Addressing purpose and subjective data labelling challenges in automated lameness detection for cattle with machine learning and micro-Doppler radar

Konstantina Linardopoulou[1], Lorenzo Viora[1], Julien Le Kernec[1], Nicholas Jonsson[1]
[1]University of Glasgow, Glasgow, Scotland, UK
Corresponding author: nicholas.jonsson@glasgow.ac.uk

Automated lameness detection in cattle using micro-Doppler radar technology has shown proof of concept. To develop it for commercial application it is essential to address specific issues concerning the labelling process and notification objectives for these systems. Labelling subjectivity and the choice of a scoring system are central challenges in the development process, as the effectiveness of AI systems relies on the quality of labels on training data, which is usually generated through human observations and annotations. Labelling of mobility can be subjective due to variations in human assessors' interpretations of lameness severity, leading to inconsistencies in the training dataset. To address this, it is necessary to establish clear and standardised guidelines for assessors, examine inter-assessor agreement, and employ labelling approaches that can reduce variation. The scoring system used to label mobility is an important determinant of the sensitivity and specificity of the automated system. The choice of scoring system must align with the specific objectives of the automated system. For instance, a scoring system optimised for high sensitivity might emphasise early lameness detection with the acceptance of false positives. In contrast, a system designed for high specificity would prioritise minimising false alarms, even if this results in missing some lameness cases. Clearly defining the system's intended purpose and notifying users of its functional requirements is critical to managing expectations and optimising the automated system's performance. To address these challenges for the micro-Doppler system, alternative approaches are being explored. One pivotal shift involves the consideration of unsupervised machine learning (ML) techniques for analysis, departing from the previously employed supervised methods. While unsupervised ML offers the potential to enhance the accuracy of lameness detection, it should be noted that this transition is constrained by the absence of a concrete reference to compare against the results. We are also investigating the application of statistical methods to extract relevant features from micro-Doppler radar data and use them to create individualised reference baselines for each animal, which may offer a basis for comparison across recordings. While this approach shows promise, its effectiveness relies on the ability to identify consistent and informative features and on the assumption that deviations from these baselines will accurately indicate lameness. It is essential to acknowledge that the success of these methods is contingent on several factors, including data quality, feature selection, and the complex nature of cattle behaviour. Ongoing research and validation studies are being investigated to determine the practical feasibility and effectiveness of these strategies in real-world farm settings.