

Feature Representation Learning for Calving Detection of Cows Using Video Frames

Ryosuke Hyodo¹ Teppei Nakano^{1,2} Tetsuji Ogawa¹, ¹Waseda University ² Intelligent Framework Lab

Abstract

Objective: Robust calving sign detection of cows using video frames for farmers' decision making

Approach: System incorporates farmers' decision-making processes into the network.

- ✓ Robustness on a small data
- ✓ Interpretability of reasons for predictions



Background

Contact-type sensor **Non-contact-type sensor**

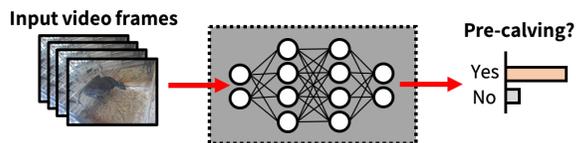


😊 High accuracy

😊 Both cows- & farmers- friendly

😞 Physical & operational burden

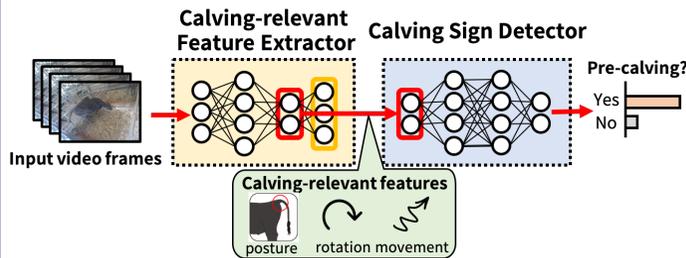
End-to-End System



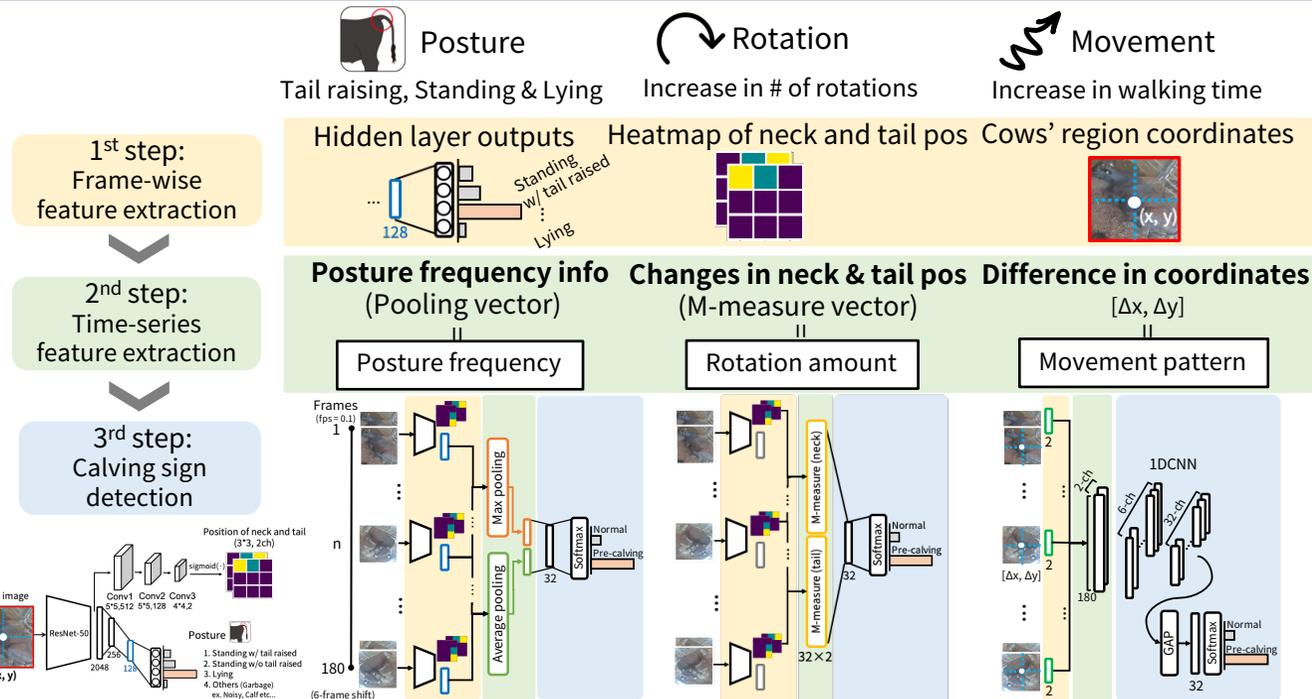
1. Large and well-organized data are necessary.
2. Low interpretability (called **Black-box** system)

Proposed System - Overview

Proposed system incorporates **domain knowledge for farmers' decision-making processes** into the network.



Proposed System



Experiment

Dataset: 15 calving scenes recorded in Kagoshima, Japan.



Systems to compare:

- E2E (End-to-End system without explicit feature extraction about calving signs)
 - Architecture is the same as tandem-posture, with frame-by-frame features derived from ImageNet-trained ResNet-50.
- SS - {posture, rotation, movement} [proposed]

Results & Discussion:

- ✓ Outperformed the E2E system on a small data.
- With explicit feature extraction, the proposed systems suppress obvious **false positives**.



Table. Detection performance on test data

System	AUC	F1-score	Precision	Recall
E2E	0.82	0.73	0.71	0.76
SS - Posture	0.88	0.79	0.77	0.80
SS - Rotation	0.84	0.74	0.80	0.70
SS - Movement	0.86	0.78	0.79	0.76

Fig. Transition of posterior probabilities in one scene of normal state.

