



13-14 November, 2025

MACHINE LEARNING BASED ANALYSIS OF DAIRY COW VOCAL BEHAVIOUR

Madalina MINCU-IORGA, Dinu GAVOJDIAN*

*Cattle Production Systems Laboratory, Research and Development Institute for
Bovine, Balotesti, Romania*

*Corresponding author: gavojdian_dinu@animalsci-tm.ro

Abstract:

• Introduction

Understanding animal vocalizations through multi-source data fusion is key to evaluating emotional states and improving welfare in precision livestock systems. This study examined dairy cattle contact calls using integrated acoustic pattern extraction, sequence analysis and feature profiling.

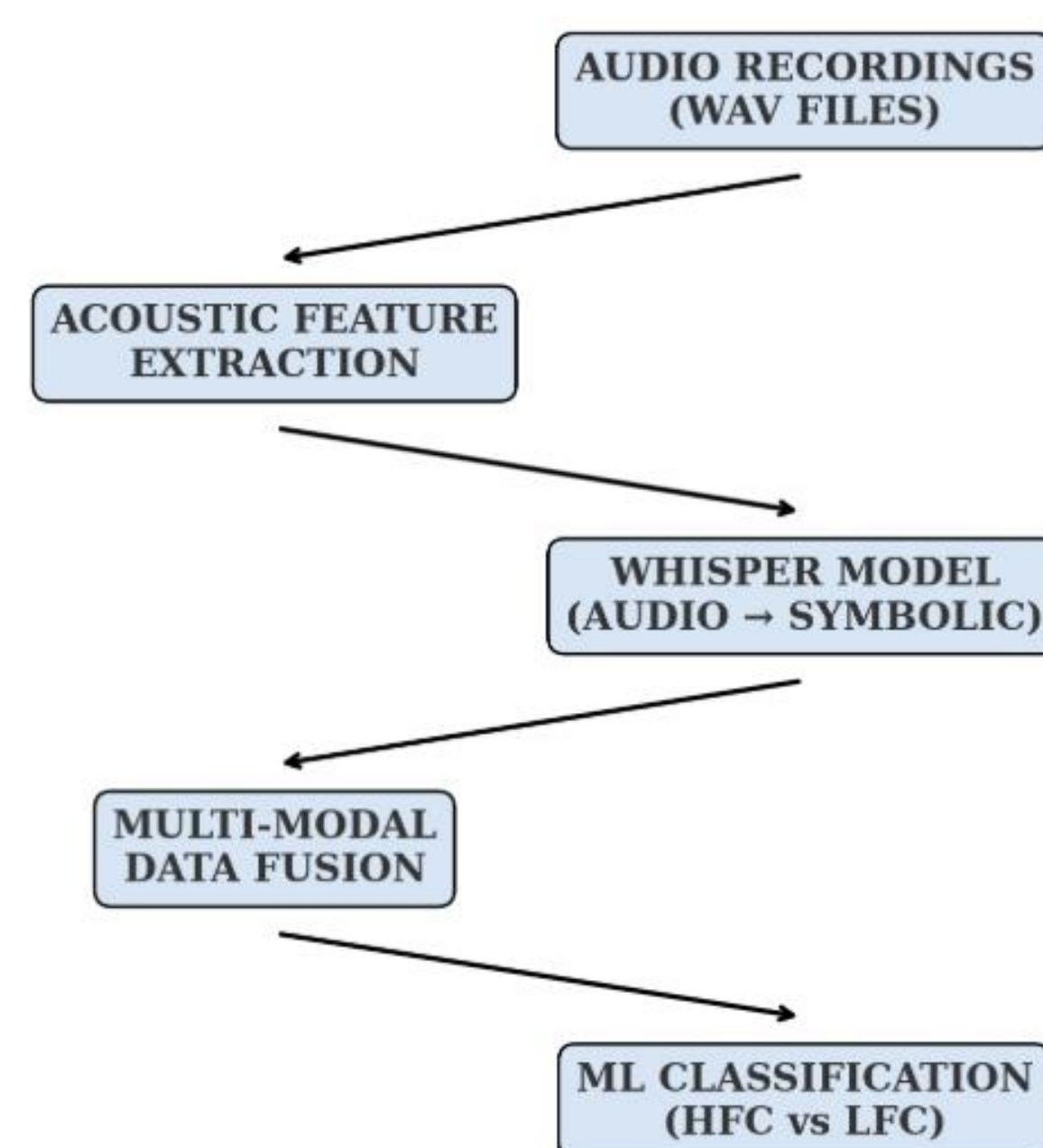


Figure 1: Methodological Workflow for Dairy Cattle Vocalization Classification

Material and method

The Natural Language Processing-based WHISPER model was employed to convert audio recordings of cow vocalizations into acoustic-to-text sequences. We categorized vocalizations into high-frequency calls (HFC) putatively associated with arousal or agitation, and low-frequency calls (LFC) indicative of calm or content states. Evaluating acoustic and sequential patterns from vocalizations of 20 lactating cows facilitated differentiation in calling behaviors (Figures 2, 3).

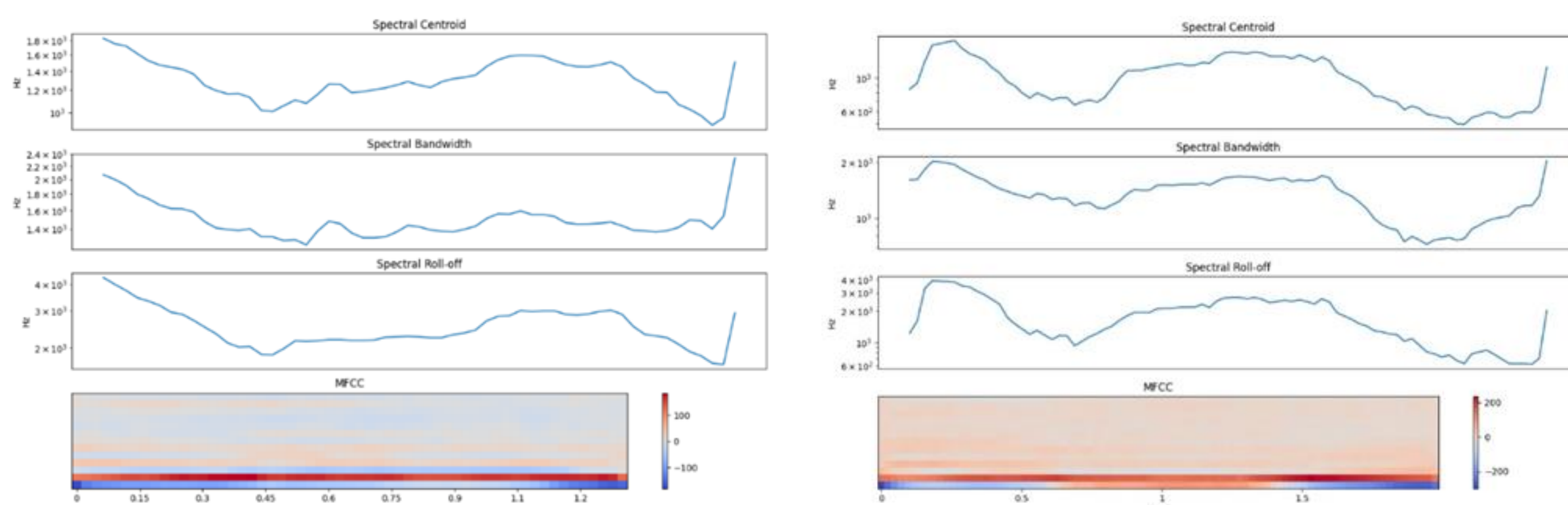


Figure 2: Left side – Spectral analysis of a typical Low-Frequency Call (LFC); Right side – Spectral analysis of a typical High-Frequency Call (HFC)

• Results and discussions

Advanced machine learning methods, including Random Forest, Support Vector Machine (SVM), and Recurrent Neural Networks (RNN), effectively analyzed and fused multi-source acoustic data, resulting in classification accuracies of 97.25% (Random Forest), 98.35% (SVM), and 88.00% (RNN). The F1-scores for identifying arousal were 0.98 (Random Forest) and 0.99 (SVM) (Table 1).

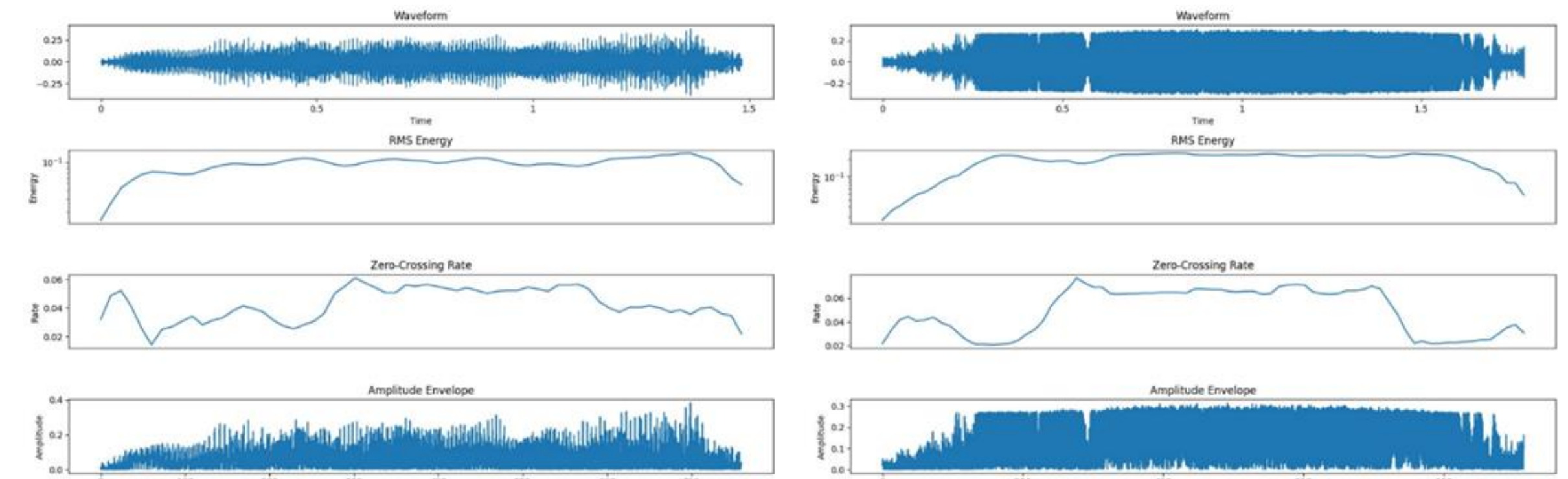


Figure 3: Left side – Amplitude and energy profile of a Low-Frequency Call (LFC), showing lower RMS energy and smoother transitions supporting a calm vocal classification; Right side – Amplitude and energy profile of a High-Frequency Call (HFC), with higher RMS energy and variability reflecting vocal strain and emotional intensity.

Table 1: Performance Evaluation Result of Random Forest, Support Vector Machine and RNN

Model	Class	Precision	Recall	F1-Score	Accuracy
Random Forest	HFC	0.98	0.99	0.98	0.9725
	LFC	0.95	0.93	0.94	
	Macro Avg	0.97	0.96	0.96	
	Weighted Avg	0.97	0.97	0.97	
SVM	HFC	0.99	0.99	0.99	0.9835
	LFC	0.98	0.96	0.97	
	Macro Avg	0.98	0.97	0.98	
	Weighted Avg	0.98	0.98	0.98	
RNN	HFC	0.89	0.97	0.93	0.88
	LFC	0.88	0.62	0.73	
	Macro Avg	0.88	0.80	0.83	
	Weighted Avg	0.88	0.88	0.88	

• Conclusions

To our knowledge, this is the first study that combines acoustic and symbolic data to analyze vocalizations of dairy cows in order to develop a non-invasive classification model using multi-modal machine learning.