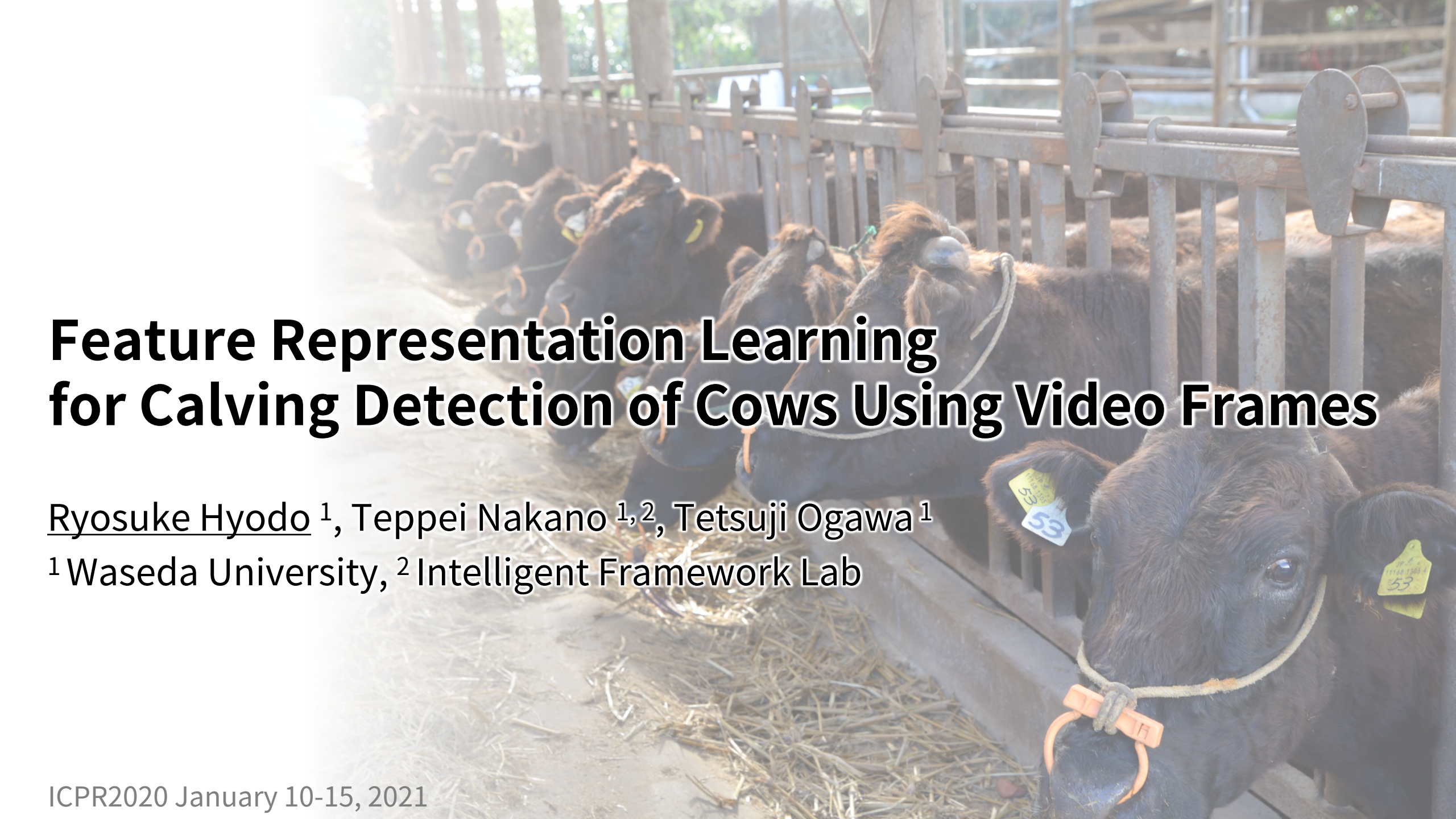


Feature Representation Learning for Calving Detection of Cows Using Video Frames

Ryosuke Hyodo [1], Teppei Nakano [1,2], Tetsuji Ogawa [1]

[1] Waseda University, [2] Intelligent Framework Lab

ICPR2020 January 10-15, 2021

Abstract

Objective:

Robust calving detection of cows using video frames for farmers' decision making.

Approach:

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

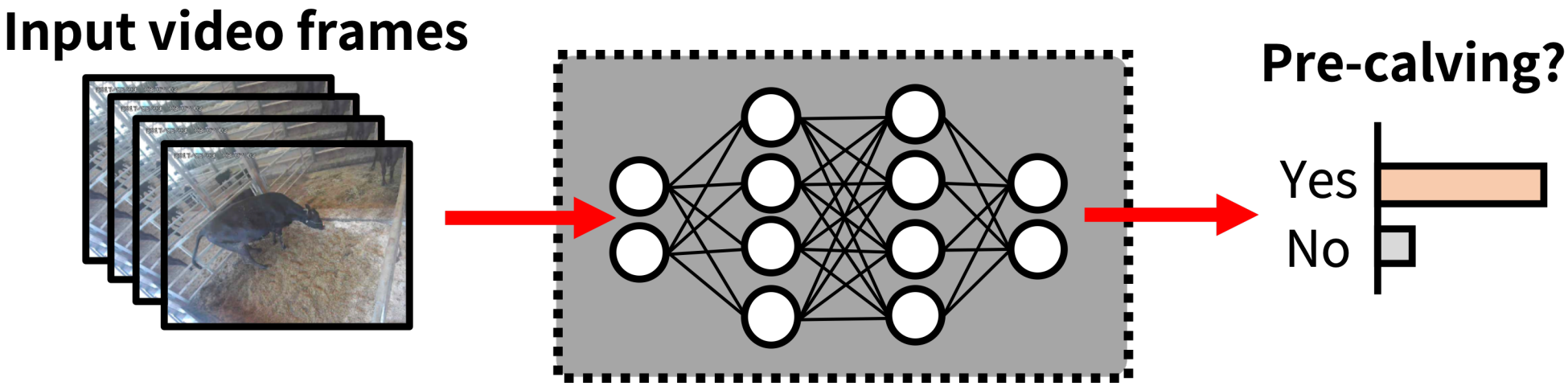
✓ Robustness on a small data

✓ Interpretability of reasons for predictions

Assisting cows during calving is important for preventing **fatal accidents**.

Camera–based detection system ⋯ **Both cows- and farmers-friendly**

End-to-End system

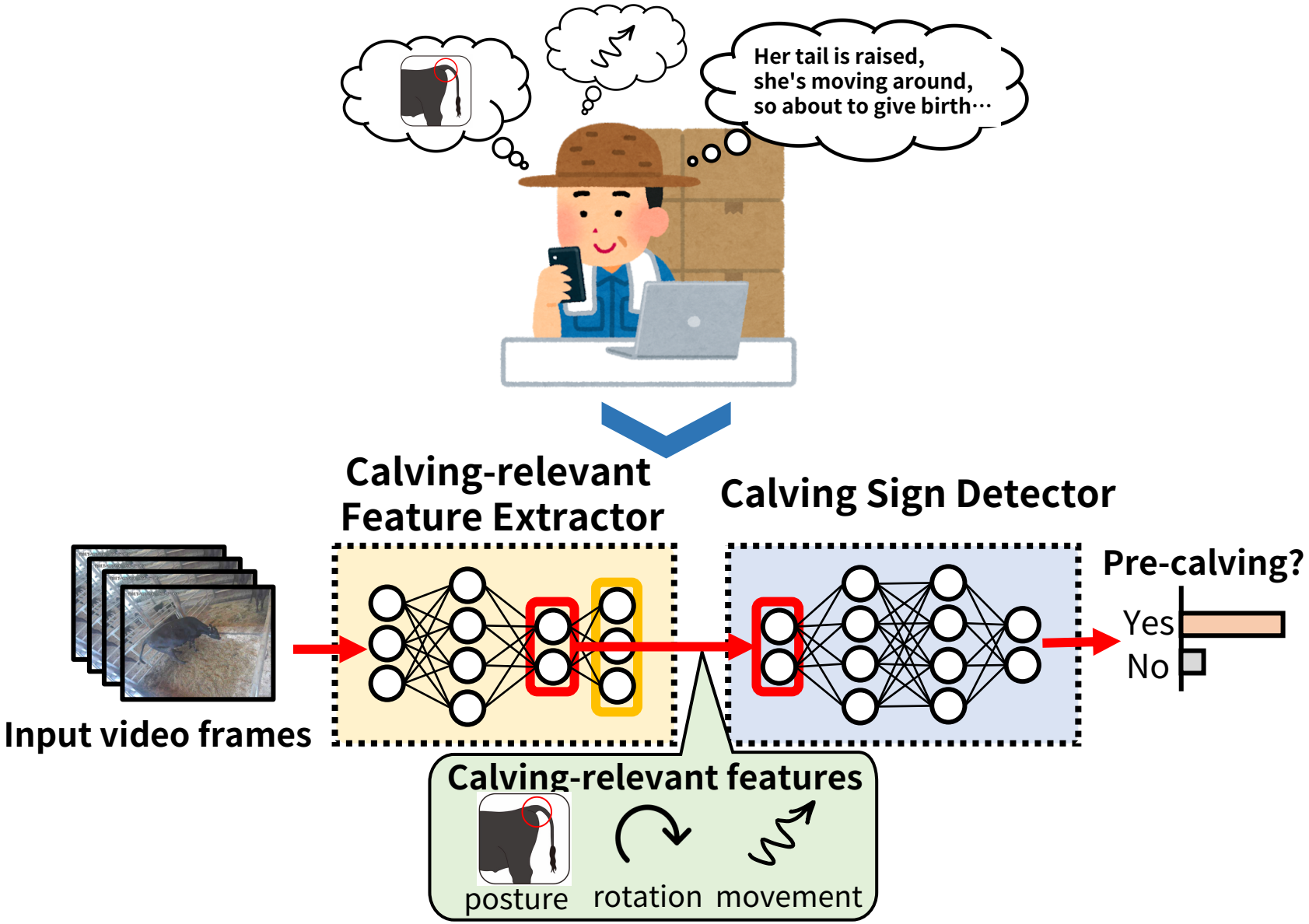


😀 Simple approach that does not require domain knowledge.

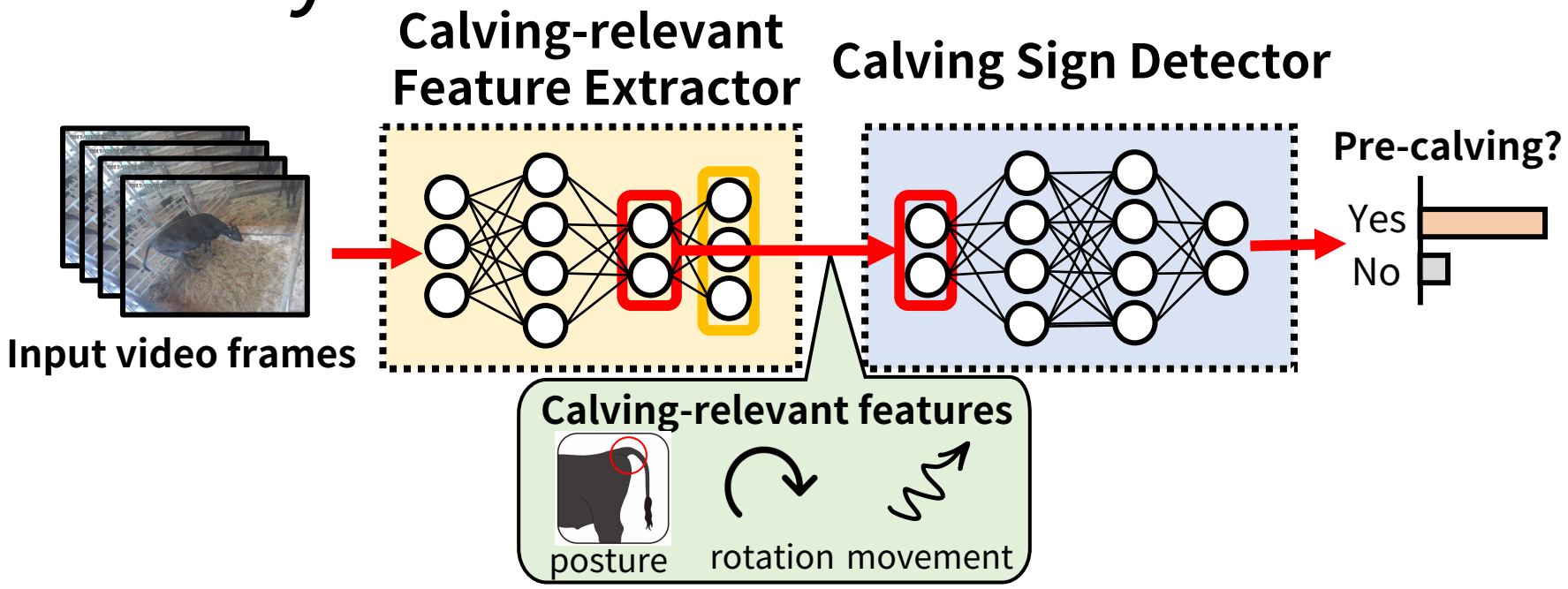
😩
1. Large and well-organized data are necessary.
2. Low interpretability (called **Black-box** system)

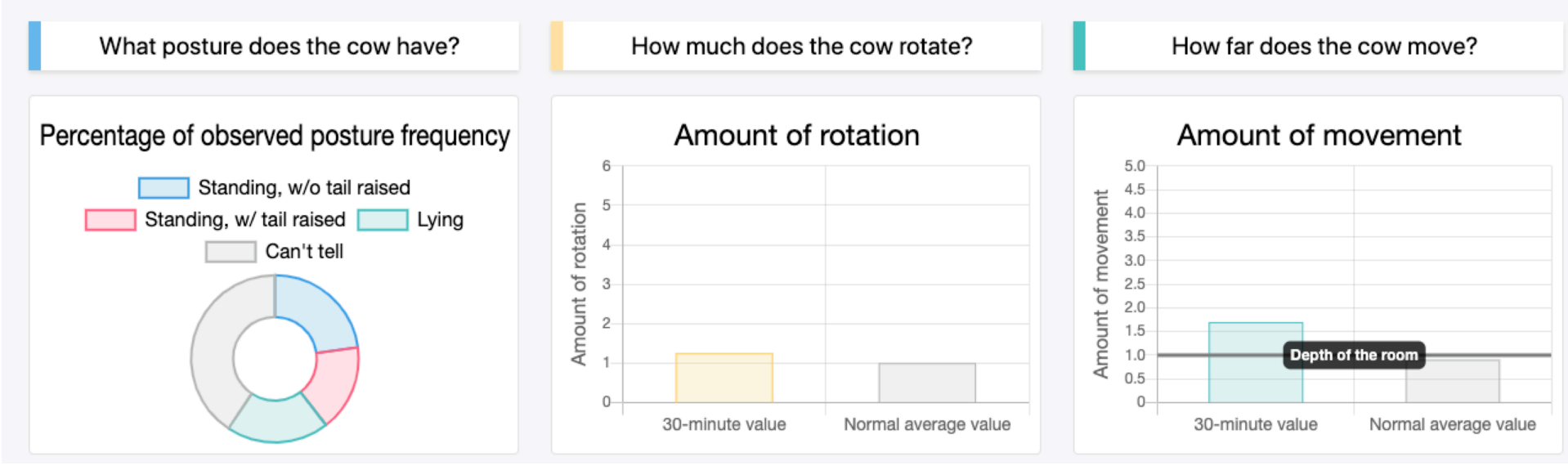

Her tail is raised,
she's moving around,
so about to give birth…

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

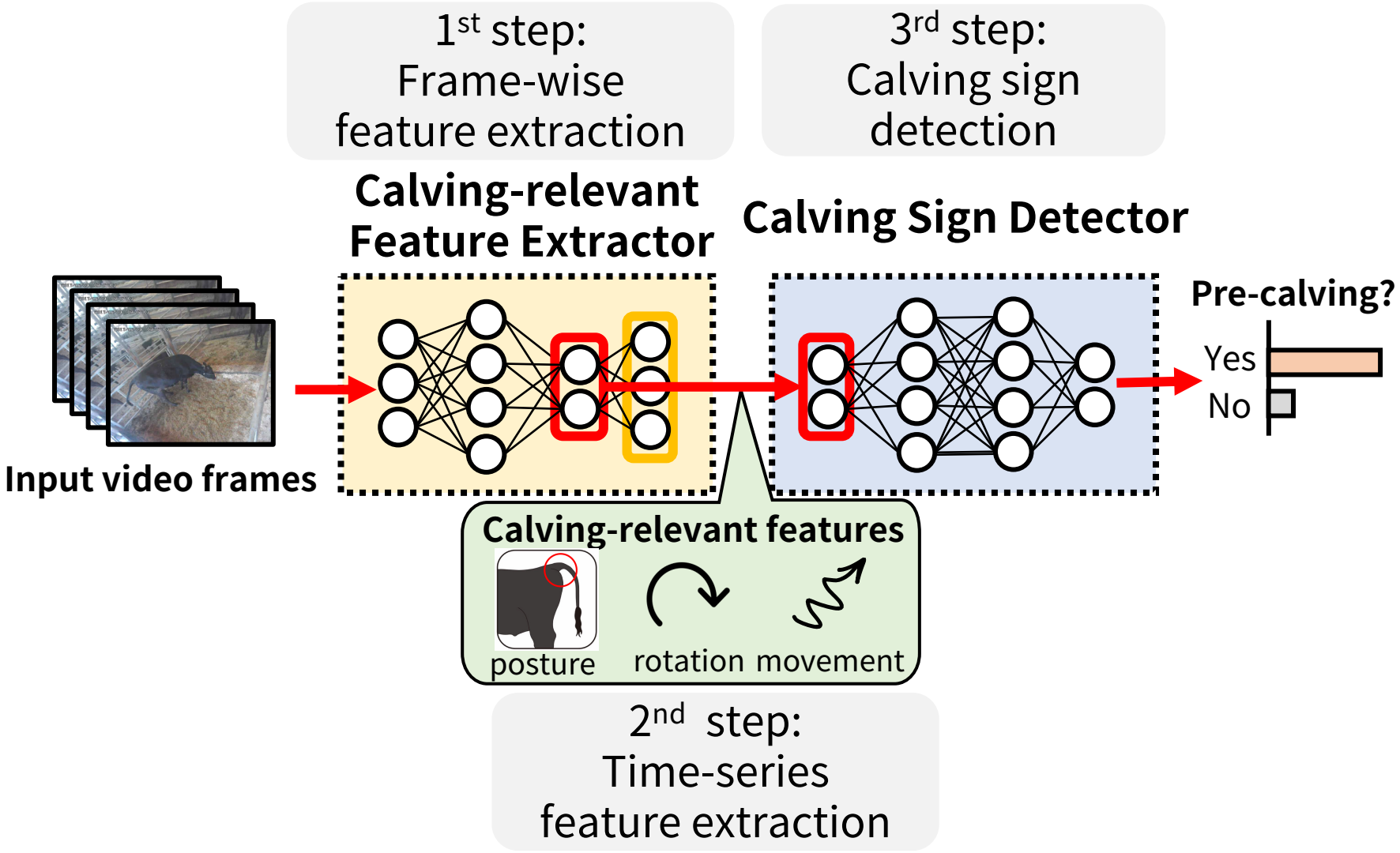


Proposed system

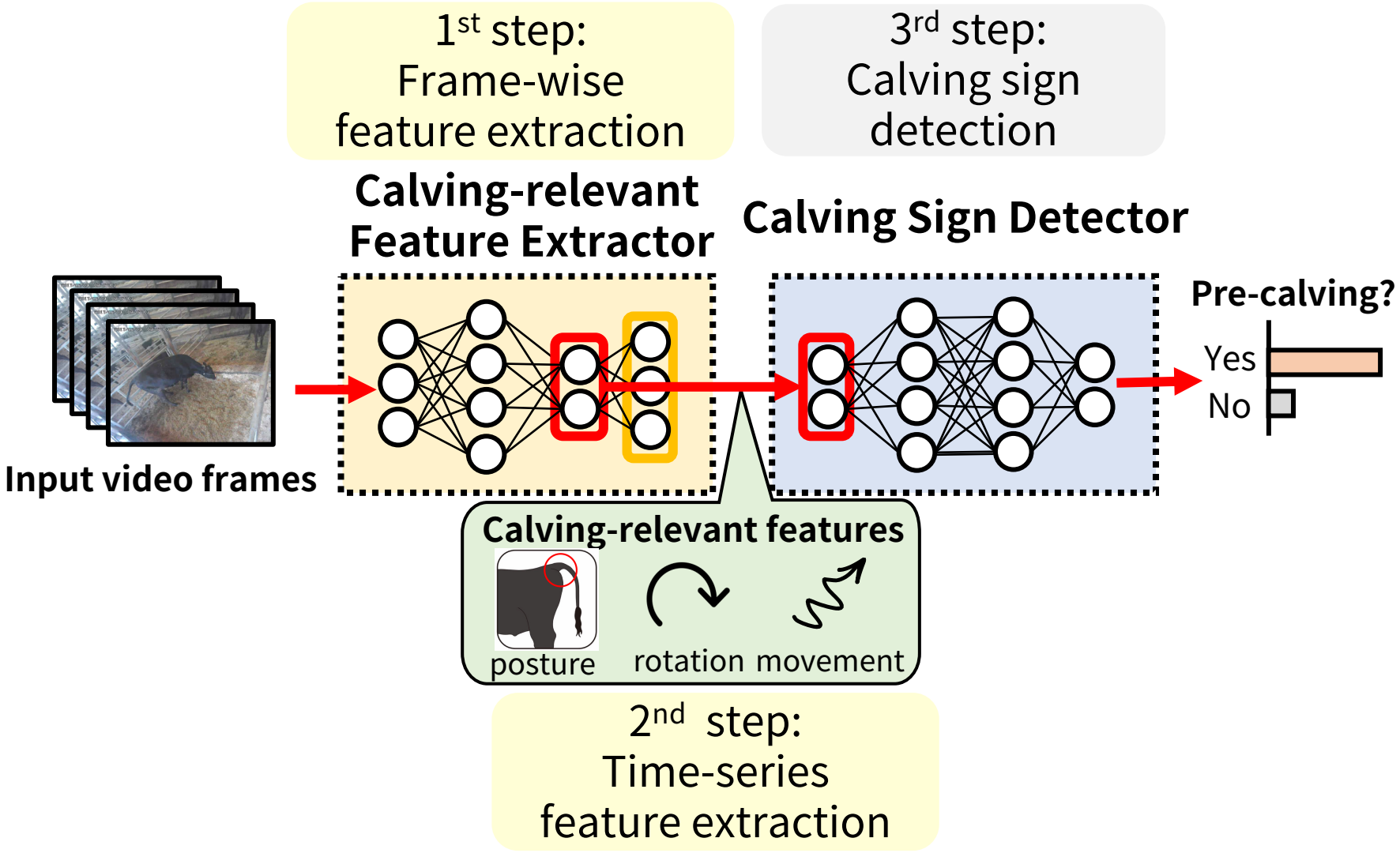




Proposed system

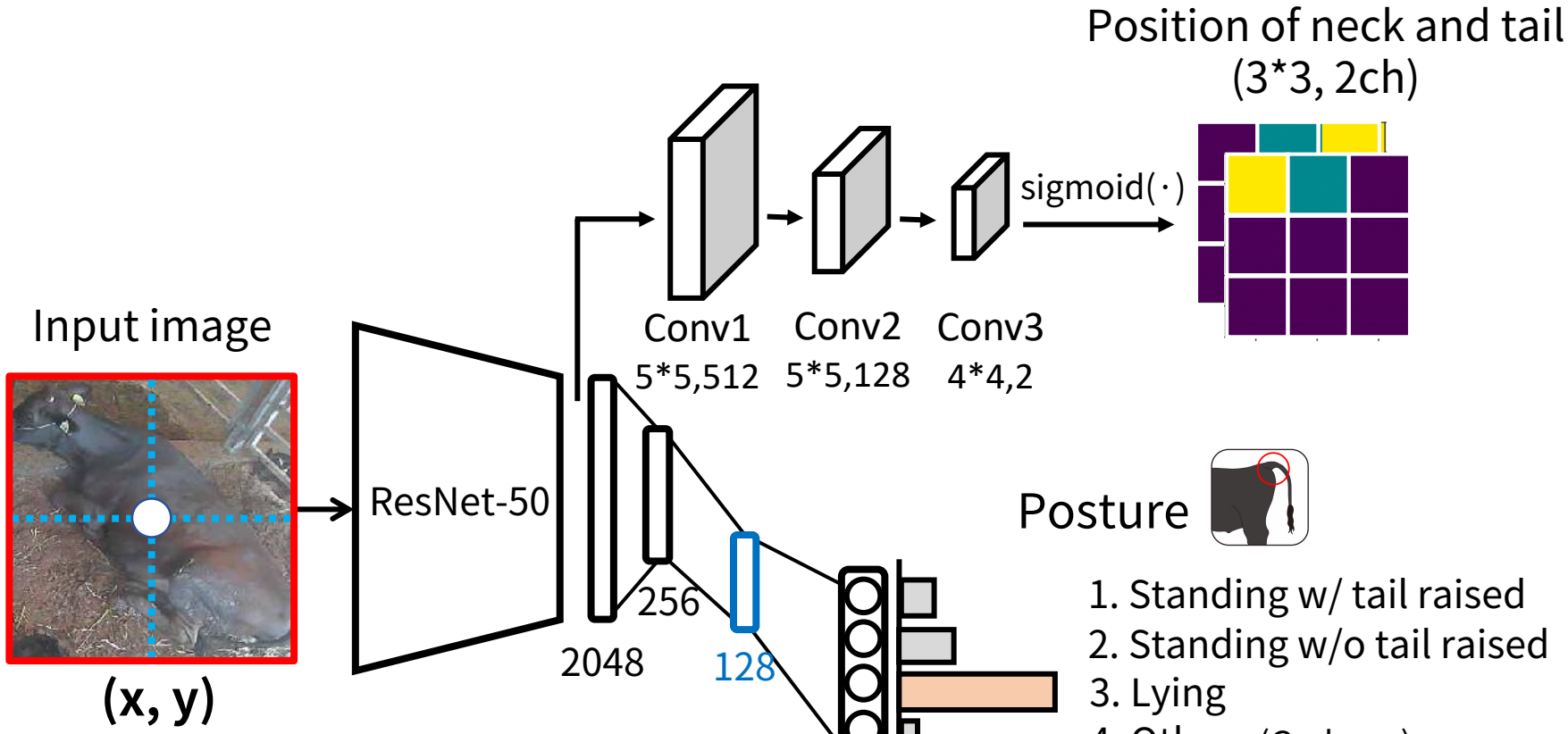


Proposed system



1st step: Frame-wise feature extraction

Posture-based feature: Hidden layer outputs for posture classification tasks (128dim)
Rotation-based feature: Heatmap of the neck and tail position estimation tasks (3*3, 2ch)
Movement-based feature: Cows' region coordinates detected by YOLO v3 (2dim)



2nd step: Time-series feature extraction

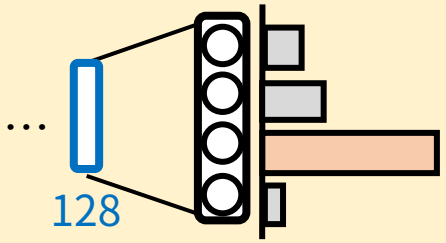
Posture

Rotation

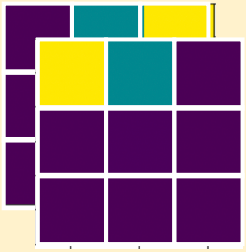
Movement

Frame-wise features

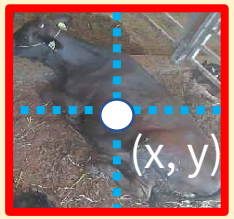
Hidden layer outputs



Heatmap of neck and tail position



Cows' region coordinates



Time-series features

Posture frequency information
(Pooling vector)

||

Posture frequency

Changes in neck and tail position
(M-measure vector)

||

Rotation amount

Difference in coordinates
[Δx, Δy]

||

Movement pattern

Tail raising

Switching between standing & lying postures

Increase in # of rotations

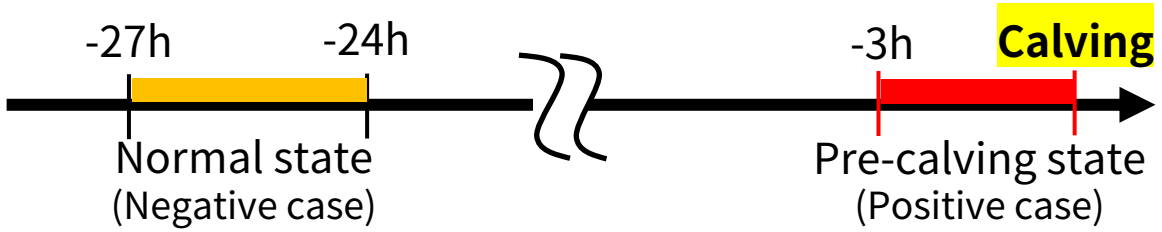
Increase in aimless walking time

Experiment

Comparison of detection performance
with and without explicit feature extraction about calving signs.

Data:

- 15 calving scenes recorded in Kagoshima, Japan



Systems to compare:

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

Evaluation metrics:

- AUC, F1-score, Precision, Recall

Results

Evaluation 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

Results

Evaluation 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

✓ Performance of proposed systems was better over E2E system

Results

Evaluation 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

✗ **E2E system detected calving signs frequently in a normal state.**

Summary

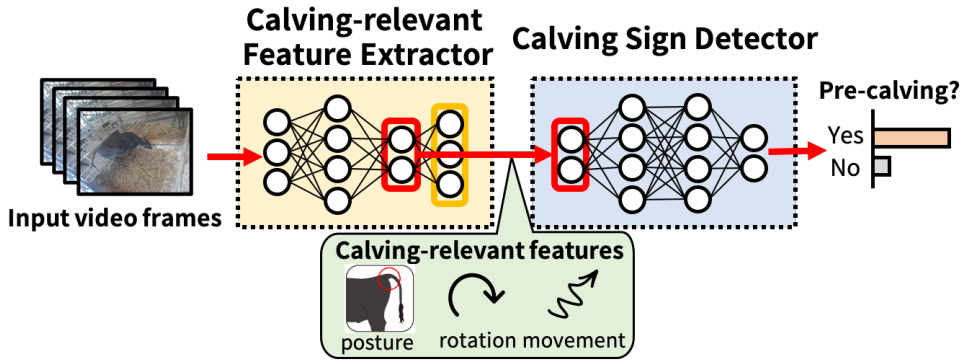
Objective:

Robust calving 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



Results:

Outperformed the E2E system on a small data.

- With explicit feature extraction, the proposed systems suppress obvious false positives.