta un tramo de sonido donde los eventos detectados presentan

- amplitudes máximas con media y varianza bajas, sesgo negativo y curtosis positiva;
- duraciones con media grande, una varianza pequeña, sesgo positivo y curtosis positiva;
- periodos intereventos con varianza y sesgo pequeños y curtosis positiva; y
- variabilidad de las amplitudes máximas y periodos con medias y varianza muy pequeñas (casi nulas);

es altamente probable que el animal estuviera rumiando. Por el contrario, si se detecta un tramo de sonido donde los eventos detectados presentan

- amplitudes máximas con media y varianza altas, sesgo negativo y curtosis negativa;
- duraciones con media, varianza y sesgo grandes y curtosis pequeña (próxima a cero);
- periodos intereventos con varianza grande, sesgo negativo y curtosis negativa;
- y
- variabilidad de las amplitudes máximas con varianza grande y periodos con medias pequeña (casi nulas);

es altamente probable que el animal esté pastoreando.

Una vez identificadas las ventanas de actividades y sus parámetros, se procede a identificar las actividades realizadas por el animal en un periodo largo (varias horas). Antes de integrar la información para clasificar la actividad alimentaria en periodos largos de tiempo se analiza cada ventana en relación con su secuencia temporal. Es decir, se analiza la categoría asignada en relación con las ventanas anteriores y posteriores para remover resultados anómalos (filtrado de mediana o suavizado). Se analiza la secuencia de ventanas y cuando se encuentra un paquete que se aleja del comportamiento medio se cambia etiqueta. Por ejemplo, si tenemos una secuencia de paquetes de rumia con un único paquete de sin-actividad (o pastoreo) en el medio es probable que se haya cometido un error de clasificación y se procede a cambiar la etiqueta de la ventana a rumia. Sin embargo, si se produce un cambio de actividad, la secuencia de ventanas cambiará de manera consistente y no se corregirá el comportamiento. Por ejemplo, si tenemos una secuencia de ventanas de rumia seguido de una secuencia de paquetes de pastoreo.

La información corregida de las ventanas de eventos es utilizada por otro modelo estadístico para evaluar la secuencia de paquetes identificados. Las actividades identificadas son: **pastoreo**, **rumia**, **otra-actividad** y **sin-actividad**. Finalmente, una vez identificadas las actividades alimentarias realizadas por el animal y dados sus parámetros característicos (cantidad total de eventos, cantidad de eventos por clase,

energía acumulada y periodo, entre otros) se calcula la cantidad de materia seca ingerida por el animal a través de un modelo computacional que utiliza los parámetros de los paquetes para estimar la cantidad de materia seca ingerida por el animal.

En una preferida de implementación, la presente invención provee de un dispositivo que comprende sensores para adquirir los sonidos producidos por el animal. Un sistema embebido basado en microprocesadores, dispositivos lógicos programables (FPGA), circuito integrado de aplicación específica (ASIC) o combinación de ellos, adquiere, procesa y analiza los sonidos producidos por el animal. El sonido es procesado para ser acondicionado y extraer la información indicativa de cada uno de los aspectos de la conducta alimentaria del animal considerado. El proceso de detección y clasificación puede realizarse en su totalidad en el sistema embebido y la información resultante puede ser enviada a una computadora periódicamente para su registro. El dispositivo puede estar contenido parcial o totalmente dentro de una caja estanca que se colocará preferentemente en la cabeza, el cuello o la oreja del animal. Estos lugares permiten adquirir los sonidos producidos por el animal durante su alimentación a través de un sensor en la cabeza del animal (micrófono) y luego analizarlos para determinar qué actividad realiza. El sensor de sonido preferido son los micrófonos direccionales (analógicos o digitales), los cuales son capaces de adquirir sonidos muy débiles en la dirección de preferencia (frente del micrófono). Como la señal es débil, la misma debe ser acondicionada (amplificada y filtrada) a través de circuitos electrónicos específicos. Se pueden utilizar dos configuraciones posibles: i) un solo micrófono direccional con aislación acústica, de modo que la señal registrada tenga una buena relación señal/ruido, o ii) dos (ó más) micrófonos direccionales para poder realizar una cancelación activa del ruido ambiente en condiciones más desfavorables. En ambos casos, la señal de sonido es procesada previamente por un algoritmo digital de limpieza que atenúa el efecto del ruido ambiente y otras señales no deseadas.

La Figura 1 presenta el diagrama de bloques de un dispositivo de adquisición, procesamiento y análisis de sonidos, el cual incluye cinco bloques funcionales:

1. **Adquisición y procesamiento de señales** - este bloque es el encargado de adquirir los sonidos producidos por el animal. Las señales registradas son procesadas y acondicionadas, con circuitos analógicos y/o con algoritmos digitales de procesamiento, para remover las perturbaciones y ruidos que se hayan registrado junto a los sonidos y generar señales indicativas de actividad alimentaria que permitan una gestión eficiente de la energía a partir de que el dispositivo solo opere cuando el animal se está alimentando. La ganancia del sensor, en conjunto con la de los circuitos de acondicionamiento es estimada para mejorar el desempeño de los algoritmos de detección y clasificación de eventos.
2. **Detección de eventos y extracción de características** - este bloque es el

encargado de analizar los sonidos registrados para detectar en tiempo real los eventos masticatorios (EM) que realiza el animal durante las actividades alimentarias. En esta tarea también se extrae la información relevante de los EMs para que el bloque de clasificación de eventos clasifique el tipo de EM detectado y calcule sus parámetros fundamentales. Esta tarea también puede extraer la información relevante del sonido para que el bloque de análisis de actividades determine y cuantifique la conducta alimentaria del animal a lo largo del tiempo sin utilizar información de los EMs .

3. **Clasificación de eventos** - este bloque es el encargado de analizar la información de los EM detectados para clasificar en tiempo real los eventos masticatorios (EM) que realiza el animal durante las actividades alimentarias. En esta tarea también se extrae la información relevante de los EMs para que el bloque de análisis de actividades determine y cuantifique la conducta alimentaria del animal y el consumo de alimento a lo largo del tiempo.
4. **Análisis de actividades** - este bloque es el encargado de analizar paquetes de EMs clasificados para detectar, identificar y cuantificar en tiempo real las actividades alimentarias (pastoreo o rumia) realizadas por el animal. Esta tarea también extrae la información relevante para analizar y cuantificar la conducta alimentaria del animal a lo largo del tiempo. Este bloque también puede identificar y cuantificar en tiempo real las actividades alimentarias (pastoreo o rumia) realizadas por el animal a partir de la información extraída del sonido, sin clasificar EMs.
5. **Almacenamiento y comunicaciones** - este bloque es el encargado de organizar, almacenar y transmitir la información generada por el dispositivo. Se encarga de estructurar la información en bloques autocontenidos y seguros que provean información adicional de utilidad (por ejemplo lugar y hora donde se adquirieron los datos, entre otros). Además, es responsable de intercambiar información con el sistema a través un sistema de comunicación inalámbrica de baja potencia. Esto permite comunicar al sistema de gestión la información del comportamiento alimentario del animal en tiempo real.
6. **Control de operación** - este bloque es el encargado de administrar la operación del dispositivo, optimizando el uso de la energía disponible para maximizar el tiempo de operación. Esta tarea se realiza deshabilitando funciones del sistema, de acuerdo a su prioridad, y manejando la carga de la batería. Este bloque utiliza la señal indicativa de actividad generada por el bloque de adquisición y procesamiento para solo ejecutar los algoritmos cuando el animal está realizando alguna actividad alimentaria.

Estos bloques se describirán en detalle a continuación.

La Figura 2 ilustra la realización del bloque de adquisición y procesamiento de señales del dispositivo. La señal sonora adquirida por el micrófono del animal es

acondicionada por un preamplificador y su ancho de banda limitado por un filtro pasabanda. A partir de este punto la señal sigue dos caminos diferentes, uno para detectar los eventos y el otro para calcular parámetros internos del algoritmo, en los cuales recibe diferentes acondicionamientos antes de su procesamiento. En ambos caminos, la señal es amplificada para mejorar la relación señal-ruido y el rango dinámico del conversor analógico/digital, que digitaliza la señal analógica. La señal utilizada para detectar los eventos es acondicionada con un amplificador con control automático de ganancia (AGC) para obtener una señal con buena relación señal/ruido sin distorsiones. Esto permite mejorar la tasa de reconocimiento y clasificación, en relación a que se procese la señal con un amplificador con ganancia fija. Sin embargo, esto dificulta el cálculo de los parámetros relacionados con actividad alimentaria y afecta a la sensibilidad del algoritmo, que están relacionados con la energía de la señal, ya que la señal original es modificada por una ganancia variante en el tiempo y desconocida. Para calcular los cambios de ganancia introducidos por el AGC se utiliza una señal procesada por un amplificador con una ganancia conocida y así ajustar los parámetros de los algoritmos de detección y clasificación de eventos.

Para captar los sonidos producidos por el animal (micrófono de animal) se utiliza un micrófono direccional (analógico o digital) que puede estar sujeto contra la frente del animal en U1 o cerca de la oreja en U3 (Figura 3). Estos son los mejores lugares para adquirir el sonido debido a la conducción ósea de la cabeza del animal. Un segundo micrófono (micrófono de ambiente) ubicado junto a la unidad de procesamiento U2 adquiere los sonidos ambientales y es procesado por el mismo sistema que la señal del micrófono del animal, de modo que ambas señales puedan ser utilizadas por el bloque de **cancelación de ruidos** (Figura 2). Este bloque remueve los ruidos y perturbaciones medibles para mejorar la relación señal-ruido de la señal adquirida por el micrófono del animal. La señal resultante es procesada por un filtro pasabajo de frecuencia de corte muy baja (del orden del periodo de un bloque de rumia) para generar una señal indicativa de la actividad realizada por el animal. Si el animal rumia o pastorea, la salida de este filtro tendrá un voltaje mayor o igual a la mitad del voltaje de alimentación, si el animal no realiza actividad tendrá una salida nula. En caso de ruidos, la señal no será nula pero su valor será muy bajo y próximo a cero. Esta señal es utilizada por el bloque de control de operaciones para gestionar las tareas del dispositivo, ejecutando el algoritmo cuando el animal realiza actividades alimentarias (Figura 12).

El bloque de **adquisición y procesamiento** provee información al bloque **detección de eventos y extracción de características** a través de una memoria FIFO (buffer) que permite la ejecución independiente de las tareas, relajando los requerimientos de coordinación ejecución.

En el proceso de alimentación de un animal se puede establecer una primera relación entre los movimientos de la mandíbula y los sonidos que producen, y luego se puede

establecer una segunda relación entre la secuencia de movimientos de la mandíbula y las actividades alimentarias realizadas durante ese período. Durante su alimentación (pastoreo y rumia), el animal mueve la mandíbula de forma continua y se pueden diferenciar dos eventos fundamentales:

1. **Arrancar** - cuando el forraje es aprehendido y cortado, y
2. **Masticar** - cuando el forraje es triturado para reducir el tamaño de las partículas.

El evento de masticar (características mecánicas del proceso de alimentación) se efectúa en las dos actividades alimentarias realizadas por el animal (rumia y pastoreo), por lo que tendremos dos clases de masticaciones: una relacionada con el pastoreo (masticación de pastoreo) y otra relacionada con la rumia (masticación de rumia). Por lo tanto, las actividades alimentarias realizadas por un animal pueden definirse en términos de los siguientes eventos: masticación de pastoreo, arranque-masticación (evento que combina masticación con arranque), arranque, masticación de rumia, silencio y ruido.

El pastoreo es una secuencia temporal de arranques, masticaciones de pastoreo y arranques-masticaciones sin un ritmo u orden establecido (figura 4.a), con pausas cortas de duración variable intercaladas para tragar el forraje procesado (bolo) y/o buscar/seleccionar nuevo forraje. Esta secuencia temporal contiene información que está relacionada con la ingesta (tiempo de pastoreo, cantidad de materia seca ingerida, etc.). Durante la rumia el animal mueve la mandíbula de forma regular solo para masticar (masticación de rumia), con pausas regulares para tragar y regurgitar el alimento. De este modo la rumia se puede definir en términos de la secuencia temporal periódica de masticaciones de rumia, seguida de pausas regulares de degluciones/regurgitaciones de duración fija (Figura 4.b). Esta secuencia temporal contiene información que está relacionada con la rumia (como tiempo de rumia, por ejemplo).

La Figura 4 presenta la salida de un micrófono durante 200 s en pastoreo (Figura 4.a) y rumia (Figura 4.b). Como se ve en la Figura 4.a, los sonidos que se producen durante el pastoreo se caracterizan por ser de corta duración, sin un ritmo muy definido, pero alrededor de 1 Hz. Por otro lado, el sonido producido durante la rumia se caracteriza por su morfología temporal: menor energía relativa y una periodicidad más marcada a al nivel de las masticaciones durante la rumia con una frecuencia próxima a 1 Hz, y al nivel de la regurgitaciones que se repiten aproximadamente cada un minuto (Figura 4.b).

Un primer paso para analizar la actividad alimentaria es detectar, clasificar y caracterizar los eventos masticatorios (**relación movimiento-sonido**): arranque, arranque-masticación, masticación de pastoreo y masticación de rumia. Cada una de estas acciones producen sonidos cuyo contenido frecuencial y energía estarán dados por el material procesado (cantidad, contenido de fibra, humedad, etc.), mientras que la duración y morfología del sonido estarán determinados por el movimiento de la mandíbula. Entonces, el sonido producido por un **arranque** se caracteriza por una

gran amplitud y una corta duración, característico de un rasguído (Figura 5.c). La energía de la señal es alta, así como los cambios acumulados durante el evento. Otra característica a destacar es la simetría de la señal, definida como el cociente de las áreas a derecha e izquierda del punto donde la señal alcanza su valor máximo. En cambio, el sonido producido por una **masticación de pastoreo** se caracteriza por una amplitud más pequeña y una duración más larga, característico de un chirrido (Figura 5.b). La energía de la señal es más baja que la del evento de arranque, así como los cambios acumulados durante el evento. Una característica a destacar es la simetría de la señal. El movimiento combinado **arranque-masticación** es la combinación de ambos (Figura 5.d). La señal de este evento es la combinación de los dos anteriores, resultando en una amplitud y una duración mayores que los otros eventos. La energía de la señal es la más alta de todos los eventos, así como los cambios acumulados. Este evento se destaca por la falta de simetría. En cambio, el sonido producido por una **masticación de rumia** se caracteriza por su reducida amplitud (la más pequeña de todos los eventos) y una duración casi tan larga como la de una **arranque-masticación** (Figura 5.a). La energía de la señal es la más baja de todas, debido al alto contenido de humedad del material. Una característica a destacar es la simetría de la señal resultante. Finalmente, se considera ruido a un sonido que no satisface ninguna de las características antes descritas y un silencio es la ausencia de sonidos durante el periodo de un evento promedio.

De manera que los eventos masticatorios que ocurren durante la actividad de alimentación pueden caracterizarse a partir del análisis de las características temporales más relevantes de los sonidos que produce un animal cuando se alimenta (forma de la energía instantánea, energía total, velocidad de cambio, duración, periodicidad, secuencia y simetría, entre otros). Entonces, un clasificador analiza las características temporales de la señal audio para clasificar los siguientes eventos: arranque, masticación-arranque, masticación pastoreo, masticación rumia, ruido y silencio.

La Figura 6 muestra la realización del bloque de **detección de eventos y extracción de características** junto con los clasificadores de eventos y actividades. Este bloque está conformado por cuatro sub-bloques: i) un detector de eventos, ii) un extractor de características y iii) un ajuste de los parámetros y umbrales del algoritmo. El **detector de eventos** determina la posible ocurrencia, o no, de un EM a partir de analizar la envolvente del sonido decimada. La presencia de un pico denota la posibilidad de EM. Los picos se detectan a partir de la coincidencia de i) un cambio en el signo de la derivada de la envolvente y que ii) la envolvente sea mayor que un umbral $U(k)$ (Figura 7.b). El algoritmo que genera el umbral considera las características anatómicas y de comportamiento del animal, así como los cambios en las condiciones de las actividades alimentarias. Cada vez que se detecta un evento, el bloque genera una señal indicativa y una estampa del tiempo en que se detectó el potencial evento. La señal indicativa avisa de la presencia de un potencial evento a los demás bloques del sistema, de modo que las características calculadas por el

extractor de características son analizadas o almacenadas, dependiendo de las características calculadas.

El **extractor de características** procesa las señales acústicas provenientes del bloque de cancelación de ruidos. En una primera etapa, se calcula la envolvente de la señal y la energía de la señal por frame, luego ambas señales son decimadas (sub-muestreadas a fin de reducir la carga computacional). Estas señales son almacenadas en sendos buffer circulares. A partir de esta información se calculan cinco características temporales de la señal de energía instantánea x_E : la energía acumulada durante el evento, el valor absoluto de los cambios, duración, forma (signo de la derivada) y simetría (Figura 5).

El **clasificador de eventos** es el encargado de determinar qué tipo de evento (arranque, arranque-masticación, masticación de pastoreo, masticación de rumia, ruido, silencio) ocurrió y cuál es el grado de certeza de dicha clasificación. Para realizar esta tarea se puede utilizar herramientas de inteligencia computacional (redes neuronales multicapas, árboles de decisión, entre otros) entrenados para tal fin. En esta implementación se utiliza un perceptrón multicapa, que tiene dos salidas: i) la etiqueta del evento y ii) el grado de certidumbre de la clasificación del evento. El **clasificador de eventos** solo actúa si las características calculadas cumplen con los criterios de un evento (validación). Es decir, las características son validadas antes de su uso por parte del clasificador (Figura 6). Si el evento es clasificado, el bloque de **cálculo de parámetros de eventos** utiliza la información de la etiqueta del evento en combinación con la señal digitalizada para calcular la energía de los eventos. Los parámetros característicos restantes (como la duración, por ejemplo) son calculados a partir de las etiquetas y características generadas por el sub-bloque de extracción de características.

Si el bloque calcula características temporales de los eventos detectados, el reconocimiento de actividades alimentarias no necesita clasificar eventos. Entonces las características generadas por el extractor y las marcas temporales asociadas son almacenadas en una memoria temporal para su procesamiento por el reconocedor de actividades.

Finalmente, el bloque de **ajuste de los parámetros del algoritmo** adapta los parámetros internos del algoritmo a las condiciones operativas y los hábitos de los animales para maximizar la detección y reconocimiento de los eventos. Utilizando la información de las características calculadas y la etiqueta para ajustar los parámetros y umbrales de acuerdo con reglas que mejoran el desempeño del sistema.

El módulo de **detección de eventos y extracción de características** se implementa a través de un algoritmo que lee los datos de la envolvente del audio $x_e(k)$ y calcula su derivada, luego si se produce un cambio en el signo de esta última (condición necesaria para un pico) la envolvente es comparada con umbral adaptativo $U(k)$ para

detectar si se produjo un pico en $x_e(k)$ (Figura 7.b). Si transcurre un periodo T_{MAX} sin que se detecte un pico se asume que ocurrió un silencio.

El procedimiento para detectar y clasificar eventos masticatorios es el siguiente:

1. A cada muestra del bloque de sonido $x(k)$ que se lee del bloque de adquisición y procesamiento se le aplican los siguientes procesos:
 - a. Calcular la envolvente de la señal de audio;
 - b. Submuestrear la señal envolvente a 150 Hz para generar la señal $x_e(k)$;
 - c. Almacenar $x_e(k)$ en el buffer circular B_e de 150 muestras;
 - d. Calcular la energía de $x(k)$ en frames a una tasa de 150 Hz para generar la señal $x_E(k)$ (Figura 7.a);
 - e. Almacenar $x_E(k)$ en el buffer circular B_E de 150 muestras;
 - f. Calcula el signo del cambio $S(k)=sign(x_e(k)-x_e(k-1))$.
2. Actualizar el umbral adaptativo $U(k)$, si ya transcurrió el periodo refractario T_{REF} y no se alcanzó el periodo máximo T_{MAX} (Figura 8.a);
3. Si transcurrió el periodo máximo de detección T_{MAX} sin que se haya detectado un evento, se etiqueta ese periodo como silencio (“**SILENCIO**”) y $U = U_0$;
4. Si ha finalizado el periodo refractario T_{REF} , $x_e(k) > U(k)$, y $sign(|S(k)-S(k-1)|) > 0$ entonces ha ocurrido un evento y se procede a extraer las características:
 - a. Comparar los datos de B_E con el umbral U_F para delimitar el inicio (n_{INICIO}) y el fin (n_{FIN}) del evento candidato. n_{INICIO} y n_{FIN} corresponden a los valores mínimo y máximo de la secuencia de n consecutivos tal que $x_E(n) > U_F$;
 - b. Segmentar los datos almacenados en B_e y B_E utilizando n_{INICIO} y n_{FIN} para determinar el segmento de datos útil que se va a utilizar para calcular las características;
 - c. Extraer las características de clasificación;
 - d. Verificar que las características calculadas están en los rangos adecuados.
 - i. Si las características no están los rangos predefinidos se descartan e ir al Paso 5 (Adaptación de los umbrales y parámetros del algoritmo);
 - ii. Si las características están dentro de los rangos predefinidos se aplica el clasificador entrenado (clasificador de eventos) a las características calculadas y se determina el tipo de evento.
5. Una vez que se ha clasificado el evento, se actualizan los parámetros de los algoritmos de umbral variable de la envolvente ($U(k)$) y la energía instantánea (U_F).
6. Cuando se procesaron N muestras se construye y almacena la **ventana de eventos**.

Una vez identificado cada evento, el módulo de detección y extracción de características procede a ajustar sus parámetros. Se extraen valores internos de los búferes B_e y B_E para actualizar los parámetros asociados con los **generadores de**

umbral de energía (U_F) y detección ($U(k)$). Internamente, los búferes B_E y B_e están divididos en tres segmentos: el primer y tercer segmento se utilizan para estimar los niveles de ruido, mientras que los segmentos medios son utilizados para la detección y extracción de características de los eventos.

Una vez que el evento ha sido clasificado, se adaptan los parámetros del umbral de detección $U(k)$ en tres pasos sucesivos:

1. Se calcula el tiempo transcurrido entre el último evento clasificado y el actual, y se calculan los valores máximo y mínimo de la envolvente;
2. Los datos son utilizados para calcular los parámetros del generador de umbral $U(k)$:
 - **Periodo sin respuesta** (T_{REF}) - período de tiempo después de clasificar el final de un evento válido en el que el algoritmo ya no busca un nuevo evento. Se calcula para cada evento clasificado como el 25% del tiempo promedio de los últimos cuatro eventos;
 - **Periodo máximo** (T_{MAX}) - periodo de tiempo máximo que puede durar un evento masticatorio. Se calcula para cada evento clasificado como el tiempo promedio de los últimos cuatro eventos;
 - **Umbral máximo** (U_{MAX}) - máximo valor del umbral esperado después de que expire T_U en la búsqueda de un nuevo evento. Se calcula como el 50% del promedio del máximo de la envolvente de los últimos cuatro eventos;
 - **Umbral mínimo** (U_{MIN}) - mínimo valor del umbral esperado en la búsqueda de un nuevo evento. Se calcula como:

$$U_{MIN} = \min \left\{ \left(U_{MIN,avg} + 0.25U_{MAX} - U_{MIN,avg} \right), \left(4U_{MIN,avg} \right) \right\}$$

donde $U_{MIN,avg}$ es el valor promedio del mínimo de la envolvente de los últimos cuatro eventos;

- **Paso del umbral** (ΔU) - es el paso de decaimiento del umbral $U(k)$ desde U_{MAX} hasta U_{MIN} una vez que T_U expira, y sirve para aumentar la sensibilidad a la detección de nuevos eventos. Se calcula como $(U_{MAX} - U_{MIN}) / T_M$;
- **Periodo de decaimiento** (T_D) - es posible que luego de un largo período de inactividad alimentaria (ni rumia, ni pastoreo), cuando el animal reinicie nuevamente con la actividad continua (rumia o pastoreo), el nivel del umbral calculado previamente sea muy alto, no detectando los nuevos eventos. Esta situación no deseada podría ocurrir, por ejemplo, si el animal está pastoreando (caso en que la envolvente tiene gran valor) y después de un período de inactividad el animal comience a rumiar (caso en que la envolvente tiene un valor pequeño). Por lo tanto, T_D es el período de tiempo que debe transcurrir después de detectar el último evento válido para que el umbral $U(k)$ disminuya

su valor para mejorar su sensibilidad. Se calcula como 8 veces el tiempo promedio entre dos eventos clasificados, de los últimos cuatro eventos.

- **Periodo de aumento** (T_i) - es posible que durante un período de silencio el nivel de umbral sea muy sensible a la detección de falsos eventos debido a ruidos o tendencias. Por lo tanto, T_i es el período de tiempo que debe transcurrir después de detectar el último evento válido para que el umbral $U(k)$ aumente su valor si ocurre una detección de un falso evento. Se calcula como 4 veces el tiempo promedio entre dos eventos clasificados, de los últimos cuatro eventos.

3. El nuevo valor del umbral de detección $U(k)$ se actualiza de la siguiente manera:

$$U(k) = U_{MAX} \quad \forall k < T_{REF} ;$$

$$U(k) = U(k - 1) - \Delta U, \quad \forall T_{REF} > k < T_{MAX};$$

$$U(k) = U \frac{(k-1)}{2} \quad \text{si } k = k^* + T_D$$

$$U(k) = 1.5 * U(k - 1) \quad \text{si } k = k^* + T_i \text{ y ocurre una falsa detección;}$$

$$U(k) = U(k - 1) \quad \text{ninguno de los anteriores.}$$

dónde k^* es la posición k donde se detectó el último evento. Luego de que un evento es clasificado, se estiman la potencia de ruido y la potencia de la señal. La potencia del ruido se calcula como el valor medio de las muestras del primer y tercer segmento de cada buffer. Finalmente, el valor del umbral de energía U_F se actualiza

$$U_F = \rho T_{NFEL},$$

donde ρ es un parámetro de sensibilidad que se modifica en valores discretos de acuerdo con el valor promedio del nivel del SNR estimado para los últimos cuatro eventos clasificados, y N_{FEL} es el promedio de la potencia de ruido de fondo para los últimos cuatro eventos EM clasificados.

El siguiente paso para analizar la actividad alimentaria es caracterizar grupos de eventos masticatorios (**relación movimientos-actividad**). Para ello, se organizan los eventos individuales identificados en ventanas temporales de duración fija (entre 3 y 15 minutos). Las ventanas de eventos se clasifican según la actividad que denotan: pastoreo, rumia, otra-actividad y sin-actividad. La información de los eventos (etiqueta y grado de certidumbre) es utilizada por un modelo estadístico basado en herramientas de inteligencia computacional para evaluar la secuencia de eventos identificados y asignarle una de las cuatro etiquetas propuestas (Figura 9). Por ejemplo, si se detecta una secuencia de masticaciones de rumia, con algunos arranques y masticaciones de pastoreo aisladas es altamente probable que el animal esté rumiando. Por el contrario, si se detecta una secuencia de arranques, masticaciones-arranques y masticaciones de pastoreo, con algunas masticaciones de

rumia y ruidos es altamente probable que el animal esté pastoreando. Una vez determinado el tipo de actividad de la ventana, el bloque **cálculo de parámetros** calcula los parámetros característicos de la ventana de eventos (duración del paquete, cantidad de masticaciones, arranques y masticación-arranques, duración promedio de los eventos, energía de las arranques, entre otros). Finalmente, una vez identificadas las actividades alimentarias realizadas por el animal y sus parámetros característicos, el **modelo de consumo** estima la cantidad de materia seca ingerida por el animal (Figura 9).

La longitud de la ventana **N** es un parámetro importante ya que determina la exactitud y precisión de las estimaciones. Si la ventana es muy corta (menos de 3 minutos) el sistema va a ser capaz de realizar un seguimiento de los cambios de actividad pero con cambios significativos en los resultados obtenidos en ventanas sucesivas debido a los pobres estadísticos calculados. Por el contrario, si la ventana es larga (más de 15 minutos) las actividades entre ventanas van a ser consistentes (sólo habrá variaciones cuando cambie la actividad), debido a los buenos estadísticos calculados, pero se introducirán errores cuando cambien las actividades, introduciendo sobre estimaciones en la actividad alimentaria dominante.

La Figura 10 muestra el diagrama de flujo de un algoritmo para clasificar las **ventanas de eventos** a partir de la información provista por el módulo de detección y clasificación de eventos. En un primer paso se calculan los porcentajes de masticaciones de rumia (P_{CR}) y pastoreo (P_{CP}), arranques (P_B) y arranques-masticaciones (P_{CB}) de cada ventana con respecto al número total de eventos (N_E) almacenados en la ventana de longitud N. Luego, en función de un conjunto de reglas heurísticas, o un modelo computacional basado en aprendizaje maquina (redes neuronales, árboles de decisiones, etc.), se clasifican las ventanas de actividades (pastoreo, rumia, silencio y ruido) y se acumulan los parámetros de los paquetes para caracterizar las actividades alimentarias. El procedimiento para clasificar los paquetes de actividades es el siguiente:

1. Inicializar $N_E=0$; $Cont_{CR}=0$; $Cont_{CP}=0$; $Cont_B=0$; $Cont_{CB}=0$; $Cont_S=0$; $Cont_N=0$;
 $P_{CR}=0$; $P_{CP}=0$; $P_B=0$; $P_{CB}=0$; $P_S=0$; $P_N=0$; $m=0$;
2. $m=m+1$; Leer datos del buffer;
3. Incrementar el contador de eventos totales (N_E);
4. Si la etiqueta del evento $L(m)$ es una “**MASTICACION DE RUMIA**” entonces:
 - a. Incremente el contador de eventos de masticaciones ($Cont_{CR}$);
 - b. Acumule los parámetros del evento en el buffer de resultados;
5. Si la etiqueta del evento $L(m)$ es una “**MASTICACION DE PASTOREO**” entonces:
 - a. Incremente el contador de eventos de masticaciones ($Cont_{CP}$);
 - b. Acumule los parámetros del evento en el buffer de resultados;

5. Si la etiqueta del evento $L(m)$ es una “**ARRANQUE**” entonces:
 - a. Incremente el contador de eventos de arranques ($Cont_B$);
 - b. Acumule los parámetros del evento en el buffer de resultados;
6. Si la etiqueta del evento $L(m)$ es una “**ARRANQUE-MASTICACION**” entonces:
 - a. Incremente el contador de eventos de masticación ($Cont_{CB}$);
 - b. Acumule los parámetros del evento en el buffer de resultados;
7. Si la etiqueta del evento $L(m)$ es una “**SILENCIO**” entonces:
 - a. Incremente el contador de eventos de silencios ($Cont_S$);
 - b. Acumule los parámetros del evento en el buffer de resultados;
8. Si la etiqueta del evento $L(m)$ es una “**RUIDO**” entonces:
 - a. Incremente el contador de eventos de silencios ($Cont_N$);
 - b. Acumule los parámetros del evento en el buffer de resultados;
9. Si $m < N$ entonces vaya al Paso 2;
10. Calcular los parámetros estadísticos del paquete de eventos.
11. Si el clasificador de paquete de eventos utiliza un método de aprendizaje maquina:
 - a. Ejecutar el modelo del clasificador
 - b. Asignar la etiqueta de salida del clasificador a la ventana;
 - c. Calcular los parámetros del paquete;
12. Si el clasificador de paquete de evento utiliza reglas heurísticas:
 - a. Si el porcentaje de masticaciones de rumia (P_{CR}) es mayor que el 80% y el número total de eventos (N_E) es mayor que un número mínimo (N_{MIN}) de eventos entonces:
 - i. Asignar la etiqueta “**RUMIA**” a la ventana;
 - ii. Calcular los parámetros del paquete;
 - b. Si el porcentaje de masticaciones de pastoreo (P_{CP}) es mayor que el 20%, la suma de los porcentajes de arranques (P_B) y arranque-masticaciones (P_{CB}) es mayor que el 70% y el número total de eventos (N_E) es mayor que un número mínimo (N_{MIN}) de eventos entonces:
 - i. Asignar la etiqueta “**PASTOREO**” a la ventana;
 - ii. Calcular los parámetros del paquete;
 - c. Si el porcentajes de silencios (P_S) es mayor que el 70% y el número total de eventos (N_E) es mayor que un número mínimo (N_{MIN}) de eventos entonces:
 - i. Asignar la etiqueta “**NO ACTIVIDAD**” a la ventana;
 - ii. Calcular los parámetros del paquete;

d. Si no cumple con ninguna de las condiciones anteriores entonces:

i. Asignar la etiqueta “**OTRA ACTIVIDAD**” a la ventana;

13. Construir y almacenar el paquete de datos correspondiente a la ventana de eventos.

14. Ir al Paso 1

Una vez determinada la actividad de cada paquete se calculan sus parámetros característicos (cantidad de eventos, cantidad de masticaciones de pastoreo, tiempo total de masticación de pastoreo, amplitud máxima promedio de masticaciones de pastoreo, energía promedio de masticaciones de pastoreo, cantidad de arranques, tiempo total de arranque, amplitud máxima promedio de arranques, energía promedio de arranques, cantidad de masticaciones-arranques, cantidad de masticaciones de rumia, tiempo total de masticación de rumia, amplitud máxima promedio de masticaciones de rumia, entre otras). Luego se estima la DMI de cada animal (Figura 9). Las variables necesarias para estimarla son:

- i. **Intensidad de masticación** (V_C) - es la energía sonora asociada a una masticación pastoreo detectada en la señal registrada. Se calcula como una de las características de clasificación (Figura 5) e incluye a la parte masticación de los eventos masticación-arranque;
- ii. **Energía total de masticación** (E_C) - es la energía acumulada de todos los eventos de masticación pastoreo y la parte masticación de los eventos masticación-arranque;
- iii. **Energía de masticación por masticación** (E_{CC}) - es la energía media de los eventos de masticación de pastoreo. Se la calcula como el cociente entre E_C y la cantidad de masticaciones de pastoreo $Cont_{CP}$ más la cantidad de masticaciones-arranque $Cont_{CB}$ (Figura 10);
- iv. **Energía de masticación por unidad de tiempo de comida** (E_{TC}) - es la energía media masticación por unidad de tiempo. Se la calcula como el cociente entre E_{CC} y el tiempo total de masticación pastoreo (T_C), el cual se calcula acumulando la duración de todas las masticación pastoreo en el paquete;
- v. **Intensidad de arranque** (V_B) - es la energía sonora asociada a un arranque detectado en la señal registrada. Se calcula como una de las características de clasificación (Figura 5) e incluye a la parte arranque de los eventos masticación-arranque;
- vi. **Energía de densidad de flujo de arranque por arranque** (E_B) - es la energía acumulada de todos los eventos de arranque y la parte arranque de los eventos masticación-arranque;
- vii. **Energía total de densidad de flujo de masticación por arranque** (E_{BB}) - es la energía media de los eventos de arranque. Se la calcula como el cociente

entre E_B y la cantidad de masticaciones de pastoreo $Cont_B$ más la cantidad de masticaciones-arranque $Cont_{CB}$ (Figura 10);

Por último, la cantidad de materia seca consumida (DMI) es estimada con un modelo de regresión no lineal multivariada, que se puede implementar usando herramientas de inteligencia computacional como perceptrones multicapas o máquinas de aprendizaje extremo (Figura 9).

La Figura 11 muestra una realización alternativa del bloque de análisis de actividades alimentarias en la cual la actividad alimentaria es caracterizada (relación movimientos-actividad) a partir de un conjunto de características temporales, estadísticas y espectrales de la envolvente del sonido producido por el animal durante la alimentación. Estas características no contienen información explícita del tipo de evento masticatorio realizado por el animal en cada instante de tiempo, pero permiten caracterizar las actividades realizadas por el animal. Para ello, se organizan las características extraídas en ventanas temporales de duración fija (entre 3 y 15 minutos). Las ventanas son procesadas para extraer los estadísticos que se utilizarán para clasificar la actividad que realiza el animal: pastoreo, rumia, otra-actividad y sin-actividad. Los parámetros estadísticos (media, la varianza, el sesgo y la curtosis) de las amplitudes máximas, las duraciones y los periodos entre los eventos detectados durante la ventana temporal, así como su variabilidad, son utilizados por un modelo estadístico basado en herramientas de inteligencia computacional para asignarle una de las cuatro etiquetas propuestas (Figura 11). Por ejemplo, si se identifica un tramo de sonido donde los eventos detectados presentan:

- i. amplitudes máximas con media y varianza baja, sesgo negativo y curtosis positiva;
- ii. duraciones con media grande, una varianza pequeña, sesgo positivo y curtosis positiva;
- iii. periodos intereventos con varianza y sesgo pequeños y curtosis positiva; y
- iv. variabilidad de las amplitudes máximas y periodos con medias y varianza muy pequeñas (casi nulas);

es altamente probable que el animal esté rumiando. Una vez que se determina el tipo de actividad de la ventana, el bloque cálculo de parámetros estima los parámetros característicos de la ventana temporal (duración del paquete, duración promedio de los eventos, amplitud máxima promedio, entre otros).

La Figura 12 ilustra la construcción del bloque de almacenamiento y comunicaciones. Este bloque es el encargado de organizar la información generada por el bloque de análisis de actividades y otros sensores para su almacenamiento y transmisión. Este bloque recibe la información de actividad alimentaria, producida por el bloque de análisis de actividad, la del GPS y otros sensores y las organiza en un paquete de datos autocontenido al cual se le incorpora un código de corrección de errores sin necesidad de retransmisión de la información, y se la encripta para garantizar su

seguridad. La incorporación de la información del GPS, o cualquier otro sensor, agrega un elemento de trazabilidad a la información de comportamiento alimentario generada por el dispositivo. El paquete de información es almacenado en una memoria no volátil y eventualmente transmitido a una computadora personal para su posterior registro y procesamiento. La transmisión de la información generada por el dispositivo se realizará a través de una interfaz inalámbrica de baja potencia tipo LoRa o similar. Las comunicaciones serán periódicas durante el día, para informar del estado general del animal, pero la información generada será transmitida en puntos donde haya transceptores (bebederos, comederos, tranqueras, estación de ordeño, etc.) para minimizar el consumo de energía. Se permitirán transmisiones excepcionales cuando ocurra un evento anómalo.

Finalmente, la Figura 13 ilustra la construcción del bloque de administración de energía. Este bloque es el encargado de generar las diferentes líneas de alimentación utilizadas por el dispositivo a partir de una única fuente de energía y administrar la energía disponible en las baterías para maximizar el tiempo de operación del dispositivo. Esta tarea se realiza de dos maneras complementarias: i) asegurando la máxima carga posible de la batería a partir de la energía provista por un panel solar, y/o acoplamiento inductivo, y ii) controlando las funciones del dispositivo que deben encontrarse operativas en cada instante. Este bloque consta de un regulador DC-DC de alta eficiencia combinado con reguladores lineales de baja caída de tensión, para maximizar la eficiencia energética y posibilitar la operación del dispositivo aun cuando el voltaje de la batería esté por debajo del voltaje de operación.

Una implementación alternativa consiste en que la unidad de procesamiento se implemente en un único módulo compuesto por un microcontrolador y un conjunto de circuitos electrónicos integrados especializados que se utilizan para implementar los elementos de la unidad de procesamiento (Figura 15): i) la unidad de adquisición y procesamiento de señales, ii) la unidad de análisis de actividades y iii) la unidad de almacenamiento y comunicaciones. Todas las unidades se implementan de manera similar a la propuesta anterior, con la diferencia de que hay una única unidad de procesamiento (microcontrolador).

Los sonidos producidos por los animales y el ambiente se adquieren a través de micrófonos digitales direccionales conectados a los canales del Codec (Figura 14). El micrófono del animal se conecta a dos canales de entrada, ya que se requiere implementar dos caminos de procesamiento diferentes (Figura 2). El micrófono que adquiere el sonido ambiente se ubica en la unidad de procesamiento, mientras que el micrófono que adquiere el sonido del animal se ubica en su frente (Figura 3).

El codec se encuentra implementado en un circuito electrónico integrado de bajo consumo que incluye en cada uno de sus tres canales de entrada un circuito de acondicionamiento constituido por un filtro pasabajo, preamplificadores, un conversor

analógico-a-digital (ADC) y filtros digitales. Cada preamplificador proporciona un nivel de ganancia analógico programable independiente y un ajuste de impedancia a la entrada. Esto último proporciona una reducción del ruido pero puede no ser compatible con algunos micrófonos. Además, incluye un sistema de extensión de rango dinámico (DRE) y dither de datos de salida ADC. El sistema de DRE ajusta automáticamente la ganancia del canal en función del nivel detectado de señal, maximizando el rango dinámico ADC mientras se conserva el margen de entrada determinado la ganancia del preamplificador. El conversor analógico-digital tiene 16 bits de resolución y frecuencia de muestreo programable hasta 192 KHz. El circuito integrado incluye un interfaz digital serie tipo I2C dedicada para la programación de los parámetros de operación y una interfaz digital serie I2S para lectura de los datos.

Este circuito integrado además permite el filtrado y reproducción de cualquiera de los canales de entrada. Estas conexiones y bloques se utilizan para generar la señal de **Detección de actividad** (Figura 2), y el pin de salida es conectado a una entrada de comparación del microcontrolador para controlar su operación (Figura 15).

En la implementación se utilizó una frecuencia de muestreo de 2 KHz, con los circuitos de polarización de micrófonos activados y el sistema de DRE activado para los canales 1 y 3, de modo que el micrófono del animal (canal 1) y el ambiente (canal 3) tengan las mismas ganancias. El canal 2, por el cual también entra la señal del micrófono del animal, tiene el sistema de DRE desactivado para poder medir el efecto del sistema de DRE en la señal registrada. La señal del canal 1 es direccionada a los bloques de registro y reproducción, mientras que las señales de los canales 2 y 3 solo se direccionan a los bloques de registro.

Cuando la salida del bloque de reproducción tiene un valor mayor que un nivel pre-programado, el microcontrolador comienza a leer datos del codec utilizando la interfaz I2S implementada por medio de un módulo de hardware dedicado. La programación de los parámetros del circuito de adquisición se realizó a través una interfaz I2C, implementada a través de hardware dedicado disponible en el microcontrolador.

El bloque de detección y clasificación de eventos se implementa en el microcontrolador a través de las rutinas de software utilizadas para procesar y analizar los datos obtenidos del codec. Estas rutinas implementan en el microcontrolador los algoritmos de detección y clasificación de eventos (Figura 6 y 7) y las rutinas necesarias para extraer, clasificar y cuantificar la información indicativa de las eventos individuales. Estas rutinas procesan las señales adquiridas de manera continua para simplificar su procesamiento y reducir el consumo del procesador.

Las señales de audio registradas son procesadas por un filtro adaptativo de mínimos cuadrados (LMS) para cancelar los ruidos ambientes y remover las tendencias que se pudieran generar en la envolvente. Este filtro remueve en un único paso los ruidos y

perturbaciones que tenga la señal registrada, para ello utiliza la señal adquirida del ambiente en el canal 3 del codec como perturbación. Además, se calcula el cambio de ganancia introducido por el sistema de DRE a partir de realizar el cociente entre los valores medio de los canales 1 y 2 una por segundo (Figura 2). Este procesamiento es implementado como una rutina de software en microcontrolador.

Desde este punto en adelante, las dos implementaciones realizan las mismas tareas y de forma similar, por lo que se omitirá su explicación.

Otra implementación alternativa, comprende el uso de cualquiera de las implementaciones de la unidad de procesamiento propuestas anteriormente para clasificar las actividades alimentarias realizadas por el animal directamente del sonido sin clasificar los eventos masticatorios realizados por el animal. Se realizan las mismas tareas y procedimientos hasta la detección de los eventos, incluido la actualización y ajuste de los parámetros del algoritmo de detección de eventos, por lo que se omitirá su explicación. Luego se calculan los parámetros temporales característicos de cada evento (amplitud máxima y duración) así como el periodo entre eventos sucesivos (periodo interevento), los cuales son organizados en paquetes.

Un paquete se define como el conjunto de eventos ocurridos en una ventana de tiempo de 5 minutos. Entonces, una vez armados los paquetes se calculan los estadísticos (media, varianza, sesgo y curtosis) de cada una de los parámetros temporales característicos del paquete (amplitudes máximas, duraciones y periodos intereventos detectados), sus variabilidades (shimmer y jitter) y el tacograma. Además, se pueden incluir la energía de la envolvente del sonido en la banda 1.0 Hz a 1.5Hz, que se calcula a partir de la salida de un filtro pasabanda de Butterworth de octavo orden. Esta información es almacenada y luego transmitida al módulo que implementa el sistema de análisis de actividades, el almacenamiento y las comunicaciones a través de un transceptor NFC. Se eligió este tipo de transceptor porque la cantidad de datos a transmitir es muy pequeña (128 bytes por paquete) y nos permite recargar la batería del primer módulo a través del acoplamiento.

Cuando el receptor del transceptor del módulo de análisis de comportamiento recibe un paquete, genera una interrupción y saca al microcontrolador 2 de su estado de hibernación. Recibe el paquete de eventos, lee el GPS y los otros sensores, almacena toda esta información en un buffer y vuelve al modo de operación hibernación. Repite este proceso hasta que complete una hora de registros. Entonces, con toda esta información procede a estimar la actividad de cada paquete. Para ello utiliza un clasificador basado en un perceptron multicapas entrenado al efecto y clasifica todos los paquetes a partir de presentarle los estadísticos al perceptron (Figura 11). Los resultados obtenidos son almacenados en un buffer

temporal para, una vez completada la clasificación de los paquetes disponibles se los suaviza (filtrado de mediana). Este proceso consiste en analizar la etiqueta de cada paquete en el contexto de los paquetes alrededor. Si la etiqueta del paquete se desvía mucho del comportamiento más frecuente (mediana estadística) se procede a corregirlo en función de esta. Una vez clasificada la actividad se procede a calcular los parámetros y luego se almacena toda esta información.

Desde este punto en adelante, las implementaciones realizan las mismas tareas y de forma similar, por lo que se omitirá su explicación.

Ejemplo de aplicación

A continuación, se describe la presente invención mediante una forma de realización preferida que detalla de manera amplia, en referencia a las figuras presentadas y dando alternativas constructivas de cada una de las partes que la constituyen. Se brinda toda la información necesaria, conjuntamente con las referencias del estado de la técnica, para que la invención pueda ser reproducida por un técnico medio en la materia.

La unidad de sensado comprende un micrófono analógico direccional para adquirir los sonidos producidos durante las acciones alimentarias, ubicado en la cabeza del animal a través de bozal como se indica en la Figura 3, y se ubica en la cara interior de una banda ajustable (que puede ser elástica) del bozal, que se encuentra ubicada en la parte superior de la cabeza del animal. Esta zona permite adquirir de manera simple y efectiva los sonidos producidos durante la alimentación y el animal naturaliza rápidamente su presencia. La banda ajustable se halla protegida por una cobertura de material impermeable y goma espuma, la cual protege al micrófono de las condiciones ambientales. La combinación de la direccionalidad de la zona de captación del micrófono con la aislación pasiva a los ruidos ambientes de los materiales que cubren la banda son suficientes para operar sin inconvenientes a campo abierto, que es la condición operativa que se encuentran con mayor frecuencia. En caso de operar en establos, se agrega un segundo micrófono en la unidad de procesamiento.

El sistema se divide en dos módulos (Figura 13): i) Un módulo que implementa los módulos de adquisición, procesamiento, detección y clasificación de eventos, ubicado junto al micrófono, y ii) un módulo que implementa del resto del sistema, ubicado en la oreja del animal. Estos módulos están comunicados de manera inalámbrica a través de un transceptor Bluetooth de baja energía o uno de comunicaciones de campo corto (NFC). La principal ventaja del segundo es que provee un medio alternativo al acoplamiento inductivo para cargar la batería del primer módulo.

El **módulo de adquisición y procesamiento** se implementa a través de un codec, un dispositivo hardware capaz de procesar y transformar la señal de audio en un flujo de datos, y un microcontrolador para controlar su operación e implementar el **módulo**

de detección y clasificación de eventos, a través de rutinas de software. Completan el módulo de adquisición una memoria no volátil, un transceptor NFC y un sistema de alimentación constituido por una batería recargable y un cargador conectado a la antena del NFC y un acoplamiento inductivo.

El micrófono que adquiere los sonidos producidos por el animal está conectado a dos canales de entrada del codec (Figura 14.a), ya que se requiere implementar dos caminos de procesamiento diferentes (Figura 2). El codec se encuentra implementado en un circuito electrónico integrado de bajo consumo que incluye en cada uno de sus canales de entrada de: i) un circuito de acondicionamiento constituido por un filtro pasabajos y un preamplificador, ii) un conversor analógico-a-digital (ADC) y iii) filtros digitales para su posprocesamiento. Cada preamplificador proporciona un nivel de ganancia analógico programable independiente y un ajuste de impedancia a la entrada. Esto último permite reducir el ruido, pero se puede perder la compatibilidad con algunos micrófonos. El sistema incluye un sistema de extensión de rango dinámico (DRE) y dither de datos de salida ADC. El sistema de DRE ajusta automáticamente la ganancia del canal en función del nivel de la señal, maximizando el rango dinámico ADC mientras se conserva el margen de entrada determinado la ganancia del preamplificador. El conversor analógico-digital tiene 16 bits de resolución y frecuencia de muestreo programable hasta 192 KHz. El circuito integrado incluye un interfaz digital serie tipo I2C dedicada para la programación de los parámetros de operación y una interfaz digital serie I2S para lectura de los datos.

Este circuito integrado además permite el filtrado digital y la reproducción de cualquiera de los canales de entrada. Estas conexiones y bloques se utilizan para generar la señal de **Detección de actividad** (Figura 2), y el pin de salida es conectado a una entrada de comparación del microcontrolador 1 para controlar su operación (Figura 14.a).

En la implementación se utilizó una frecuencia de muestreo de 2 KHz, con los circuitos de polarización de micrófonos activados y el sistema de DRE activado solo para el canal 1. El canal 2 tiene el sistema de DRE desactivado para medir el efecto del DRE en el canal 1 sobre la señal registrada. La señal del canal 1 es direccionada a los bloques de registro y reproducción, mientras que la señal del canal 2 solo se direcciona al bloque de registro.

Cuando la salida del bloque de reproducción tiene un valor mayor que un nivel pre-programado, el microcontrolador 1 comienza a leer datos del codec utilizando la interfaz I2S implementada por medio de un módulo de hardware dedicado. La programación de los parámetros del circuito de adquisición se realizó a través de una interfaz I2C, implementada a través de hardware dedicado disponible en el microcontrolador.

El módulo de detección y clasificación de eventos se implementa en el microcontrolador 1 a través de las rutinas de software utilizadas para procesar y analizar los datos obtenidos del codec. Estas rutinas implementan en el microcontrolador los algoritmos de detección y clasificación de eventos (Figura 6 y 7) y las rutinas necesarias para extraer, clasificar y cuantificar la información indicativa de los eventos individuales. Estas rutinas procesan las señales adquiridas de manera continua para simplificar su procesamiento y reducir el consumo del procesador. Una de estas rutinas es la encargada de calcular el cambio de ganancia introducido por el sistema de DRE a partir de realizar el cociente entre los valores medio de los canales 1 y 2 una por segundo (Figura 2).

A continuación, se realiza la detección síncrona de la señal, que consiste en multiplicar, muestra a muestra, la señal acondicionada por sí misma. La información resultante se almacena en un buffer temporal de 14 muestras para su posterior procesamiento. Los datos resultantes se procesan siguiendo dos caminos (Figura 6):

- i) la señal resultante es filtrada, con un filtro de coeficientes enteros de segundo orden, submuestreada a 150 Hz y se almacena en el buffer de envolvente B_e ;
- ii) las muestras de la señal resultante son acumuladas en cada periodo de submuestreo (14 muestras) para calcular la **energía de frame** (Figura 7.a), es submuestreada a 150 Hz y se almacena en el buffer de energía B_E .

La envolvente del audio submuestreada es procesada con un filtro adaptativo de mínimos cuadrados (LMS) para remover las tendencias y offsets que se pudieran generar en la envolvente debido a ruidos y perturbaciones no canceladas por sistema de acondicionamiento. Con este procesamiento se completa la tarea de cancelación de ruidos y acondicionamiento de la señal de audio. Luego, se procede a detectar la ocurrencia de un evento masticatorio. Para ello se detecta la ocurrencia simultánea de un pico en la envolvente de audio (Figura 7.b) y que la envolvente sea mayor que un umbral $U(k)$ (Figura 8). El umbral $U(k)$ es generado por un algoritmo que toma en cuenta las siguientes características anatómicas y comportamiento animal (Figura 8.a): i) hay un periodo mínimo de tiempo entre dos movimientos de mandíbula consecutivos, ii) la duración de los movimientos está restringido a un periodo máximo y iii) la intensidad del sonido está restringida a un máximo dado. Teniendo en cuenta todas estas particularidades de los eventos, el algoritmo de detección de picos genera el umbral $U(k)$ con las siguientes particularidades:

- Un periodo insensible (T_{REF}) - es un periodo de tiempo después de detectar un evento en el cual el algoritmo no busca nuevos eventos. Se calcula como la fracción α ($0.5 < \alpha < 1$) de la duración promedio de los últimos cuatro eventos detectados.
- Un periodo máximo (T_{MAX}) - es el máximo tiempo que un evento se calcula como la fracción β ($1 \leq \beta < 2$) de la duración promedio de los últimos eventos detectados.

- El valor máximo esperado del umbral (U_{MAX}) - es el valor máximo esperado en el próximo pico. Se calcula como la fracción γ ($0.5 \leq \gamma < 1$) del promedio móvil de los últimos picos detectados en la envolvente.
- La velocidad de crecimiento del umbral (ΔU): es la cantidad que el umbral $U(k)$ es decrementado en cada muestra después del periodo insensible T_{REF} ha expirado. El umbral $U(k)$ solo cambia durante el periodo T_{REF} y T_{MAX} .

El valor del umbral de detección $U(k)$ se actualiza de la siguiente manera

$$\begin{aligned}
 U(k) &= U_{MAX} && \forall k < T_{ref}; \\
 U(k) &= U(k - 1) - \Delta U, && \forall T_{ref} > k < T_{MAX}; \\
 U(k) &= U \frac{(k-1)}{2} && \text{si } k = k^* + T_D \\
 U(k) &= 1.5 * U(k - 1) && \text{si } k = k^* + T_1 \text{ y ocurre una falsa detección;} \\
 U(k) &= U(k - 1) && \text{ninguno de los casos anteriores.}
 \end{aligned}$$

dónde k^* es la posición k donde se detectó el último evento.

Una vez detectado un potencial evento se procede determinar los datos que se utilizarán para calcular las características para clasificar el evento. Los datos que se utilizarán son aquellos que en el buffer de energía son mayores al umbral U_F y son consecutivos (Figura 8.b). Con este subconjunto de datos se procede a calcular las cinco características utilizadas para clasificar los eventos masticatorios (Figura 5): i) energía total; ii) duración; iii) simetría; iv) rapidez acumulada y v) cambio de signo.

Antes de proceder a la clasificación se **validan las características calculadas**, de modo de verificar si los valores están dentro del rango de un evento (Figura 6). Si las características no pasan la prueba de validación se procede a **ajustar los parámetros del algoritmo y sus umbrales**, pero si las características son validadas se procede a clasificar el evento con un clasificador, en esta implementación se utiliza un perceptron multicapa. El clasificador produce dos salidas a partir de las características: i) una etiqueta del evento (arranque, arranque-masticación, masticación pastoreo, masticación rumia y ruido) y ii) un número entre 0 y 1 indicando el grado de confianza de la clasificación.

Luego, hayan sido validadas o no, las características son utilizadas para ajustar los parámetros del algoritmo, se estiman la potencia de ruido y la potencia de la señal. La potencia del ruido se calcula como el valor medio de las muestras del primer y tercer segmento de cada buffer y el valor del umbral de energía U_F se actualiza

$$U_f = \rho T_{NFEL},$$

donde ρ es el parámetro de sensibilidad que se modifica en valores discretos de acuerdo con el valor promedio del nivel del SNR estimado para los últimos cuatro eventos clasificados.

Luego se calculan los parámetros característicos de cada evento (amplitud máxima, energía, etc.). Solo se calcula la energía de los arranques y las masticaciones de pastoreo. La energía de las masticaciones-arranques es separada; la energía de la parte de masticación (primer parte del evento Figura 5.d) se asigna a las masticaciones de pastoreo y la energía de la parte de arranque (segunda parte del evento Figura 5.d) se asigna los arranques. Esta información, combinada con la cantidad de arranques y su duración es utilizada por el modelo de consumo para estimar la cantidad de materia seca ingerida por el animal en el periodo de tiempo considerado (Figura 9).

Una vez detectados, clasificados y cuantificados los eventos individuales, otra rutina toma los resultados del clasificador de eventos y los organiza en paquetes. Un paquete se define como el conjunto de eventos ocurridos en una ventana de tiempo de 5 minutos. Entonces, una vez armados los paquetes se calculan los estadísticos del paquete (cantidad de masticaciones de pastoreo, tiempo total de masticación de pastoreo, amplitud máxima promedio de masticaciones de pastoreo, energía promedio de masticaciones de pastoreo, cantidad de arranques, tiempo total de arranque, amplitud máxima promedio de arranques, energía promedio de arranques, cantidad de masticaciones-arranques, cantidad de masticaciones de rumia, tiempo total de masticación de rumia, amplitud máxima promedio de masticaciones de rumia, entre otras), los cuales proveen la información para calcular el consumo de materia seca del animal. Esta información es almacenada y luego transmitida al módulo que implementa el sistema de análisis de actividades, el almacenamiento y las comunicaciones a través de un transceptor NFC. Se eligió este tipo de transceptor porque la cantidad de datos a transmitir es muy pequeña (128 bytes por paquete) y nos permite recargar la batería del primer módulo a través del acoplamiento.

Cuando el receptor del transceptor del módulo de análisis de comportamiento recibe un paquete, genera una interrupción y saca al microcontrolador 2 de su estado de hibernación. Recibe el paquete de eventos, lee el GPS y los otros sensores, almacena toda esta información en un buffer y vuelve al modo de operación hibernación. Repite este proceso hasta que complete una hora de registros. Entonces, con toda esta información procede a **estimar la actividad de cada paquete**. Para ello utiliza un clasificador basado en un perceptron multicapas entrenado al efecto y clasifica todos los paquetes a partir de presentarle los estadísticos relacionados con las etiquetas de los eventos (cantidad total de eventos, cantidad de masticaciones de pastoreo, cantidad de arranques, cantidad de masticaciones-arranques, cantidad de masticaciones de rumia, cantidad de ruidos, cantidad de silencios) al perceptron (Figura 9). Los resultados obtenidos son

almacenados en un buffer temporal para, una vez completada la clasificación de los paquetes disponibles se los suaviza (filtrado de mediana). Este proceso consiste en analizar la etiqueta de cada paquete en el contexto de los paquetes alrededor. Si la etiqueta del paquete se desvía mucho del comportamiento más frecuente (mediana estadística) se procede a corregirlo en función de esta. Una vez clasificada la actividad se procede a calcular los parámetros y la cantidad de materia seca ingerida y luego se almacena toda esta información.

La materia seca ingerida se estima, si el paquete es de pastoreo, utilizando una máquina de aprendizaje extremo (ELM) y la información estadística del paquete relacionada con la cantidad y energía de los eventos. La ELM se construye tomando como base la estructura del perceptrón de una capa oculta con N_H neuronas, los pesos de las neuronas son fijados aleatoriamente (por lo que no necesitan ser entrenadas), mientras que los pesos de salida se pueden obtener resolviendo un problema de mínimos cuadrados. Las variables de entrada al ELM son los estadísticos utilizados para estimar las variables relacionadas con la DMI (energía de masticación por masticación; energía de masticación por unidad de tiempo de comida y energía de arranque por arranque): la cantidad total de eventos, la cantidad, el tiempo total y la energía promedio de masticaciones de pastoreo, la cantidad, el tiempo total y la energía promedio de arranques, y la cantidad de masticaciones-arranques. Esta información se presenta directamente a la ELM, evitando los errores, problema numéricos y complejidad del código introducidos por las divisiones.

La información de la DMI es agregada al reporte del paquete e integrado con la información extra registrada (posición, temperatura, etc.), luego se la codifica para corregir los errores que se puedan producir durante su almacenamiento y transmisión. La información más importante (tiempo de rumia, pastoreo, cantidad de materia seca ingerida y posición) es transmitido a la estación base a través de un transceptor inalámbrico de baja potencia (Figura 12). La memoria no volátil del dispositivo del dispositivo está constituida por al menos 4 Gigabytes de memoria tipo Nand flash de bajo consumo. El módulo de comunicaciones puede implementar conexiones punto a punto y estrella con velocidades de transmisión de hasta 250 kbit/s y un alcance de 2000 metros a campo abierto, implementando el estándar LoRa. Tanto los factores de forma como el consumo de los módulos son pequeños y compatibles con los requerimientos de autonomía. El microcontrolador 2 se comunica con la memoria no volátil a través una interfaz SPI, implementada a través de hardware dedicado disponible en el microcontrolador. Por otro lado, el módulo de comunicación implementa una red punto a punto en configuración estrella con un nodo de comunicación dedicado, el cual se comunica con el microcontrolador a través una interfaz serie asíncrona tipo UART, implementada a través de hardware dedicado disponible en el microcontrolador.

Finalmente, el bloque de alimentación es el encargado de generar las diferentes líneas de alimentación utilizadas por el dispositivo a partir de una única fuente de energía y administrar la energía disponible en las baterías para maximizar el tiempo de operación del dispositivo. Esta tarea se realiza de dos maneras complementarias: i) asegurando la máxima carga posible de la batería a partir de la energía provista por un panel solar y ii) controlando las funciones del dispositivo operativas en cada instante. El control de las funciones del dispositivo se lleva a cabo deshabilitando funcionalidades del dispositivo de acuerdo a sus prioridades. Es decir, que en la medida en que la carga de la batería disminuye la frecuencia del uso de las comunicaciones inalámbricas se disminuye, hasta anularse en casos extremos, con el objeto de garantizar el registro y análisis de las actividades del animal. Este bloque es el encargado de gestionar el funcionamiento de los módulos y un reguladores DC-DC de alta eficiencia combinado con reguladores lineales de baja caída de tensión, para maximizar la eficiencia energética de este bloque y posibilitar la operación del dispositivo aun cuando el voltaje de la batería este por debajo del voltaje de operación del dispositivo.

Ejemplo de aplicación 2. Ensayo experimental

El sistema propuesto fue evaluado utilizando un conjunto de datos registrado en experimentos de campos sin ruido (DS1) y con ruido (DS2), en distintas combinaciones: i) entrenamiento con DS1 y testeo con DS1 (denominado DS11); ii) entrenamiento con DS1 y testeo con DS2 (denominado DS12); iii) entrenamiento con DS2 y testeo con DS1 (denominado DS21); iv) entrenamiento con DS2 y testeo con DS2 (denominado DS22). Las tasas de reconocimientos obtenidas en cada caso se muestran en las matrices de confusión mostradas en la Figura 15. Los resultados muestran que las masticaciones de rumia tienen un muy alto nivel de reconocimiento permitiendo diferenciarlos de todos los otros eventos asociados al pastoreo (masticaciones pastoreo y arranques). Además, la tasa de reconocimiento promedio es superior al 87.4 % en todas las combinaciones.

Para comparar los resultados obtenidos por el sistema propuesto con el sistema propuesto en la patente AR103025A1, ambos algoritmos fueron configurados, evaluados y testeados en las mismas condiciones, utilizando en ambos algoritmos un clasificador MLP. Además, en el sistema propuesto se agruparon los eventos masticaciones rumia y masticaciones pastoreo en una sola clase de evento masticaciones, debido a que no es posible realizar una comparación directa entre ambos algoritmos.

Tabla 1: Comparación de resultados para los cuatro conjuntos de datos.

	DS11		DS12		DS21		DS22	
	CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA	CBIA	CBEBA
Recall	98.8	98.4	98.5	98.0	98.0	97.9	98.7	98.4
Precision	98.8	98.2	91.7	95.8	98.0	98.7	91.9	96.1
F1-Score	98.8	98.8	95.0	96.9	97.8	98.3	95.1	97.2
Recall _M	91.7	91.8	72.9	90.5	82.0	92.9	86.7	90.0
Precision _M	93.7	92.1	79.6	85.9	83.0	85.1	86.8	90.2
F1-Score _M	92.5	91.9	77.5	88.0	82.9	88.8	87.3	90.0
Recall _G	90.8	91.5	68.0	90.2	80.2	92.7	85.3	89.5
Precision _G	93.5	91.6	78.5	84.3	81.6	83.3	85.7	89.6

La Tabla 1 muestra los resultados obtenidos con los sistemas propuestos en la patente AR103025A1 (etiquetados como CBIA) y en la presente (etiquetados como CBEBA) respectivamente. Se comprobó si las discrepancias de los resultados obtenidos eran estadísticamente significativas ($p < 0,05$) usando la prueba de los rangos con signo de Wilcoxon. Se observa que, al evaluar los algoritmos en presencia de ruido (DS12 y DS22), la detección de evento del sistema propuesto logra mejores resultados que el sistema propuesto por la patente AR103025A1 (CBIA), gracias a una menor cantidad de falsos eventos detectados (falsos positivos), lo cual se cuantifica con el Precision. Esto indica que el sistema propuesto detecta eventos de forma superior que la patente AR103025A1 (CBIA) en ambientes ruidosos.

Al evaluar las tasas de reconocimiento, ambos algoritmos logran tasas similares de clasificación en ausencia de ruido (DS11), pero el sistema propuesto logra mejores resultados que el de la patente AR103025A1 (CBIA) en presencia de ruido, tanto para entrenar, testear o ambas. Observando la disminución de las tasas de clasificación al comparar los resultados al testear con ruido y sin ruido (DS12 contra DS11), se observa que la disminución en el algoritmo propuesto es menos de la mitad producida en el algoritmo de la patente AR103025A1 (CBIA). Para los casos de condiciones de entrenamiento con ruido (DS21 y DS22), el algoritmo propuesto (CBEBA) logró mejores métricas generales de clasificación que el algoritmo de la patente AR103025A1 (CBIA), respectivamente, lo que indica su mayor sensibilidad al ruido ambiental.

Las tasas de reconocimiento de actividades obtenidas para cada método y las variantes propuestas se muestran en la Figura 17. El algoritmo de la patente AR100013B1 (Figura 17.d) es capaz de reconocer mejor el pastoreo que el algoritmo

de la patente AR102025A1 (Figura 17.c) a costa de un peor reconocimiento de la rumia y otro, obteniendo en ambos casos tasas de reconocimiento promedio similares. Sin embargo, el costo computacional del algoritmo de la patente AR102025A1 es 47 veces menor que el de la patente AR100013B1, permitiendo implementarlo en un dispositivo embebido.

Finalmente, las dos variantes del algoritmo propuesto (Figura 17.a y 17.b) obtuvieron mejores tasas de reconocimiento de rumia y pastoreo, teniendo una mínima confusión entre ellas. El algoritmo basado en eventos identificados (Figura 17.a) obtuvo tasas de reconocimiento superiores al 90% para ambas actividades, con confusiones despreciables (menores al 2%) entre ellas. La mayor confusión se dio con otras actividades. Por otro lado, el algoritmo basado en características temporales de los eventos detectados (Figura 17.b) obtuvo tasas de reconocimiento similares a los de las patentes anteriores en todas las clases. Es decir, tuvo un buen reconocimiento de ambas actividades alimentarias con tasas superiores al 78% para ambas actividades, con confusiones del orden del 10% entre ellas. Los resultados obtenidos con esta variante son más consistentes que los de las patentes anteriores.

Por otro lado, se realizó un ensayo para comparar la robustez de los métodos ante distintos niveles de ruido. Se utilizó ruido blanco gaussiano sumado a cada una de las señales utilizada en las proporciones adecuadas para lograr la relación señal a ruido (SNR) deseada. En la Figura 18 se muestran las tasas de reconocimiento promedio obtenido para distintos niveles SNR. Los resultados muestran que el sistema propuesto siempre alcanza una tasa de reconocimiento superior al sistema propuesto en la patente AR103025A1 lo cual se debe, por un lado, a que las características de actividad fueron menos afectadas por los errores de clasificación de eventos masticatorios, y por el otro, a que el algoritmo de clasificación de eventos masticatorios propuesto es más robusto en condiciones de ruido que el propuesto en la patente AR103025A1. Por otro lado, la tasa de reconocimiento del sistema propuesto en la patente AR103025A1 disminuye considerablemente para SNR menores a 20 dB mostrando poca robustez al ruido aditivo, mientras que en el algoritmo propuesto la tasa de reconocimiento se mantiene prácticamente constante hasta 0 dB, evidenciando una gran tolerancia frente al ruido.

Referencias

1. Clapham W, Fedders J, Beeman K and J Neel. "Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle", *Computers and Electronics in Agriculture*, vol. 76, pp. 96-104, (2011).
2. Milone D, Rufiner H, Galli J, Laca E and C Cangiano. "Computational method for segmentation and classification of ingestive sounds in sheep", *Computers and Electronics in Agriculture*, vol. 65, pp. 228 - 237, (2009).
3. Milone D, Galli J, Cangiano C, Rufiner H and E Laca. "Automatic recognition of ingestive sounds of cattle based on hidden Markov models", *Computers and Electronics in Agriculture*, vol. 87, pp. 51 - 55, (2011).
4. Navon S, Mizrach A, Hetzroni A and E Ungar. "Automatic recognition of jaw movements in free-ranging cattle, goats and sheep, using acoustic monitoring", *Biosystems Engineering*, vol. 114(4), pp. 474-483 (2013).
5. Chelotti J, Vanrell S, Milone D, Utsumi S, Galli J, Rufiner L and L Giovanini. "A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle", *Computers and Electronics in Agriculture*, 127, 64-75 (2016).
6. Deniz N, Chelotti J, Galli J, Planisich A, Larripa M, Rufiner L and L Giovanini. "Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals", *Computers and Electronics in Agriculture*, 138, 167-174 (2017).
7. Martiskainen P, Järvinen M, Skön J, Tiirikainen J, Kolehmainen M and J Mononen. "Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines", *Appl. Anim. Behav. Sci.*, 119, 32-38 (2009).
8. Diosdado J, Barker Z, Hodges H, Amory J, Croft D, Bell N and E Codling, "Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system", *Animal Biotelemetry*, 3, 1-14 (2015).
9. Rayas-Amor A, Morales-Almaráz E, Licon-Velázquez G, Vieyra-Alberto R, García-Martínez A, Martínez-García C, Cruz-Monterrosa R and G Miranda-de la Lama, "Triaxial accelerometers for recording grazing and ruminating time in dairy cows: An alternative to visual observations", *Journal of Veterinary Behavior*, 20, 102-108 (2017).
10. Benaissa S, Tuytens F, Plets D, De Pessemier T, Trogh J, Tanghe E, Martens L, Vandaele L, Van Nuffel A, Joseph W and others, "On the use of on-cow accelerometers for the classification of behaviours in dairy barns", *Research in veterinary science*, 125, 425-433 (2019).
11. Benaissa S, Tuytens F, Plets D, Cattrysse H, Martens L, Vandaele L, Joseph W and B Sonck, "Classification of ingestive-related cow behaviours using RumiWatch halter and neck-mounted accelerometers", *Applied Animal Behaviour Science*, 211, 9-16 (2019).
12. Shen W, Cheng F, Zhang Y, Wei X, Fu Q and Y Zhang, "Automatic recognition of ingestive-related behaviors of dairy cows based on triaxial acceleration", *Information Processing in Agriculture* 7, 427-443 (2020).

13. Vanrell S, Chelotti J, Galli J, Utsumi S, Giovanini L, Rufiner H and Milone D, "A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle". *Computers and Electronics in Agriculture*, 151, 392-402 (2018).
14. Chelotti J, Vanrell S, Galli J, Giovanini L and L Rufiner. "A pattern recognition approach for detecting and classifying jaw movements in grazing cattle", *Computers and Electronics in Agriculture*, 145, 83-91 (2018).
15. Chelotti J, Vanrell S, Rau L, Galli J, Planisich A, Utsumi S, Milone D, Giovanini L and L Rufiner. An online method for estimating grazing and rumination bouts using acoustic signals in grazing cattle. *Computers and Electronics in Agriculture*, 173, 105443 (2020).
16. Li G, Xiong Y, Du Q, Shi Z and R Gates. "Classifying Ingestive Behavior of Dairy Cows via Automatic Sound Recognition", *Sensors*, 21(15), 5231 (2021).

Reivindicaciones

Habiendo así especialmente descrito y determinado la naturaleza de la presente invención y la manera de llevarla a la práctica, se declara reivindicar como de propiedad y derecho exclusivo:

Reivindicaciones:

1. Un método automático para monitorear las actividades alimentarias realizadas por animales rumiantes caracterizado porque comprende:
 - a. adquirir y registrar los sonidos y/o vibraciones producidos por un animal;
 - b. identificar en dichos sonidos y/o vibraciones del paso a.), aquellos que son producidos por dicho animal durante su alimentación;
 - c. detectar en la señal de sonido del paso b) los movimientos mandibulares que realiza dicho animal durante la alimentación;
 - d. establecer una relación entre las características temporales del sonido y los movimientos de la mandíbula durante su alimentación;
 - e. identificar en la señal de sonido del paso b), eventos masticatorios en base a la relación establecida del paso d), donde dichos eventos son clasificados en base al período, la amplitud y la energía del sonido;
 - f. establecer una relación entre las características estadísticas del sonido y los porcentajes de movimientos de la mandíbula con las actividades alimentarias durante su alimentación;
 - g. identificar en los eventos masticatorios del paso e) las actividades alimentarias en base a la relación establecida del paso f), donde dichas actividades son clasificadas en base a las características estadísticas del sonido del paso b) y los porcentajes de los eventos identificados en una ventana de tiempo fija durante su alimentación;
2. El método de la reivindicación 1, porque además comprende un paso adicional h. caracterizado porque comprende establecer una relación entre los movimientos mandibulares y los parámetros característicos del sonido (el periodo y energía del sonido) con la cantidad y tipo de alimento ingerido durante la alimentación;
3. El método de la reivindicación 1 y 2, porque además comprende un paso adicional i. caracterizado porque comprende cuantificar los eventos del paso e. y estimar la cantidad de alimento ingerido en base a la relación establecida del paso h);

4. El método de las reivindicaciones 1 a 3, caracterizado porque en dicho paso c., los movimientos mandibulares son, al menos, arranque, arranque-masticación, masticación de pastoreo y masticación de rumia.
5. El método de las reivindicaciones 1 a 3, caracterizado porque en dicho paso d., la relación entre los movimientos de la mandíbula y los sonidos producidos por la mandíbula durante la alimentación está dada por las características temporales del sonido.
6. El método de la reivindicación 5, caracterizado porque dichas características temporales del sonido son la forma de la energía instantánea, energía total, velocidad de cambio, duración, periodicidad, secuencia y simetría.
7. El método de las reivindicaciones 1 a 3, caracterizado porque en el paso e., dicha identificación comprende detectar y clasificar los eventos mediante el análisis de las envolventes del sonido y su energía.
8. El método de las reivindicaciones 1 a 3, caracterizado porque en dicho paso f., las actividades alimentarias son, al menos, pastorear, rumiar y hay que decir exactamente las actividades que podemos monitorear.
9. El método de las reivindicaciones 1 a 3, caracterizado porque en dicho paso g., la relación entre las actividades alimentarias y los movimientos de la mandíbula está dada por los porcentajes de los movimientos de mandíbula que ocurren en una ventana temporal de duración fija.
10. El método de las reivindicaciones 1 a 3, caracterizado porque en dicho paso g., la relación entre las actividades alimentarias y los sonidos producidos por la mandíbula durante la alimentación está dada por las características estadísticas del sonido en una ventana temporal de duración fija.
11. El método de la reivindicación 10, caracterizado porque dichas características estadísticas de la amplitud máxima, duración, variabilidad y tacograma de la envolvente del sonido son su media, desviación estándar, sesgo y curtosis.
12. El método de las reivindicaciones 1 a 3, caracterizado porque en el paso g., la identificación comprende detectar y clasificar las actividades alimentarias mediante el análisis de los movimientos mandibulares y/o las estadísticas de la envolvente del sonido.
13. El método de las reivindicaciones 2 y 3, caracterizado porque en dicho paso h., la estimación de la cantidad de alimento ingerido comprende evaluar una regresión lineal o no lineal, o un modelo de aprendizaje automático, que utiliza el tipo de actividad alimentaria y parámetros característicos de los sonidos (periodo y energía) producidos durante el pastoreo para el valor del tiempo.

14. El método de las reivindicaciones 1 a 3, caracterizado porque dichos pasos son realizados mediante métodos de procesamiento digital de señales y un modelo de aprendizaje automático entrenado para los sonidos registrados.
15. El método de la reivindicación 9, en el que el modelo de aprendizaje automático puede ser un perceptrón multicapa, un árbol de decisiones, una máquina de soporte vectorial o cualquier otro modelo de aprendizaje automático disponible.
16. El método de la reivindicación 10, en el que el modelo de aprendizaje automático entrenado se basa en un perceptrón multicapa o un árbol de decisiones.
17. El método de la reivindicación 11, que comprende además entrenar el modelo de aprendizaje automático utilizando observaciones históricas del comportamiento del animal.
18. El método de las reivindicaciones 1 a 3 caracterizado porque comprende, además, generar una salida en una interfaz de usuario indicativa de la actividad alimentaria realizada, información relevante relacionada con la actividad y, si corresponde el valor de consumo de pasto del animal.
19. Un sistema para monitorear las actividades alimentarias y el consumo de pasto por animales rumiantes, para ejecutar el método de la reivindicación 1, caracterizado porque comprende:
 - una unidad de sensado de sonido o vibraciones espacialmente asociado con la cabeza del animal para adquirir los sonidos y vibraciones producidas a lo largo del tiempo;
 - una unidad de procesamiento que comprende: una interfaz para recibir y acondicionar los datos de sonido y vibraciones registrados; un procesador configurado para evaluar los datos de sonido y vibraciones registrados para identificar los movimientos mandibulares, el comportamiento alimentario y consumo de pasto del animal en cada periodo de tiempo;
 - una unidad de almacenamiento no volátil para identificar cada animal y almacenar la información generada;
 - una unidad de comunicación para transmitir la información generada por cada animal a un servidor; y
 - una interfaz de usuario configurada para almacenar, gestionar, analizar y reportar la información relacionada con las actividades alimentarias realizadas y la cantidad de alimento ingerido por el animal a lo largo del tiempo.

20. El sistema de la reivindicación 19 incluye múltiples micrófonos para adquirir los sonidos y vibraciones producidas por el animal y el ambiente en el cual se encuentra.
21. El sistema de cualquiera de las reivindicaciones 19 ó 20, en el que los sensores están montados mediante un cabestro sobre la cabeza de los animales y/o mediante una caravana en la oreja.

Figuras

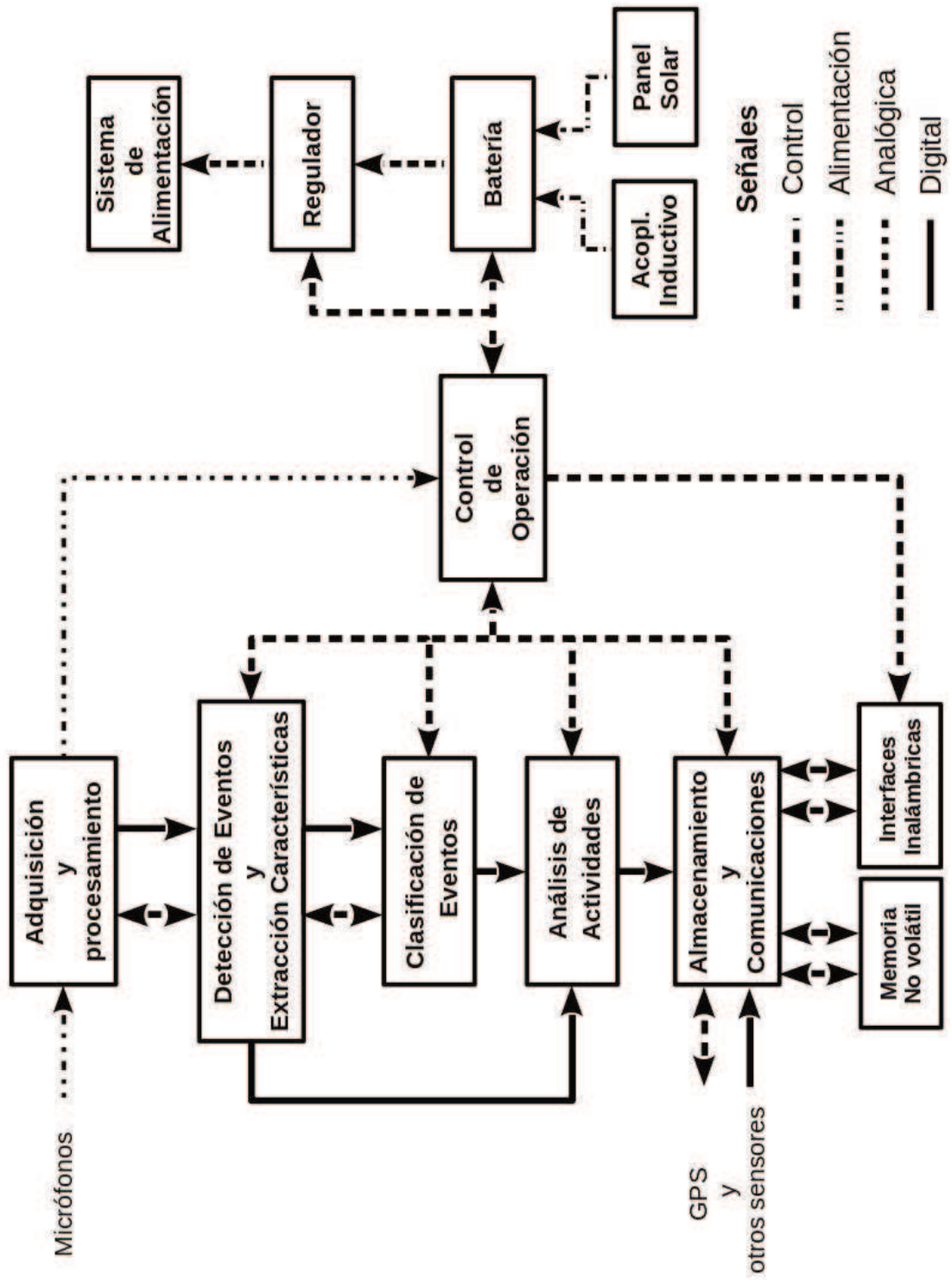


Figura 1

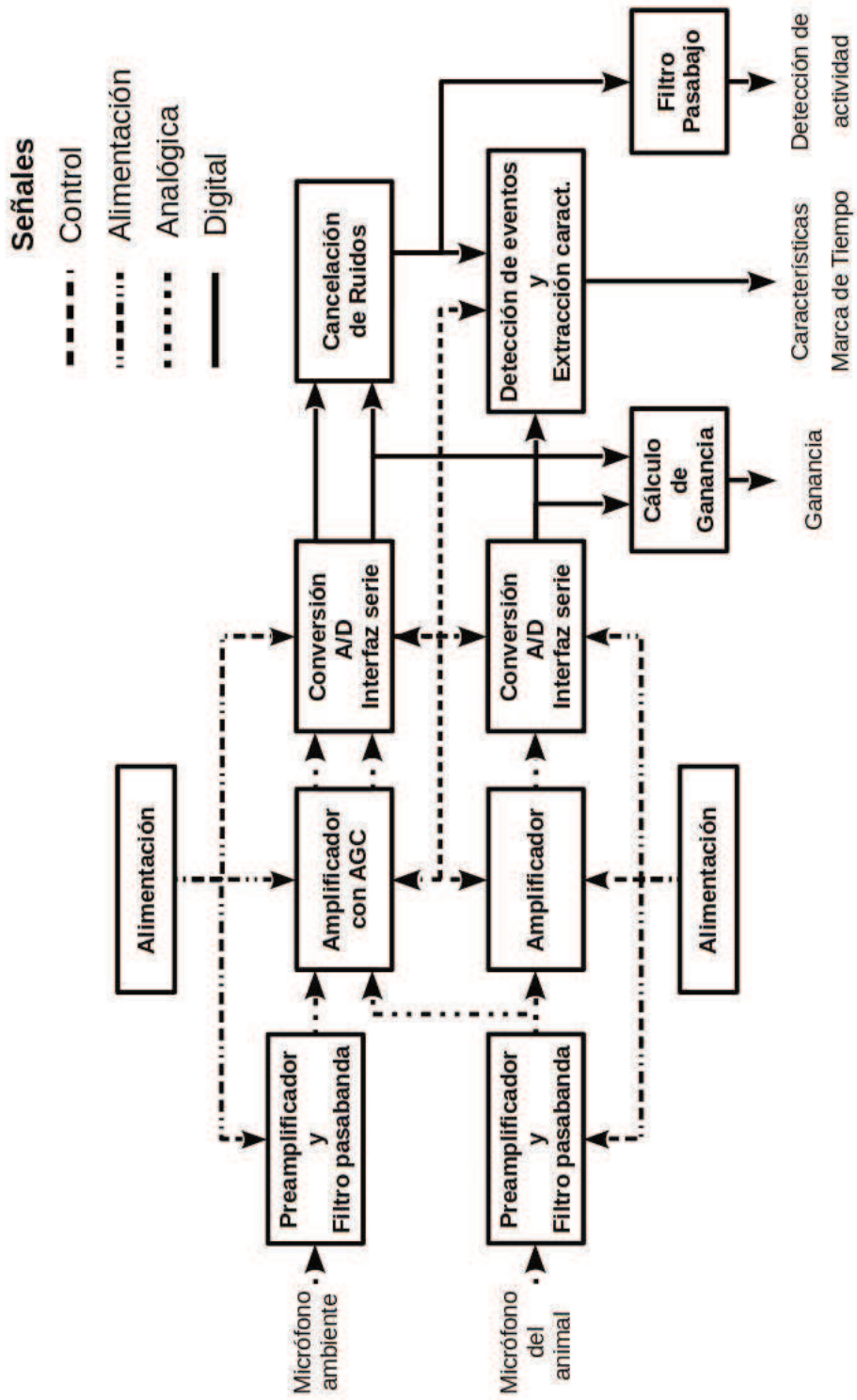


Figura 2

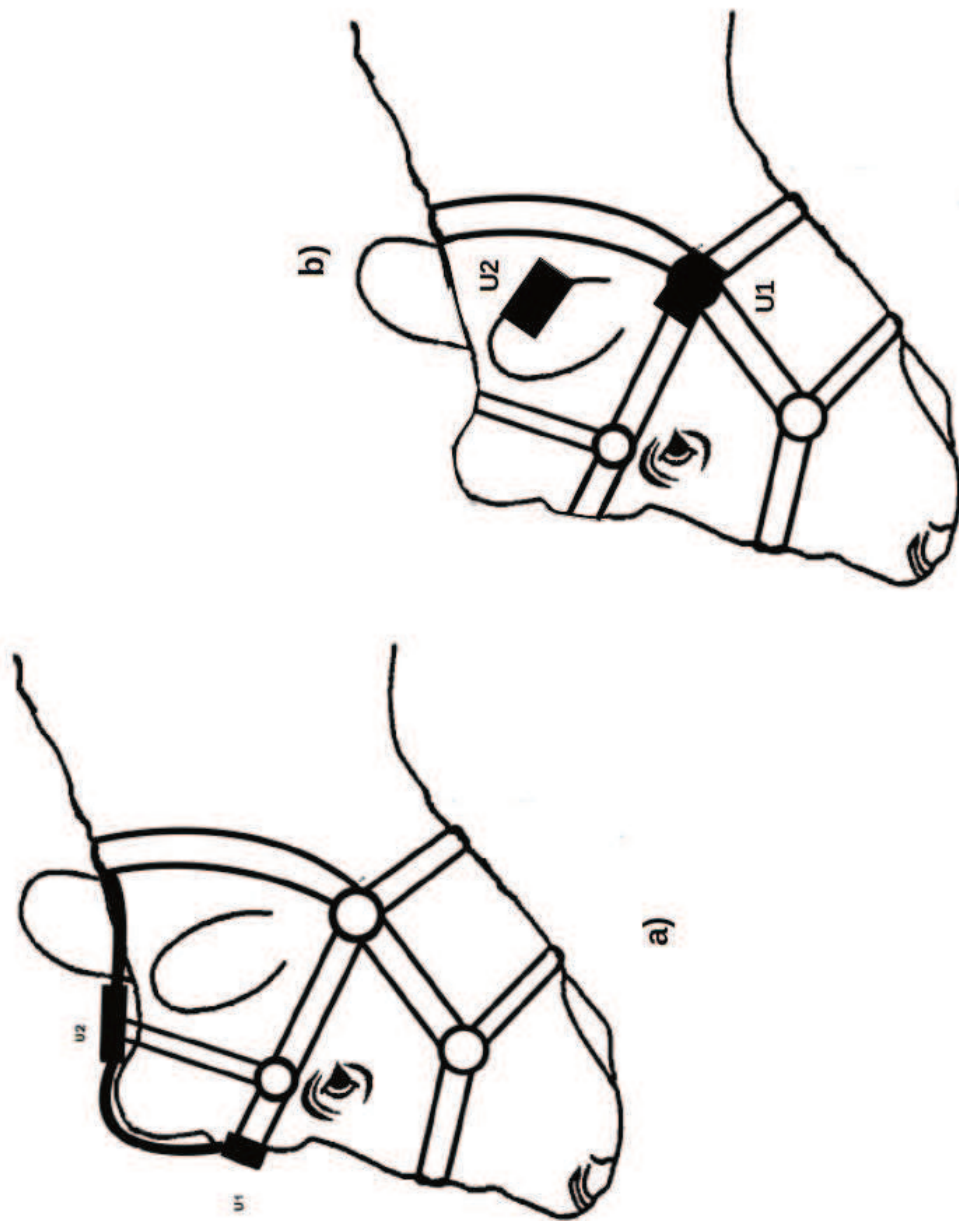


Figura 3

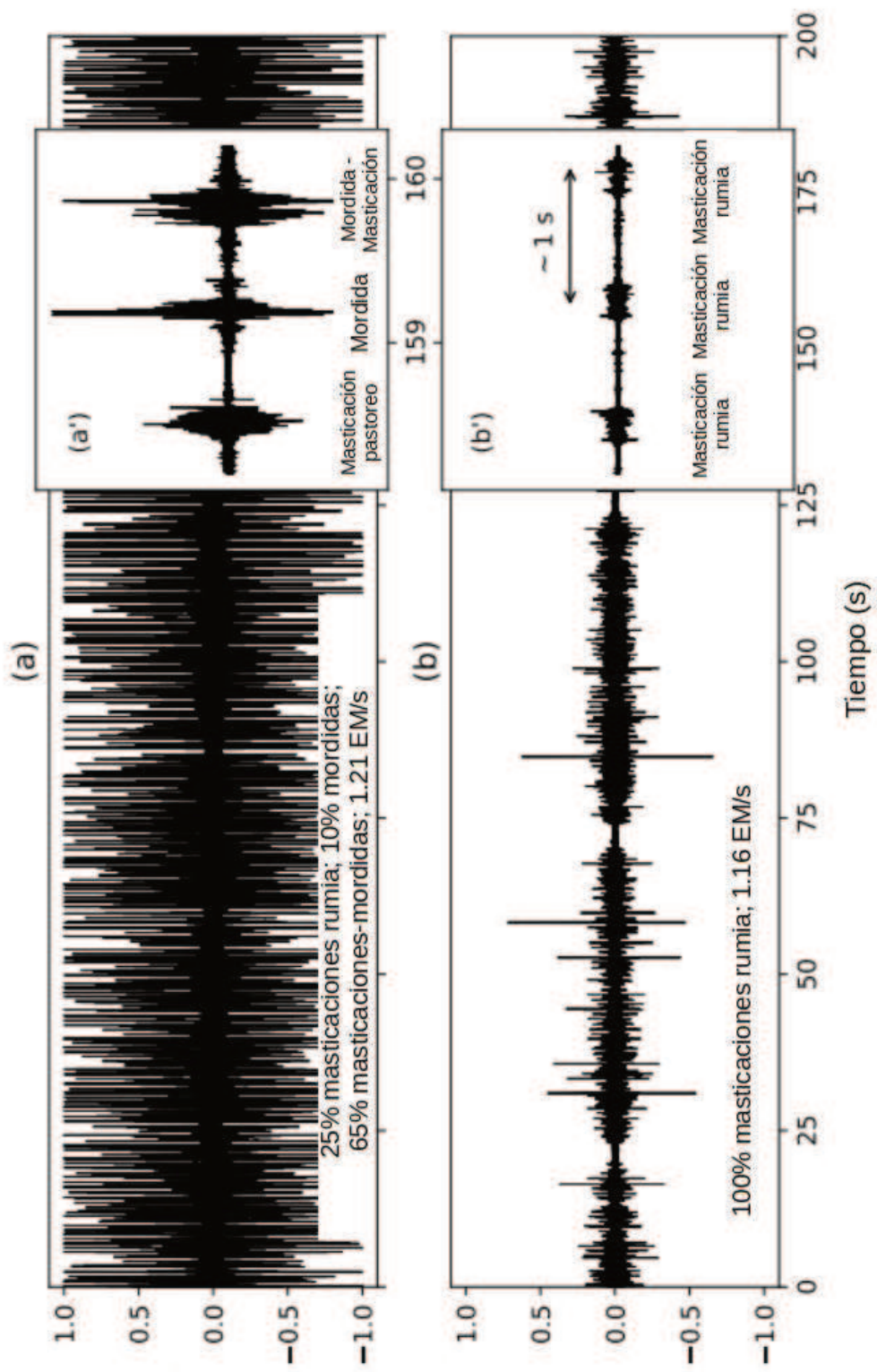


Figura 4

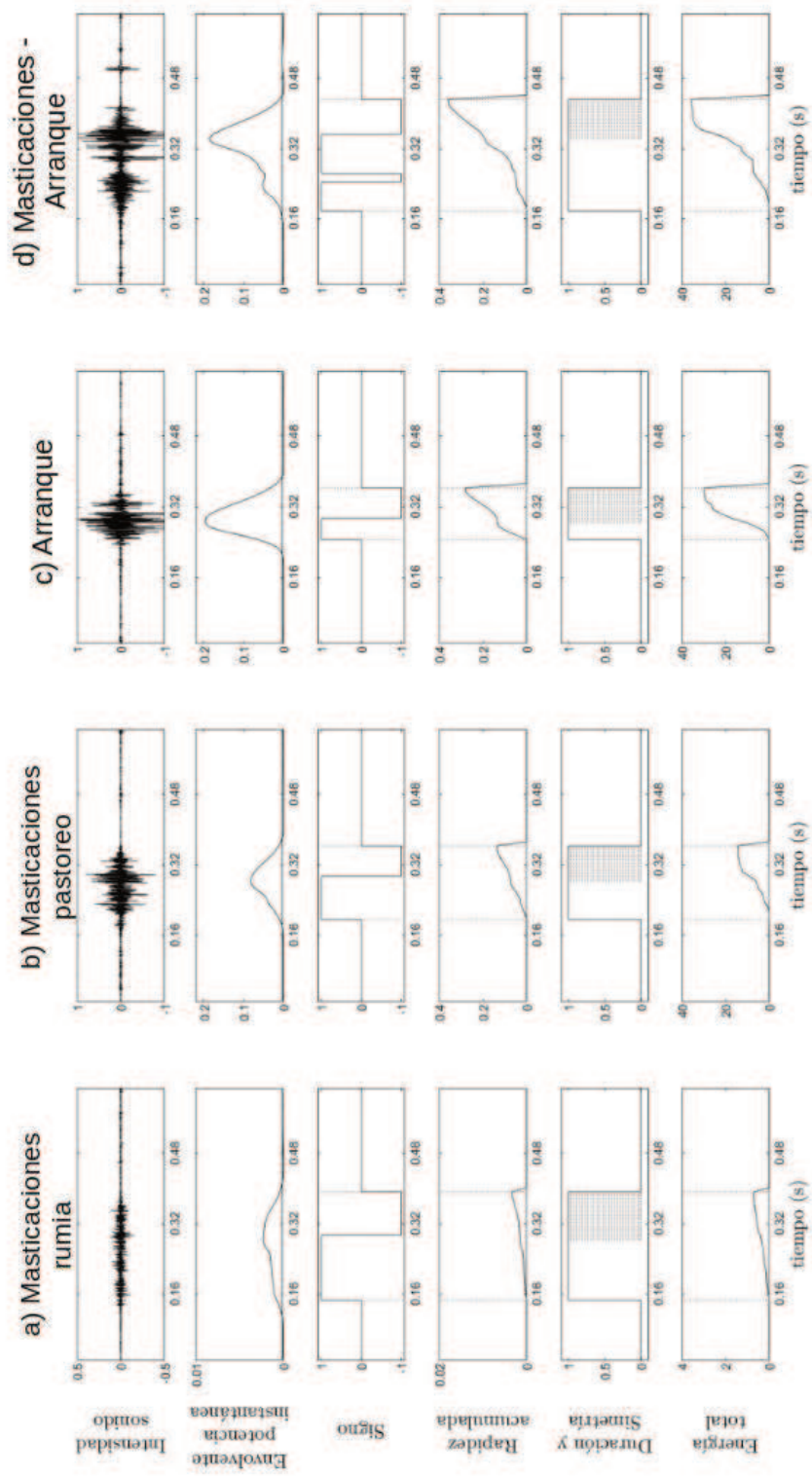


Figura 5

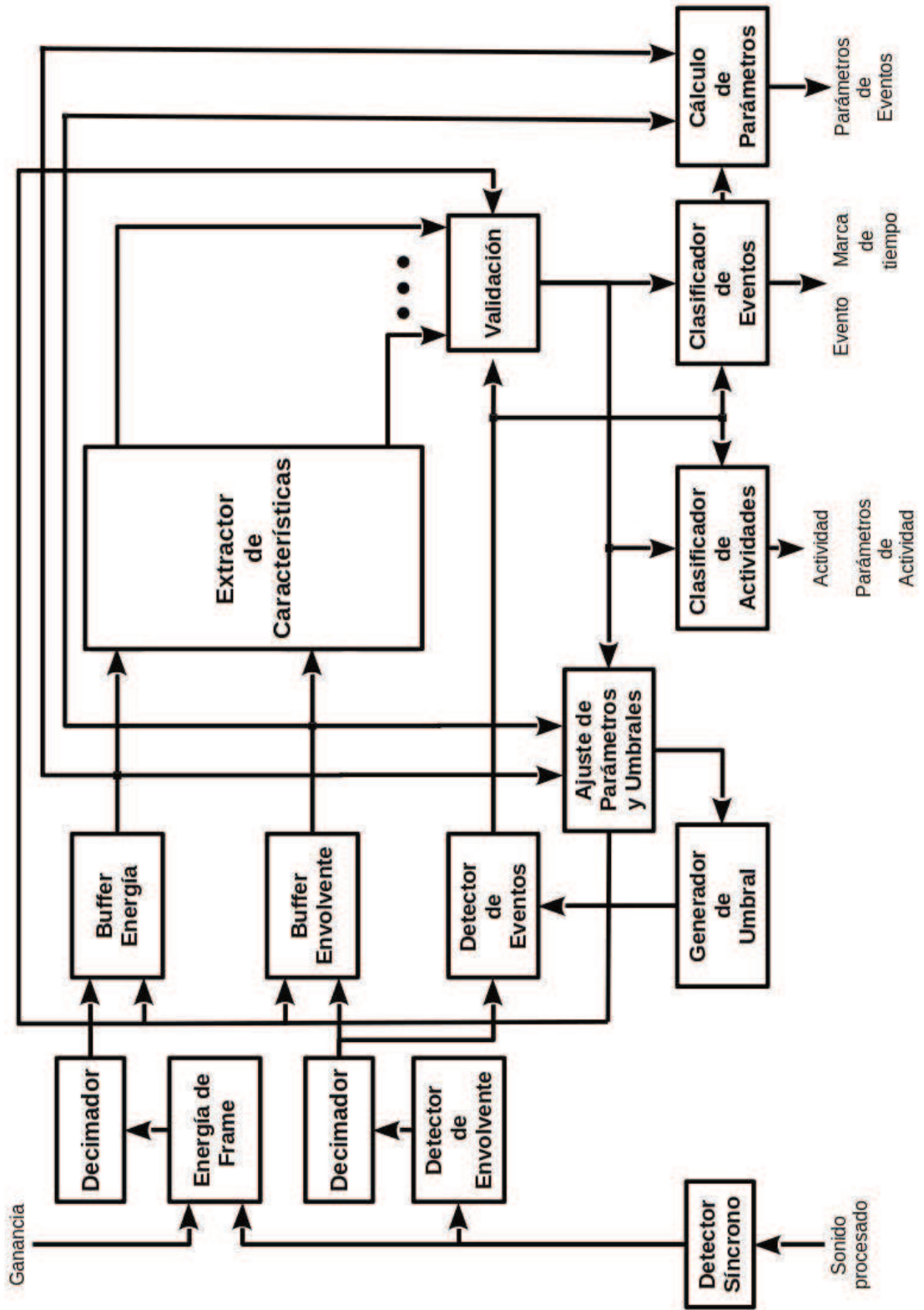


Figura 6

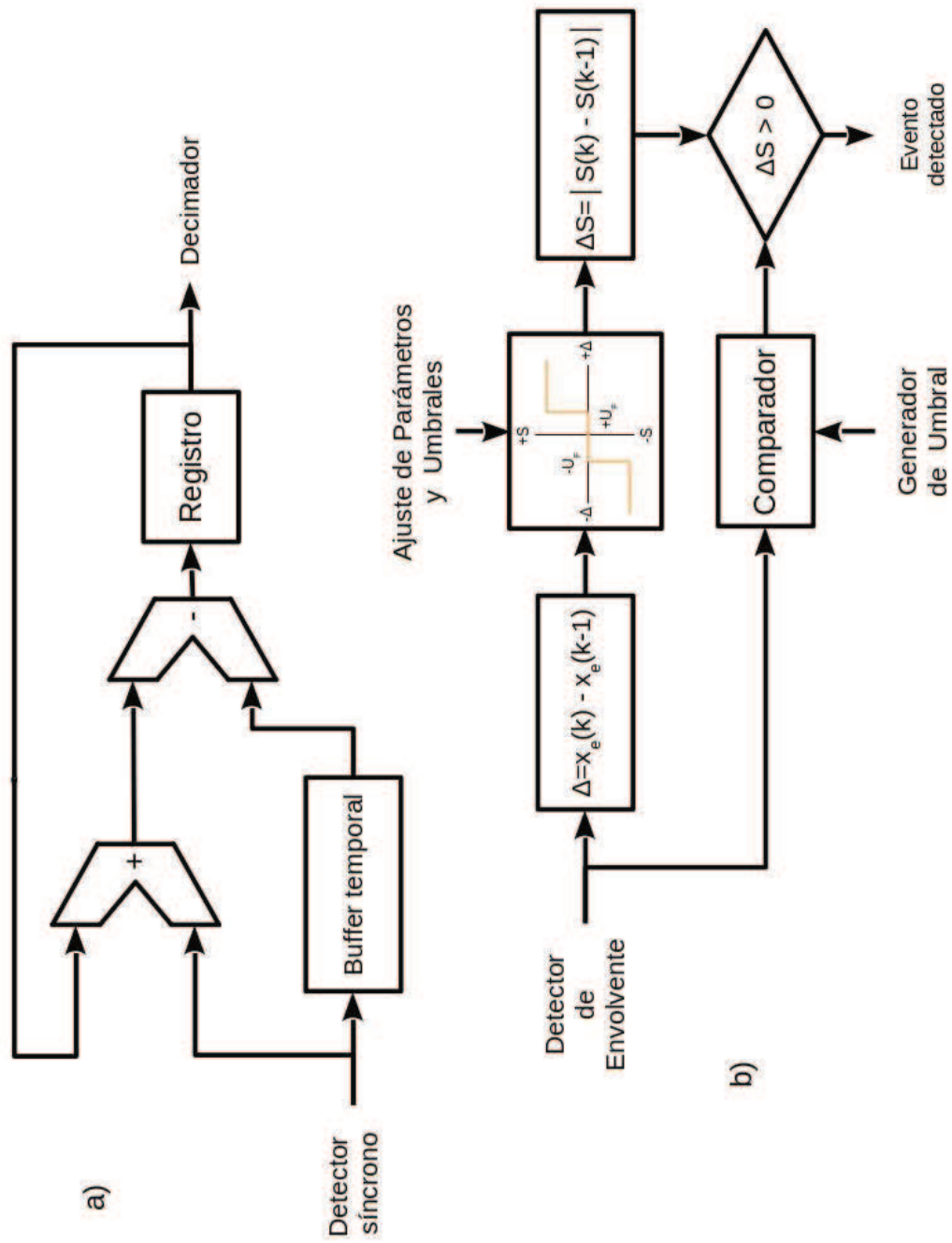


Figura 7

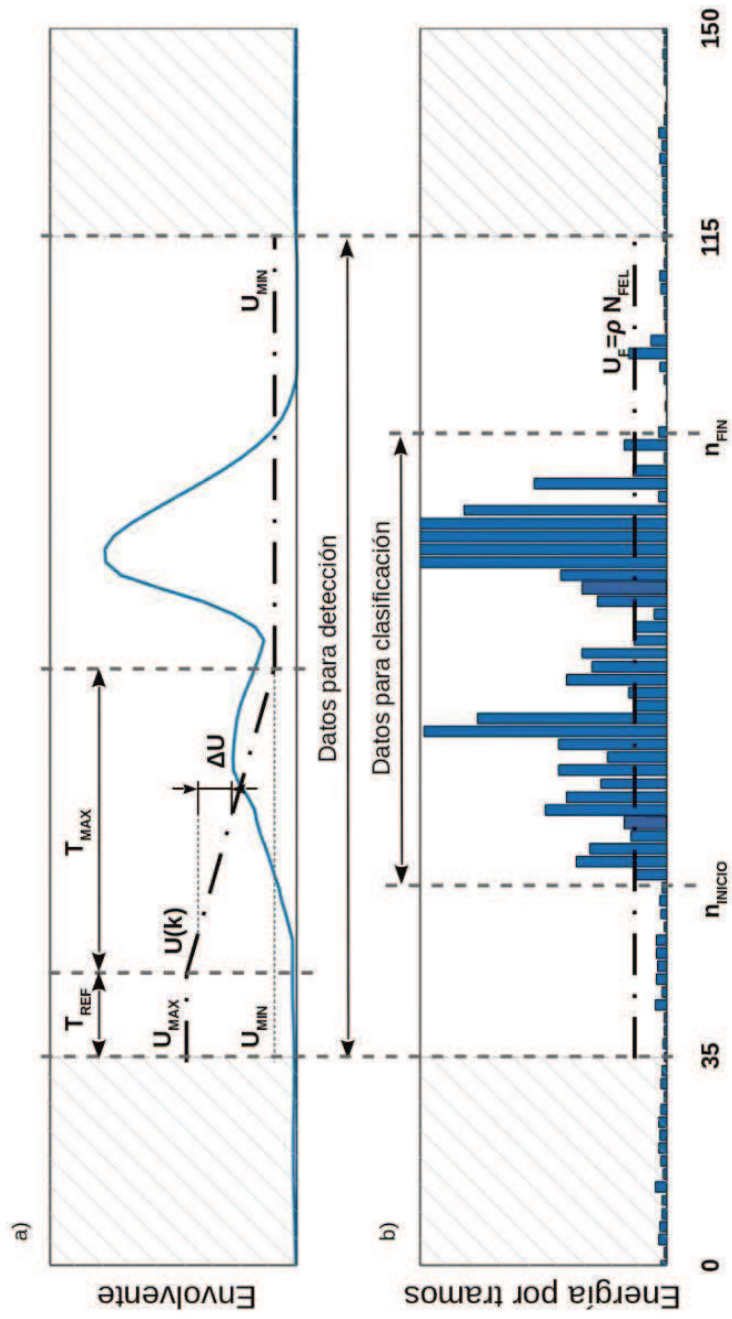


Figura 8

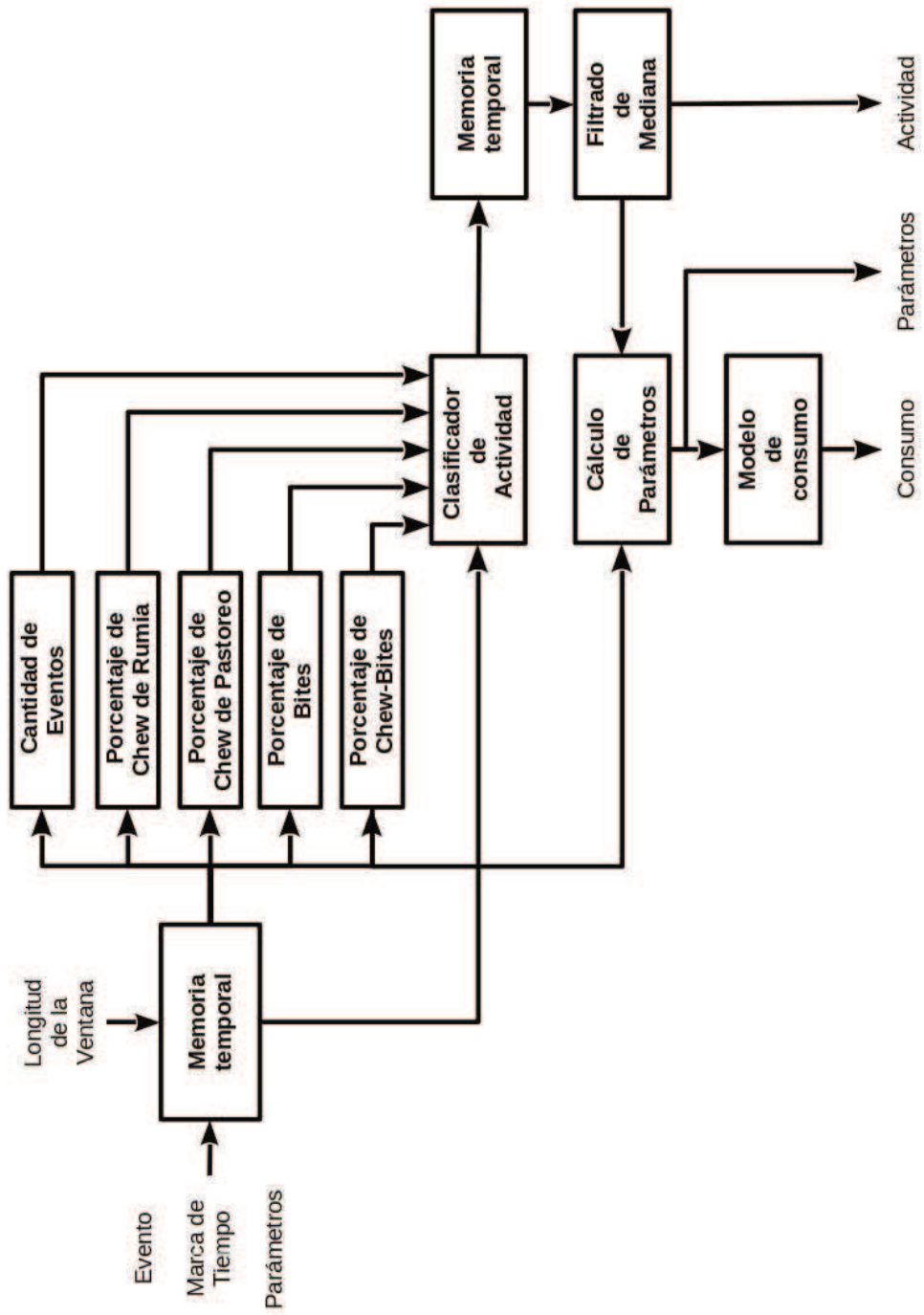


Figura 9

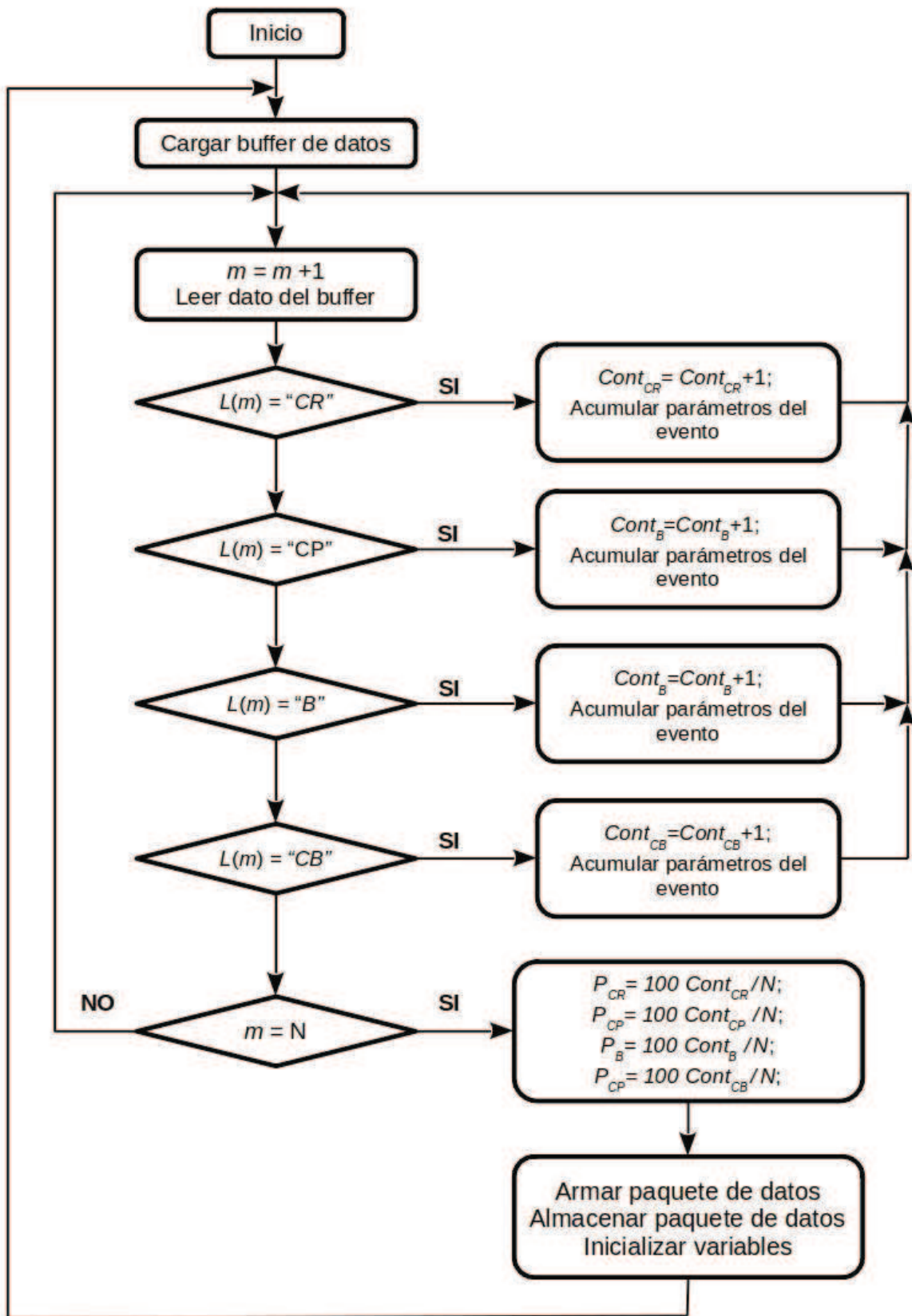


Figura 10

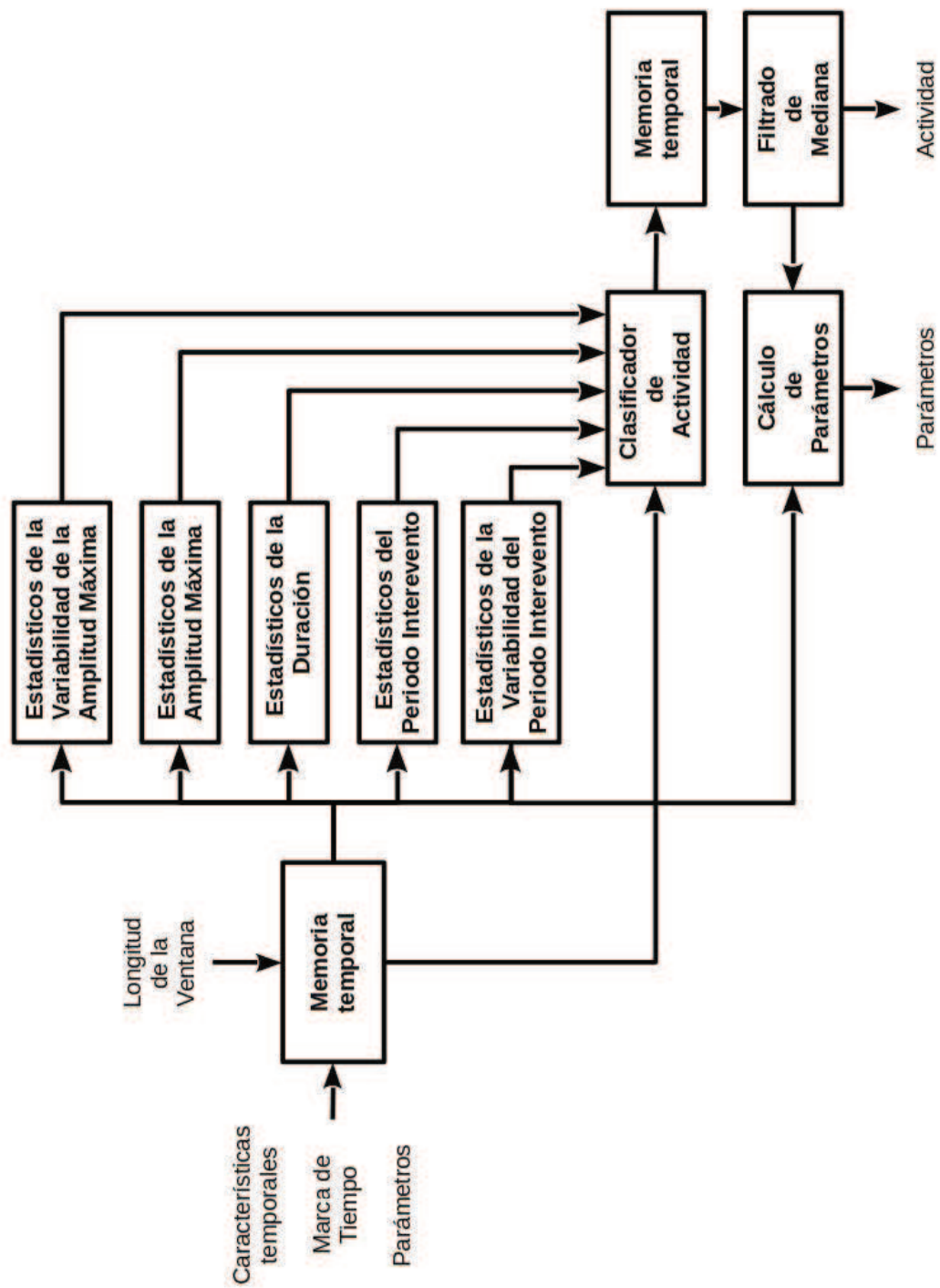


Figura 11

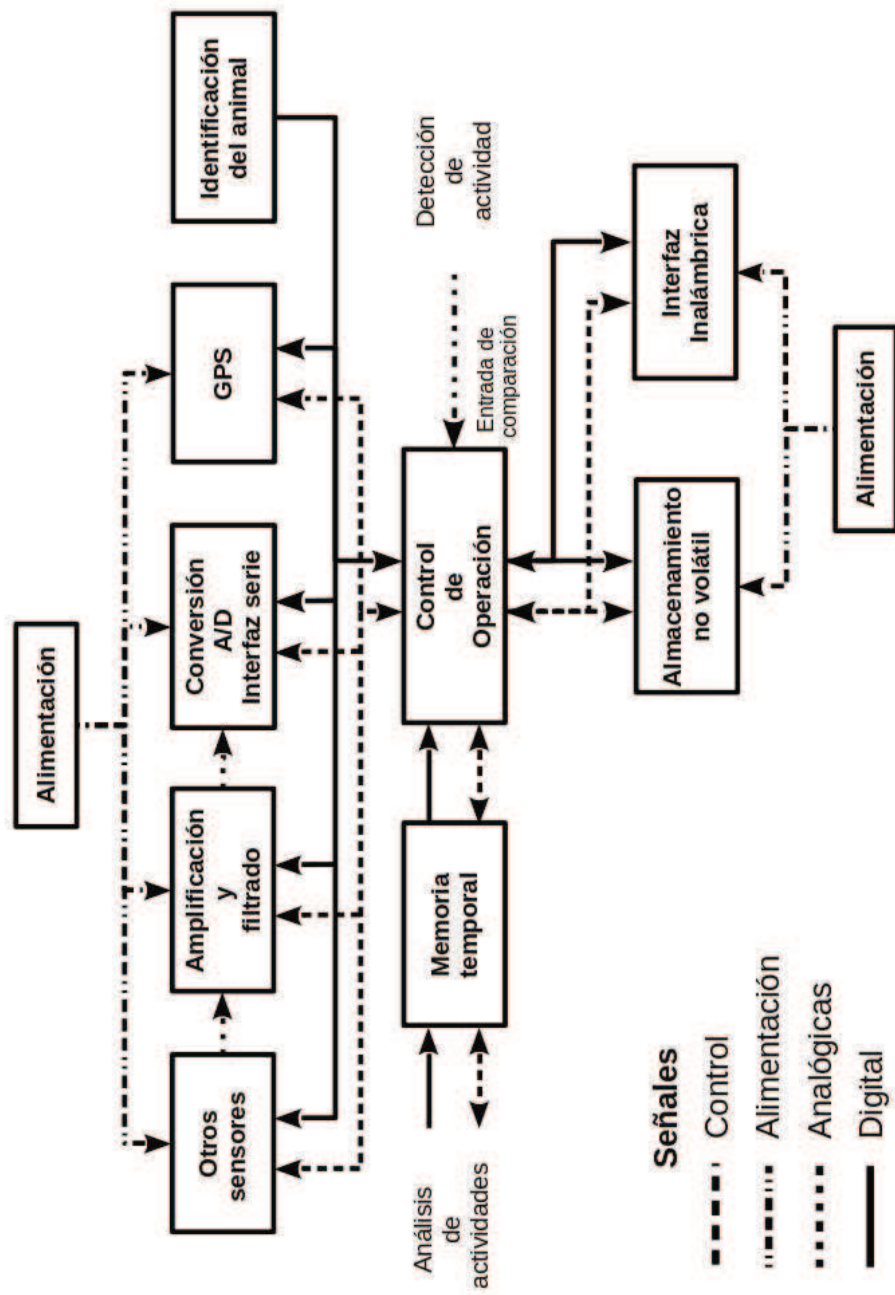


Figura 12

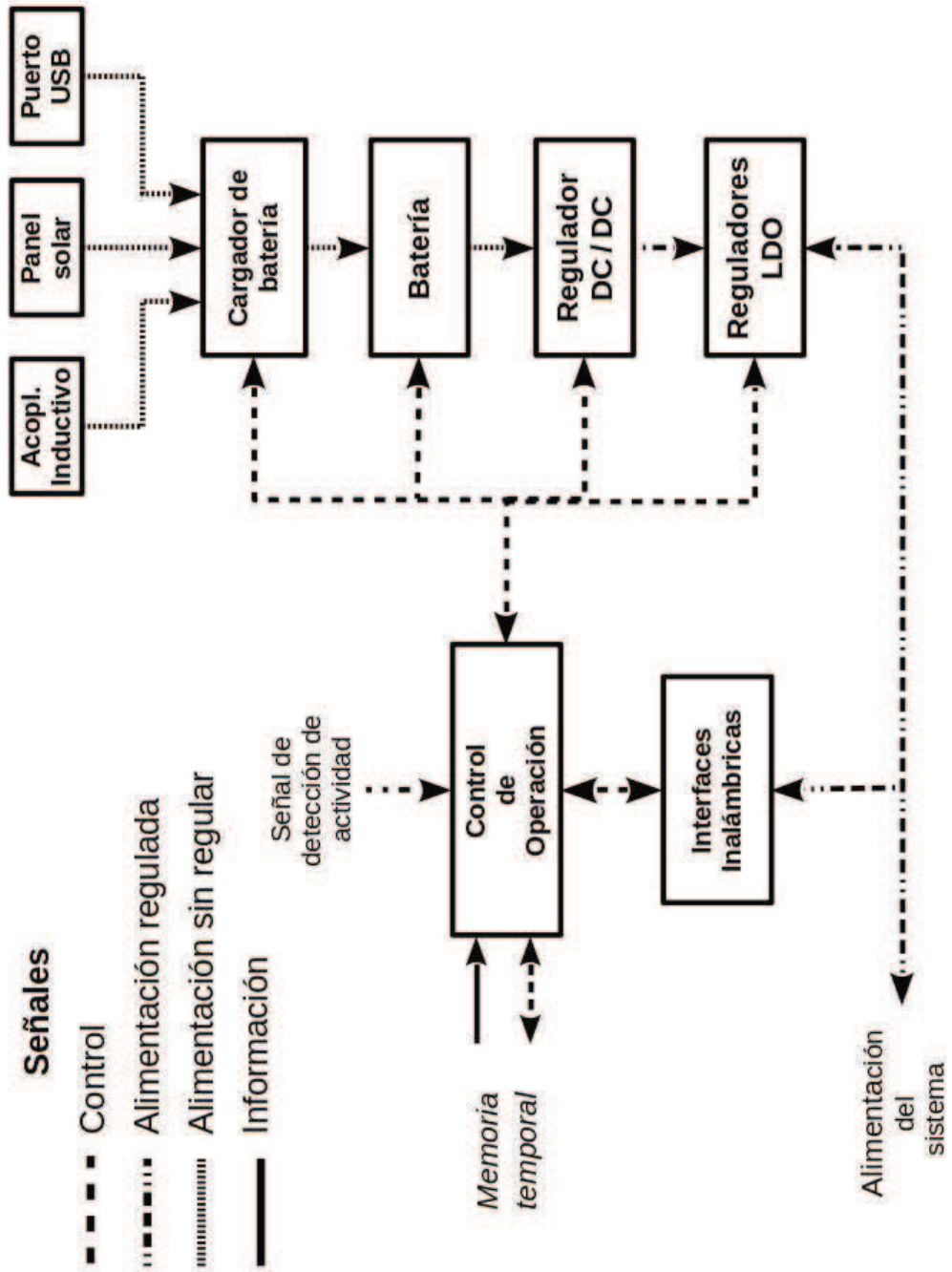


Figura 13

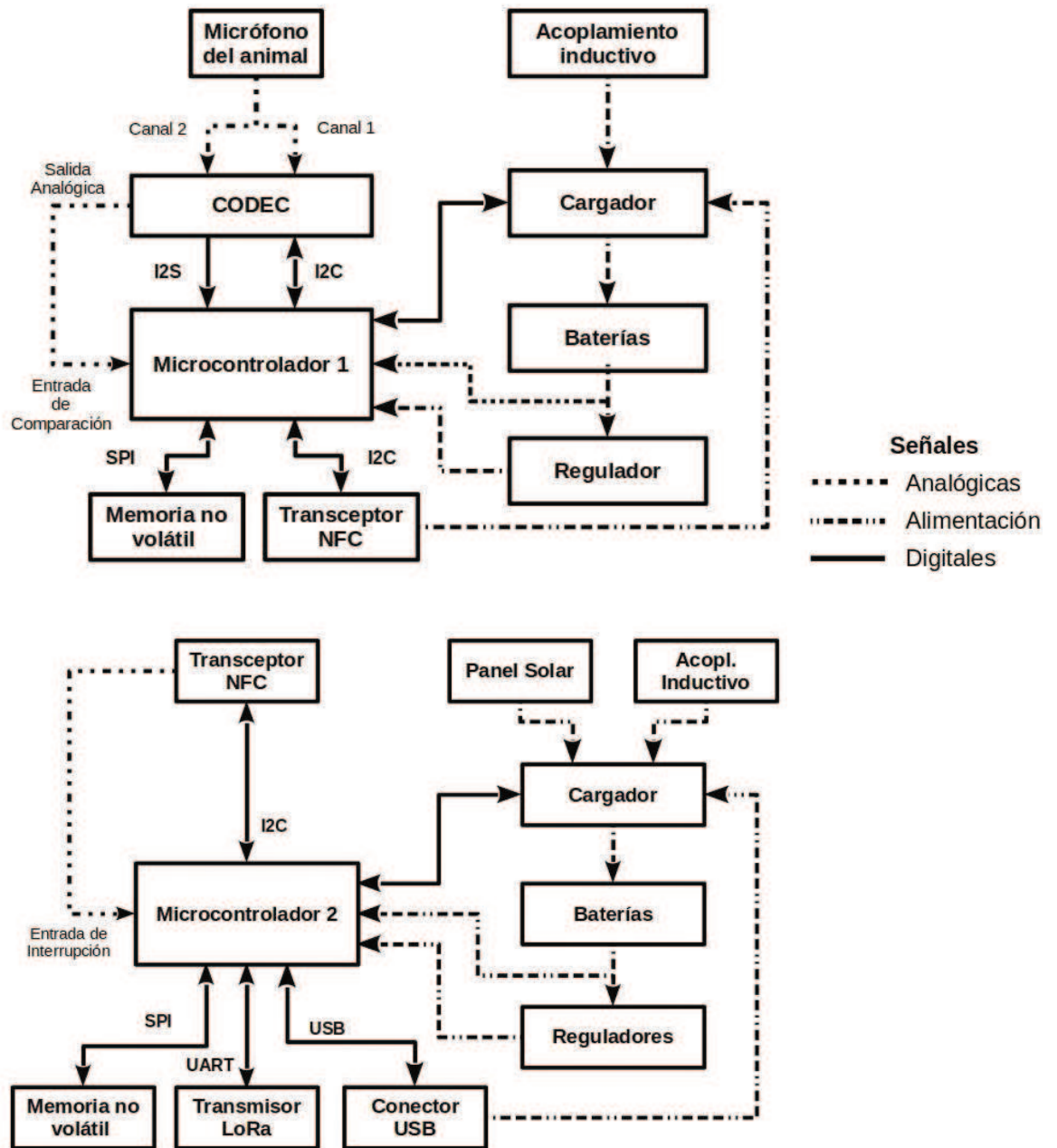


Figura 14

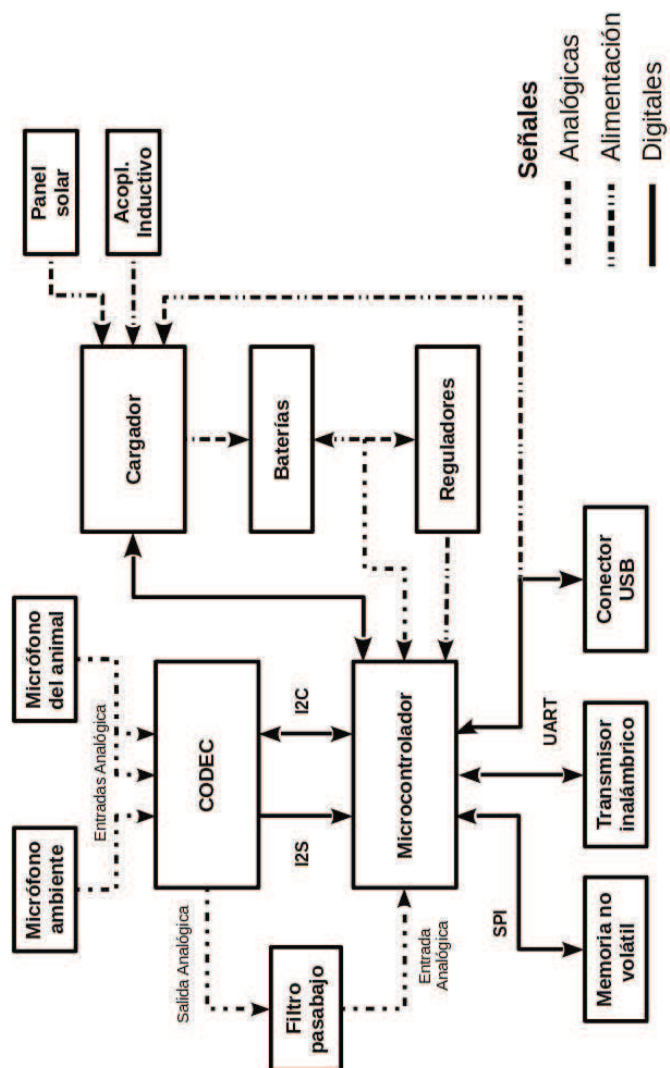


Figura 15

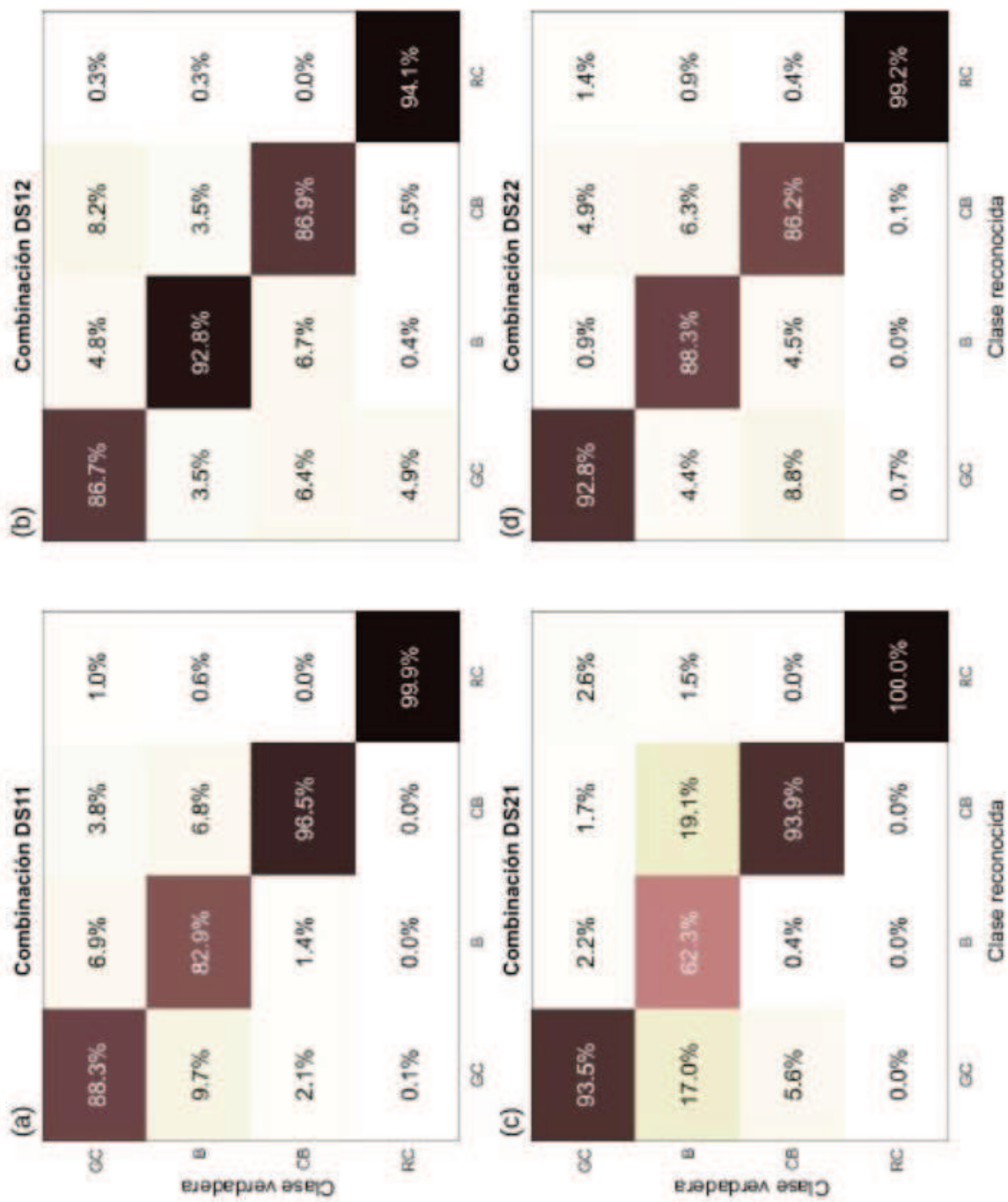


Figura 16

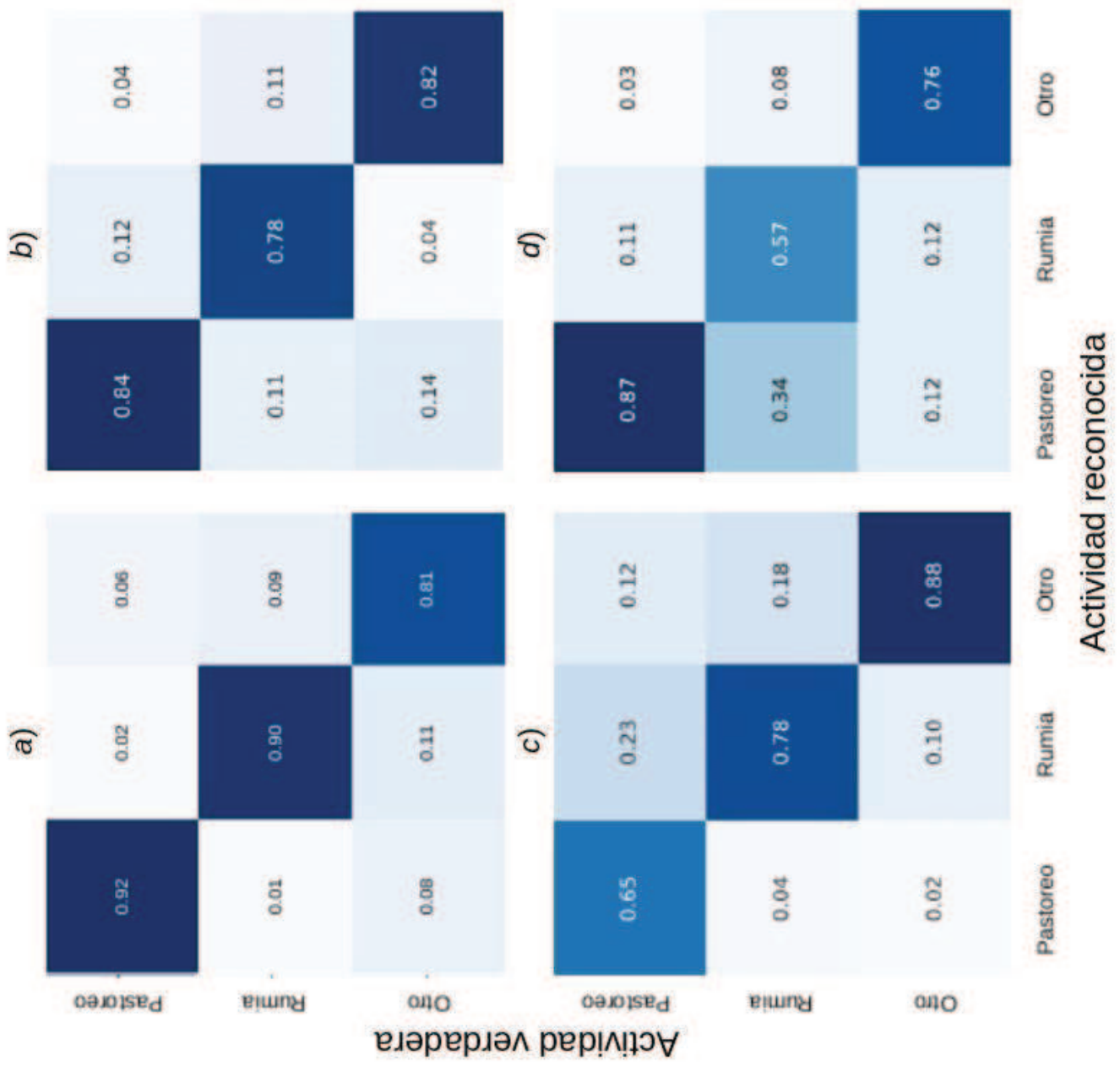


Figura 17

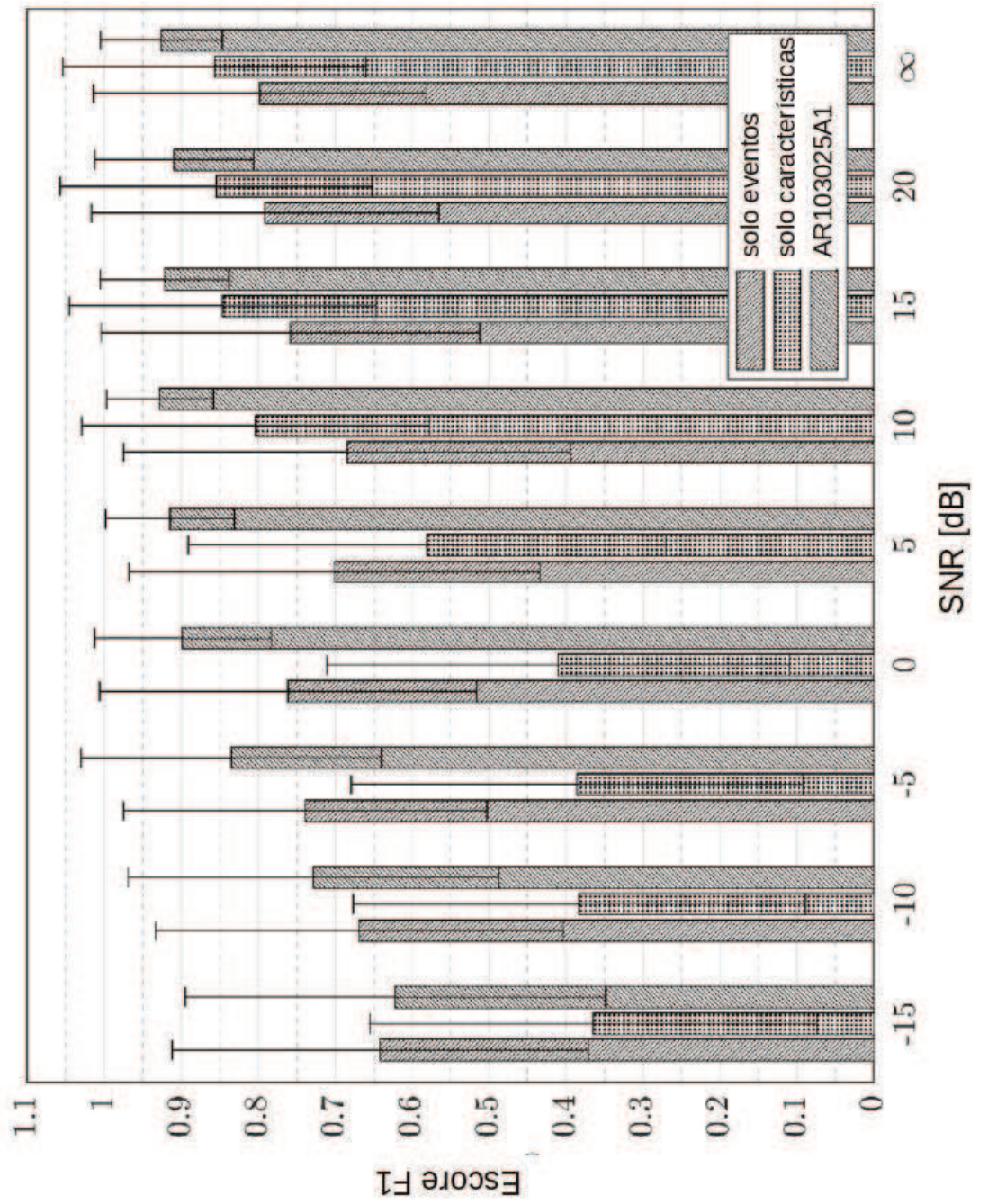


Figura 18

Resumen

Un método y un dispositivo para el monitoreo, cuantificación y evaluación de actividades de pastoreo y rumia a partir de la detección y cuantificación de los eventos de masticación y un dispositivo para realizar dicho procedimiento que solo utiliza el sonido producido por los animales al alimentarse.

Bibliografía

- [1] Daniel Berckmans. “Precision livestock farming technologies for welfare management in intensive livestock systems”. In: *Scientific and Technical Review of the Office International des Epizooties* 33.1 (2014), pp. 189–196.
- [2] National Research Council. *Nutrient Requirements of Dairy Cattle: Seventh Revised Edition, 2001*. Washington, DC: The National Academies Press, 2001. ISBN: 978-0-309-06997-7. DOI: 10.17226/9825. URL: <https://www.nap.edu/catalog/9825/nutrient-requirements-of-dairy-cattle-seventh-revised-edition-2001>.
- [3] J D Leaver. “Milk production from grazed temperate grassland”. en. In: *J. Dairy Res.* 52.2 (May 1985), pp. 313–344.
- [4] J L De Boever et al. “Chewing activity of ruminants as a measure of physical structure—A review of factors affecting it”. In: *Animal Feed Science and Technology* 27.4 (1990), pp. 281–291.
- [5] Julio Ricardo Galli et al. “Monitoring and assessment of ingestive chewing sounds for prediction of herbage intake rate in grazing cattle”. In: *animal* 12.5 (2018), pp. 973–982.
- [6] Julio R Galli et al. “Acoustic monitoring of short-term ingestive behavior and intake in grazing sheep”. In: *Livestock Science* 140.1-3 (2011), pp. 32–41.
- [7] Mette S Herskin, Lene Munksgaard, and Jan Ladewig. “Effects of acute stressors on nociception, adrenocortical responses and behavior of dairy cows”. In: *Physiol. Behav.* 83.3 (2004), pp. 411–420.
- [8] Daniel J Bristow and David S Holmes. “Cortisol levels and anxiety-related behaviors in cattle”. In: *Physiol. Behav.* 90.4 (2007), pp. 626–628.
- [9] J G Welch. “Rumination, Particle Size and Passage from the Rumen”. In: *J. Anim. Sci.* 54.4 (1982), p. 885.
- [10] M Hirata et al. “The effects of recording interval on the estimation of grazing behavior of cattle in a daytime grazing system”. In: *Asian-Australasian Journal of Animal Sciences* 15.5 (2002), pp. 745–750.

-
- [11] PC Aikman, CK Reynolds, and DE Beever. “Diet digestibility, rate of passage, and eating and rumination behavior of Jersey and Holstein cows”. In: *Journal of dairy science* 91.3 (2008), pp. 1103–1114.
- [12] C Pahl et al. “Feeding characteristics and rumination time of dairy cows around estrus”. In: *Journal of Dairy Science* 98.1 (2015), pp. 148–154.
- [13] Derek W Bailey et al. “Mechanisms that result in large herbivore grazing distribution patterns”. In: *Journal of Range Management* 49.5 (1996), pp. 386–400.
- [14] Robert J Kilgour. “In pursuit of “normal”: A review of the behaviour of cattle at pasture”. In: *Applied Animal Behaviour Science* 138.1-2 (2012), pp. 1–11.
- [15] Oscar Balocchi, Rubén Pulido, and Javier Fernández. “Comportamiento de vacas lecheras en pastoreo con y sin suplementación con concentrado”. In: *Agricultura Técnica* 62.1 (2002), pp. 87–98.
- [16] Clive Phillips. *Cattle behaviour and welfare*. John Wiley & Sons, 2008.
- [17] John Hodgson et al. *Grazing management. Science into practice*. Longman Group UK Ltd., 1990.
- [18] John Hodgson and Andrew W Illius. “The ecology and management of grazing systems”. In: (1998).
- [19] Julio Kuhn Da Trindade et al. “Potential of an acoustic method for quantifying the activities of grazing cattle.” In: *Pesquisa Agropecuaria Brasileira* 46.8 (2011), pp. 965–968.
- [20] Marcelo A Benvenuti et al. “Defoliation patterns and their implications for the management of vegetative tropical pastures to control intake and diet quality by cattle”. In: *Grass and Forage Science* 71.3 (2016), pp. 424–436.
- [21] Julio R Galli et al. “Discriminative power of acoustic features for jaw movement classification in cattle and sheep”. In: *Bioacoustics* 29.5 (2020), pp. 602–616.
- [22] EA Laca, ED Ungar, and MW Demment. “Mechanisms of handling time and intake rate of a large mammalian grazer”. In: *Applied Animal Behaviour Science* 39.1 (1994), pp. 3–19.
- [23] Eugene David Ungar and Steven Mark Rutter. “Classifying cattle jaw movements: Comparing IGER Behaviour Recorder and acoustic techniques”. In: *Appl. Anim. Behav. Sci.* 98.1 (June 2006), pp. 11–27.
- [24] Eugene David Ungar et al. “The implications of compound chew–bite jaw movements for bite rate in grazing cattle”. In: *Applied animal behaviour science* 98.3-4 (2006), pp. 183–195.

-
- [25] F J Cordova, Joe D Wallace, and Rex D Pieper. “Forage Intake by Grazing Livestock: A Review”. In: *J. Range Manage.* 31.6 (1978), pp. 430–438.
- [26] ARM Chambers, J Hodgson, and JA Milne. “The development and use of equipment for the automatic recording of ingestive behaviour in sheep and cattle”. In: *Grass and Forage Science* 36.2 (1981), pp. 97–105.
- [27] Diego Humberto Milone et al. “Computational method for segmentation and classification of ingestive sounds in sheep”. In: *Computers and Electronics in Agriculture* 65.2 (2009), pp. 228–237.
- [28] J-M Luginbuhl et al. “A simple electronic device and computer interface system for monitoring chewing behavior of stall-fed ruminant animals”. en. In: *Journal of dairy science* 70.6 (1987), pp. 1307–1312.
- [29] Florian Leiber et al. “Intake estimation in dairy cows fed roughage-based diets: An approach based on chewing behaviour measurements”. In: *Applied Animal Behaviour Science* 185 (2016), pp. 9–14.
- [30] J Werner et al. “Evaluation of the RumiWatchSystem for measuring grazing behaviour of cows”. In: *Journal of Neuroscience Methods* 300 (2018), pp. 138–146.
- [31] Nils Zehner et al. “System specification and validation of a noseband pressure sensor for measurement of ruminating and eating behavior in stable-fed cows”. In: *Computers and Electronics in Agriculture* 136 (2017), pp. 31–41.
- [32] Franz Nydegger, Lorenz Gyax, and Wendelin Egli. “Automatic measurement of rumination and feeding activity using a pressure sensor”. In: *International Conference on Agricultural Engineering-AgEng 2010: towards environmental technologies, 6-8 September 2010, Clermont-Ferrand, France.* 2010, pp. 1–8.
- [33] S M Rutter. “Graze: a program to analyze recordings of the jaw movements of ruminants”. en. In: *Behav. Res. Methods Instrum. Comput.* 32.1 (Feb. 2000), pp. 86–92.
- [34] S M Rutter, R A Champion, and P D Penning. “An automatic system to record foraging behaviour in free-ranging ruminants”. In: *Appl. Anim. Behav. Sci.* 54.2-3 (1997), pp. 185–195.
- [35] M Rombach et al. “Herbage dry matter intake estimation of grazing dairy cows based on animal, behavioral, environmental, and feed variables”. In: *Journal of Dairy Science* 102.4 (2019), pp. 2985–2999.
- [36] C Arcidiacono et al. “Development of a threshold-based classifier for real-time recognition of cow feeding and standing behavioural activities from accelerometer data”. In: *Comput. Electron. Agric.* 134 (2017), pp. 124–134.

-
- [37] L A González et al. “Behavioral classification of data from collars containing motion sensors in grazing cattle”. In: *Comput. Electron. Agric.* 110 (2015), pp. 91–102.
- [38] V Giovanetti et al. “Automatic classification system for grazing, ruminating and resting behaviour of dairy sheep using a tri-axial accelerometer”. In: *Livest. Sci.* 196 (2017), pp. 42–48.
- [39] Paula Martiskainen et al. “Cow behaviour pattern recognition using a three-dimensional accelerometer and support vector machines”. In: *Appl. Anim. Behav. Sci.* 119.1–2 (2009), pp. 32–38.
- [40] Andriamasinoro Lalaina Herinaina Andriamandroso et al. “Development of an open-source algorithm based on inertial measurement units (IMU) of a smartphone to detect cattle grass intake and ruminating behaviors”. In: *Comput. Electron. Agric.* 139 (2017), pp. 126–137.
- [41] P L Greenwood et al. “Use of sensor-determined behaviours to develop algorithms for pasture intake by individual grazing cattle”. In: *Crop and Pasture Science* 68.12 (2017), pp. 1091–1099.
- [42] Daniel Smith et al. “Behavior classification of cows fitted with motion collars: Decomposing multi-class classification into a set of binary problems”. In: *Computers and electronics in agriculture* 131 (2016), pp. 40–50.
- [43] Simona Büchel and Albert Sundrum. “Evaluation of a new system for measuring feeding behavior of dairy cows”. In: *Computers and Electronics in Agriculture* 108 (2014), pp. 12–16.
- [44] Daniel P Campos et al. “Short-term fibre intake estimation in goats using surface electromyography of the masseter muscle”. In: *Biosystems Engineering* 183 (2019), pp. 209–220.
- [45] Guanghai Duan et al. “Short-term feeding behaviour sound classification method for sheep using LSTM networks”. In: *International Journal of Agricultural and Biological Engineering* 14.2 (2021), pp. 43–54.
- [46] Emilio A Laca, Michiel F WallisDeVries, et al. “Acoustic measurement of intake and grazing behaviour of cattle”. In: *Grass and Forage Science* 55.2 (2000), pp. 97–104.
- [47] Julio R Galli et al. “Acoustic monitoring of chewing and intake of fresh and dry forages in steers”. In: *Animal Feed Science and Technology* 128.1-2 (2006), pp. 14–30.

-
- [48] Philip U Alkon, Yosef Cohen, and Peter A Jordan. “Towards an acoustic biotelemetry system for animal behavior studies”. In: *The Journal of Wildlife Management* (1989), pp. 658–662.
- [49] Wendy M Griffiths et al. “A video and acoustic methodology to map bite placement at the patch scale”. In: *Applied Animal Behaviour Science* 98.3-4 (2006), pp. 196–215.
- [50] E A Laca et al. “An integrated methodology for studying short-term grazing behaviour of cattle”. In: *Grass and forage science* 47.1 (1992), pp. 81–90.
- [51] Rémy Delagarde, Jean-Pierre Caudal, and Jean-Louis Peyraud. “Development of an automatic bitemeter for grazing cattle”. In: *Ann. Zootech.* 48.5 (1999), pp. 329–339.
- [52] Diego H Milone et al. “Automatic recognition of ingestive sounds of cattle based on hidden Markov models”. In: *Computers and electronics in agriculture* 87 (2012), pp. 51–55.
- [53] William M Clapham et al. “Acoustic monitoring system to quantify ingestive behavior of free-grazing cattle”. In: *Computers and Electronics in Agriculture* 76.1 (2011), pp. 96–104.
- [54] Shilo Navon et al. “Automatic recognition of jaw movements in free-ranging cattle, goats and sheep, using acoustic monitoring”. In: *Biosystems Engineering* 114.4 (2013), pp. 474–483.
- [55] José O Chelotti et al. “A real-time algorithm for acoustic monitoring of ingestive behavior of grazing cattle”. In: *Computers and Electronics in Agriculture* 127 (2016), pp. 64–75.
- [56] Nestor N Deniz et al. “Embedded system for real-time monitoring of foraging behavior of grazing cattle using acoustic signals”. In: *Computers and electronics in agriculture* 138 (2017), pp. 167–174.
- [57] José O Chelotti et al. “A pattern recognition approach for detecting and classifying jaw movements in grazing cattle”. In: *Computers and Electronics in Agriculture* 145 (2018), pp. 83–91.
- [58] Jorge A Vázquez Diosdado et al. “Classification of behaviour in housed dairy cows using an accelerometer-based activity monitoring system”. In: *Animal Biotelemetry* 3.1 (2015), pp. 1–14.
- [59] Adolfo A Rayas-Amor et al. “Triaxial accelerometers for recording grazing and ruminating time in dairy cows: An alternative to visual observations”. In: *Journal of Veterinary Behavior* 20 (2017), pp. 102–108.

-
- [60] Said Benaïssa et al. “On the use of on-cow accelerometers for the classification of behaviours in dairy barns”. In: *Research in veterinary science* 125 (2019), pp. 425–433.
- [61] Said Benaïssa et al. “Classification of ingestive-related cow behaviours using RumiWatch halter and neck-mounted accelerometers”. In: *Applied Animal Behaviour Science* 211 (2019), pp. 9–16.
- [62] Weizheng Shen et al. “Automatic recognition of ingestive-related behaviors of dairy cows based on triaxial acceleration”. In: *Information Processing in Agriculture* 7.3 (2020), pp. 427–443.
- [63] Avshalom Bar-Shalom. *Method and system for monitoring physiological conditions of, and/or suitability of animal feed for ruminant animals*. US Patent 7,350,481. Apr. 2008.
- [64] C Goldhawk, K Schwartzkopf-Genswein, and K A Beauchemin. “Validation of rumination collars for beef cattle”. In: *J. Anim. Sci.* 91.6 (2013), pp. 2858–2862.
- [65] MV Byskov et al. “Recording rumination time by a rumination monitoring system in Jersey heifers fed grass/clover silage and hay at three feeding levels”. In: *Journal of animal science* 92.3 (2014), pp. 1110–1118.
- [66] P Gregorini et al. “A note on rumination behavior of dairy cows under intensive grazing systems”. In: *Livestock Science* 158.1-3 (2013), pp. 151–156.
- [67] Kathrin Schirmann et al. “Validation of a system for monitoring rumination in dairy cows”. In: *Journal of Dairy Science* 92.12 (2009), pp. 6052–6055.
- [68] Sebastián R Vanrell et al. “A regularity-based algorithm for identifying grazing and rumination bouts from acoustic signals in grazing cattle”. In: *Computers and Electronics in Agriculture* 151 (2018), pp. 392–402.
- [69] Brahim Achour et al. “Image analysis for individual identification and feeding behaviour monitoring of dairy cows based on Convolutional Neural Networks (CNN)”. In: *Biosystems Engineering* 198 (2020), pp. 31–49.
- [70] Simona MC Porto et al. “The automatic detection of dairy cow feeding and standing behaviours in free-stall barns by a computer vision-based system”. In: *Biosystems Engineering* 133 (2015), pp. 46–55.
- [71] Xuqiang Yin et al. “Using an EfficientNet-LSTM for the recognition of single cow’s motion behaviours in a complicated environment”. In: *Computers and Electronics in Agriculture* 177 (2020), p. 105707.

-
- [72] Safa Ayadi et al. “Dairy Cow rumination detection: A deep learning approach”. In: *International Workshop on Distributed Computing for Emerging Smart Networks*. Springer. 2020, pp. 123–139.
- [73] Ran Bezen, Yael Edan, and Ilan Halachmi. “Computer vision system for measuring individual cow feed intake using RGB-D camera and deep learning algorithms”. In: *Computers and Electronics in Agriculture* 172 (2020), p. 105345.
- [74] Guoming Li et al. “Classifying Ingestive Behavior of Dairy Cows via Automatic Sound Recognition”. In: *Sensors* 21.15 (2021). ISSN: 1424-8220. DOI: 10.3390/s21155231. URL: <https://www.mdpi.com/1424-8220/21/15/5231>.
- [75] Kui Wang et al. “Identification and classification for sheep foraging behavior based on acoustic signal and deep learning”. In: *Computers and Electronics in Agriculture* 187 (2021), p. 106275.
- [76] Leonardo L Giovanini et al. *A method and a device for detecting, classifying and quantifying in real time the feeding activity of ruminant animals*. WO 2017/125805 A3 Patent. Dec. 2016.
- [77] Richard O Duda, Peter E Hart, and David G Stork. *Pattern Classification*. en. John Wiley & Sons, Nov. 2012.
- [78] Philip H Swain and Hans Hauska. “The decision tree classifier: Design and potential”. In: *IEEE Transactions on Geoscience Electronics* 15.3 (1977), pp. 142–147.
- [79] Leo Breiman et al. *Classification and Regression Trees*. en. Chapman and Hall/CRC, Jan. 1984.
- [80] J Ross Quinlan. *C4. 5: programs for machine learning*. en. Elsevier, June 2014.
- [81] Christopher M Bishop. *Pattern Recognition and Machine Learning*. en. Springer, Aug. 2006.
- [82] Marina Sokolova and Guy Lapalme. “A systematic analysis of performance measures for classification tasks”. In: *Information Processing & Management* 45.4 (2009), pp. 427–437.
- [83] Davide Ballabio, Francesca Grisoni, and Roberto Todeschini. “Multivariate comparison of classification performance measures”. In: *Chemometrics and Intelligent Laboratory Systems* 174 (2018), pp. 33–44.
- [84] Jamie A Ward, Paul Lukowicz, and Gerhard Tröster. “Evaluating performance in continuous context recognition using event-driven error characterisation”. In: *International Symposium on Location-and Context-Awareness*. Springer. 2006, pp. 239–255.

-
- [85] Laurens van der Maaten and Geoffrey Hinton. “Visualizing data using t-SNE”. In: *Journal of machine learning research* 9.Nov (2008), pp. 2579–2605.
- [86] Frank Wilcoxon. “Individual Comparisons by Ranking Methods”. In: *Biometrics Bulletin* 1.6 (1945), p. 80.
- [87] A E Stone. “Symposium review: The most important factors affecting adoption of precision dairy monitoring technologies”. In: *Journal of Dairy Science* 103.6 (2020), pp. 5740–5745.
- [88] Moritz Weißbrich, Javier Andrés Moreno-Medina, and Guillermo Payá-Vayá. “Using Genetic Algorithms to Optimize the Instruction-Set Encoding on Processor Cores”. In: *2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST)*. 2021, pp. 1–6. DOI: 10.1109/MOCAST52088.2021.9493406.
- [89] Henk Corporaal. *Microprocessor architectures - from VLIW to TTA*. en. John Wiley & Sons, Dec. 1997, p. 428. ISBN: 978-0-471-97157-3.
- [90] Pekka Jääskeläinen et al. “HW/SW Co-design Toolset for Customization of Exposed Datapath Processors”. In: *Computing Platforms for Software-Defined Radio*. Ed. by Waqar Hussain et al. Springer International Publishing, 2017, pp. 147–164. ISBN: 978-3-319-49679-5. DOI: 10.1007/978-3-319-49679-5_8. URL: https://doi.org/10.1007/978-3-319-49679-5_8.

Doctorado en Ingeniería
mención inteligencia computacional, señales y sistemas

Título de la obra:

**Desarrollo de un sensor inteligente
para el monitoreo continuo de
animales en pastoreo para
ganadería de precisión**

Autor: Luciano Sebastián Martínez Rau

Lugar: Santa Fe, Argentina

Palabras Claves:

Ganadería de precisión,
Procesamiento de señales,
Reconocimiento de patrones,
Aprendizaje maquina,
Monitoreo acústico,
Sistemas embebidos,
Sensor acústico,
FPGA,
Ejecución en tiempo real.