Photo: Thomas Suchanek/Vetmeduni

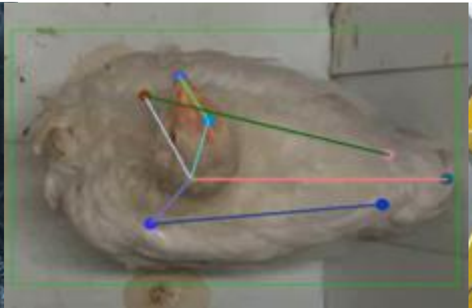
vetmeduni

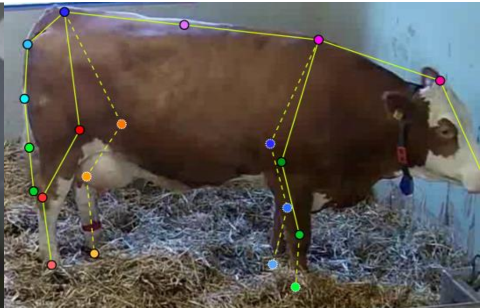
Vision-based Detection of Pain and Nest-building Behaviors in Sows within Commercial Farrowing Pens

Peter Helf, Maciej Oczak

Precision Livestock Farming Hub
Clinical Department for Farm Animals and Food System Transformation
University of Veterinary Medicine Vienna

June 30, 2026



Sensor based management of sows

Automated monitoring of gestating sows to…

- improve health and welfare
- increase productivity

Need for monitoring during gestation and in farrowing pens

Reduce stillbirths, piglet losses, and sow deaths

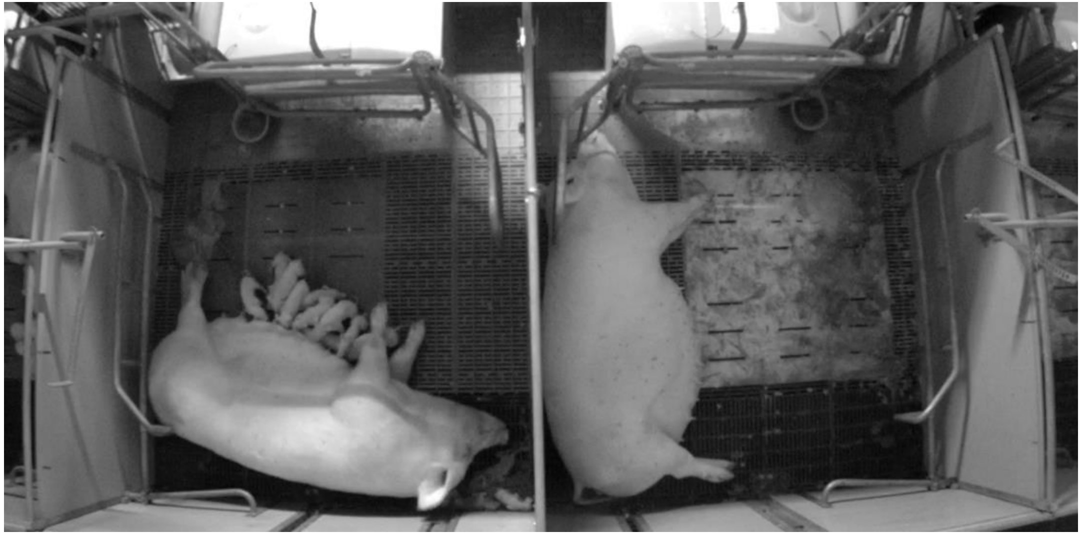
Time to farrowing estimation

Predict the onset of farrowing

By knowing the time of farrowing in advance farmers can be prepared to assist

Prevent piglet crushing

Due to new EU legislation the use of farrowing crates will be limited

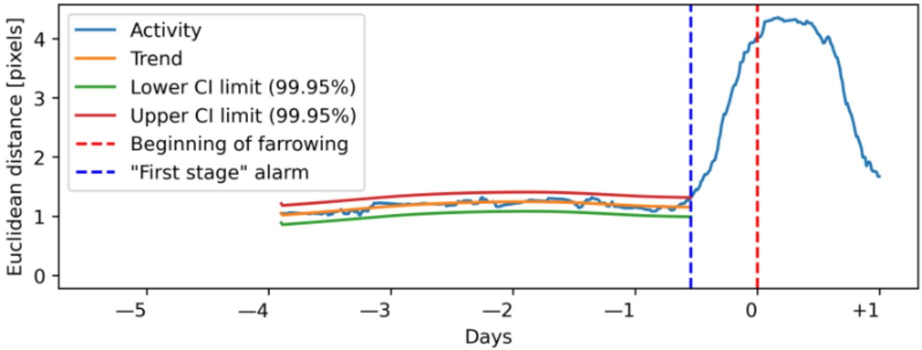


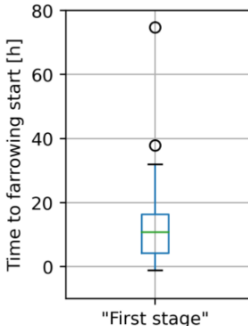
Time to farrowing estimation using activity

Determine how active a sow is using sensors
- Accelerometers in ear tags to measure movement
- Use of object tracking with computer vision

Median time of first stage alarm: ~13h
1st quartile of first stage alarm: ~4h
3rd quartile of first stage alarm: ~16h

13 out of 44 sows did not raise any alarm



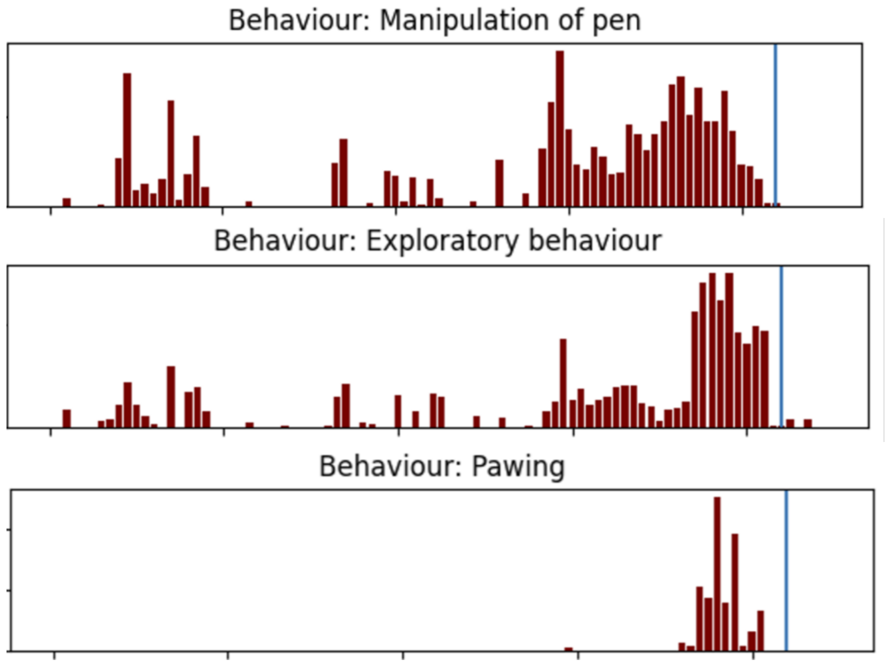


Modified from "M. Oczak et al., Implementation of computervision-based farrowing prediction in pens with temporary sow confinement, Veterinary Sciences,10(2):109, 2023."

Time to farrowing estimation using behaviors

Sow behaviour changes before the onset of farrowing

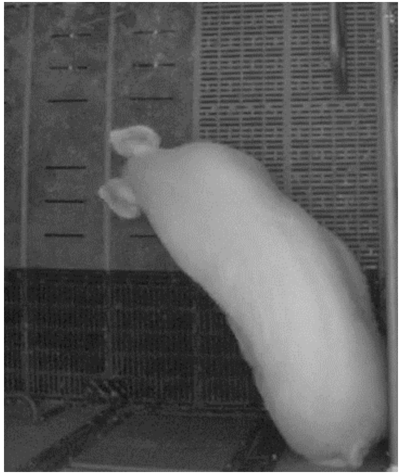
Automatically detected behaviors can be used to estimate time to farrowing



Nest-building behaviors

Pawing

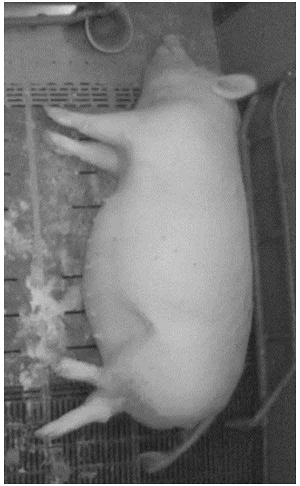
Manipulation of pen



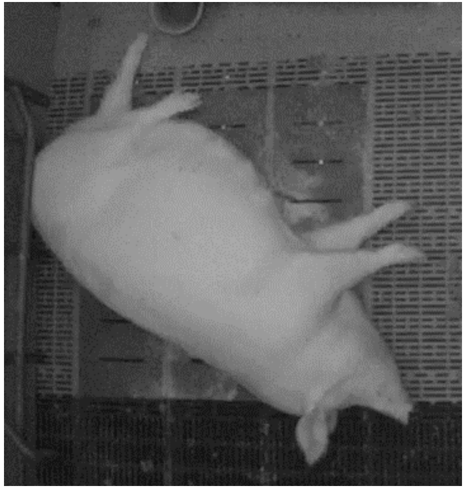
Exploration

Pain-associated behaviors

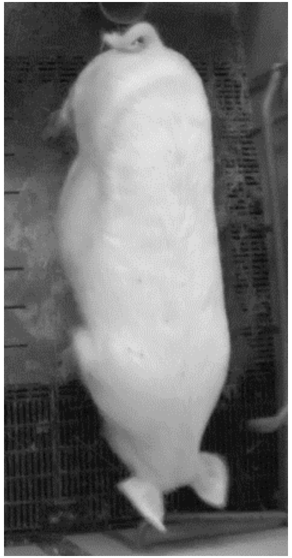
Tremble



Back leg forward



Back arch



Tail flick

Labelled dataset

11 farrowings have been labelled by human observers, 68 hours per farrowing, 748 total

46,010 behavior events have been labelled

Back leg forward makes up over 50% of the dataset

Behavior	Count
Pawing	1,950
Manipulation of pen	5,428
Exploratory behavior	4,938
Tremble	1,267
Back leg forward	28,355
Back arch	2,045
Tail flick	2,027
Total	46,010

Labelled dataset

Downsampled back leg forward events

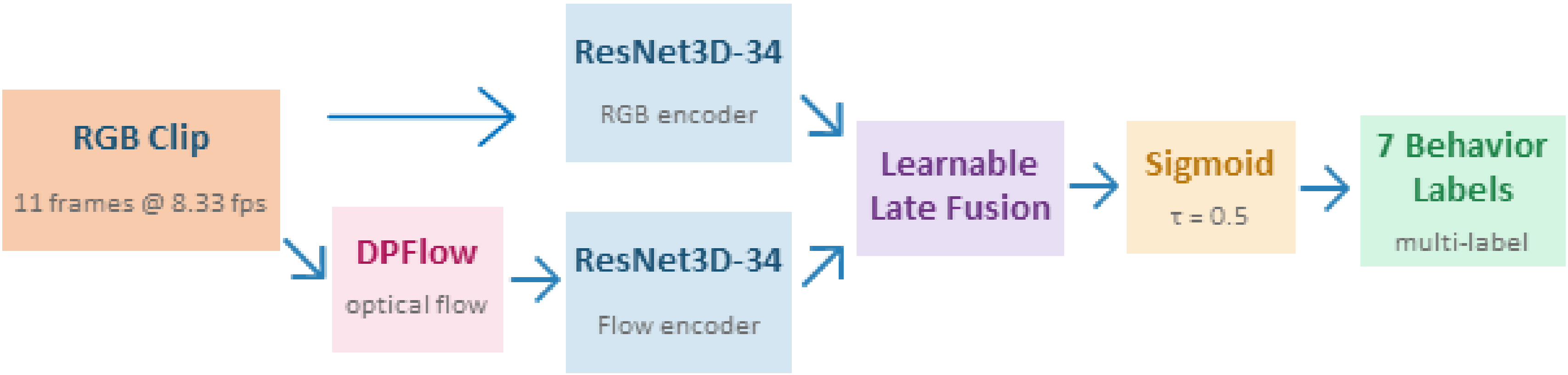
Split the dataset on sow level:

- 8 sows in training set
- 3 sows in validation set

Behavior	Count
Pawing	1,950
Manipulation of pen	5,428
Exploratory behavior	4,938
Tremble	1,267
Back leg forward	5,671
Back arch	2,045
Tail flick	2,027
Total	23,326

Behavior detection pipeline

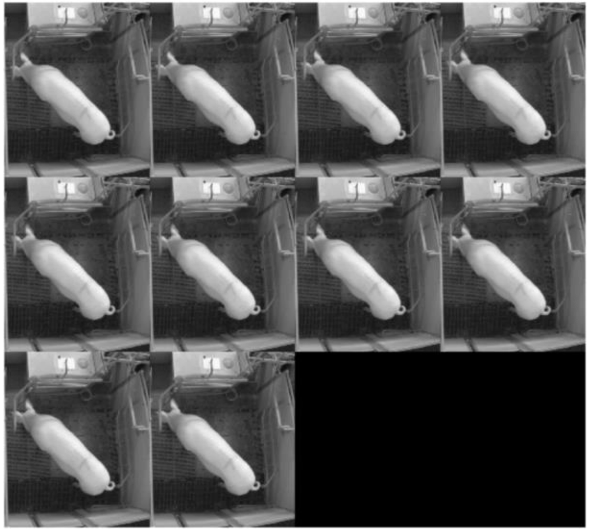
Adapted from "Bohnslav et al. (2021). DeepEthogram, a machine learning pipeline for supervised behavior classification from raw pixels. elife, 10, e63377."

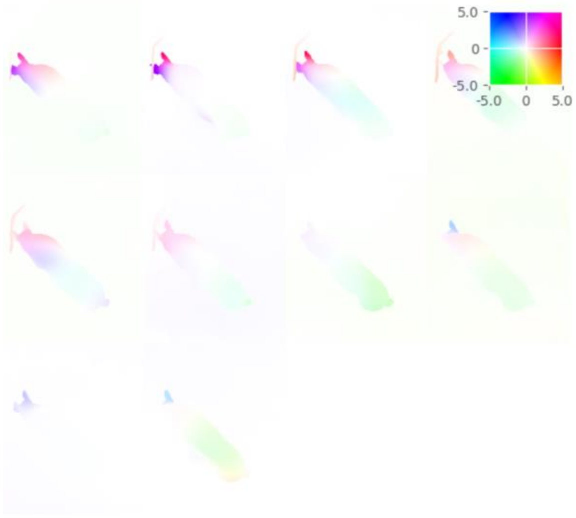


Optical flow extraction

The DPFlow architecture was used to estimate the optical flow between frames

Pretrained weights, trained on the Spring dataset, were used





Feature extraction

The RGB images and the optical flow data are both sent through independent ResNet3D-34 models

ResNet3D is designed to work on multiple images

The output of the ResNets is a value for each behavior

Feature fusion

The features of the two models are combined using weighted averaging

$$l_i = \alpha_i l_i^{RGB} + (1 - \alpha_i)\, l_i^{Flow}, \qquad y_i = \sigma(l_i)$$

$\alpha_i \in [0, 1]$ are learnable fusion weights, and $\sigma(\cdot)$ denotes the sigmoid

The final output is the probability for each behavior, a threshold of 0.5 was used during validation for the calculation of metrics

Input Processing

- 25 fps → every 3rd frame (≈8.33 fps)
- 11-frame clips (≈1.32 s duration)
- Resized to 416 × 499 px
- Apply normalization

Loss & Optimization

- Focal loss
- Initial LR: 10^{-3}
- ReduceLROnPlateau (patience: 5, factor: 0.1)

Augmentation

- Random horizontal flips
- Rotations and scaling
- Brightness/contrast jitter

Pretrained Weights

- ResNet3D-34 pretrained on Kinetics700
- DPFlow pretrained on Spring dataset

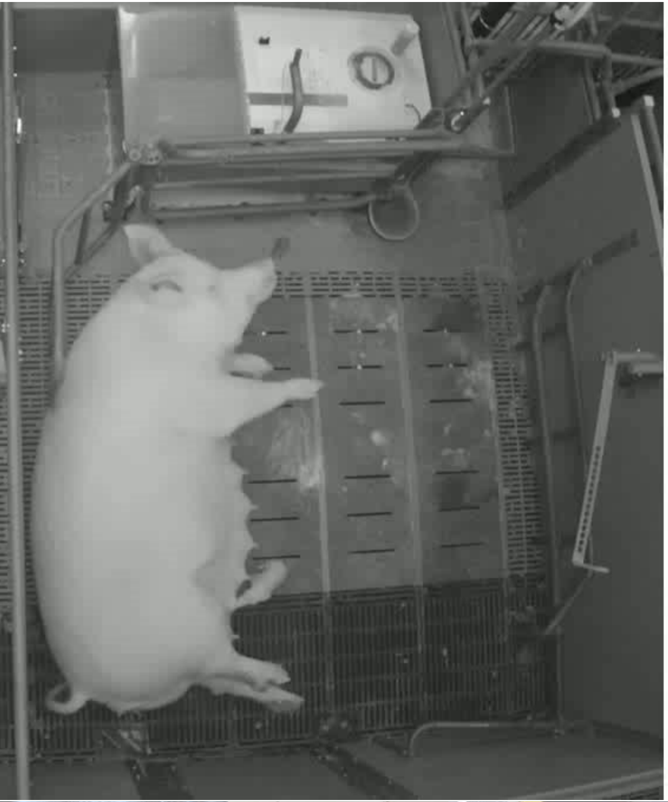
Results on validation dataset

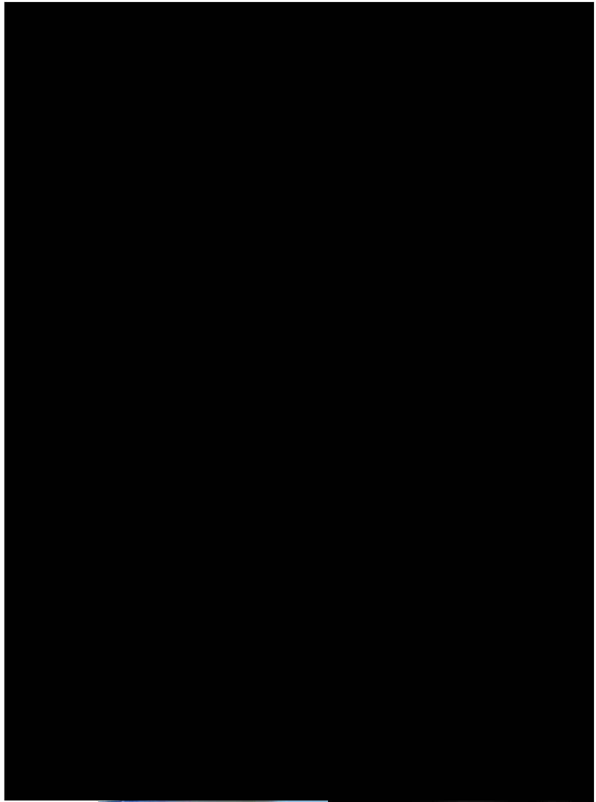
Behavior	F1	Precision	Recall
Manipulation of pen	0.875	0.835	0.919
Back leg forward	0.871	0.907	0.838
Exploratory behavior	0.820	0.827	0.814
Tail flick	0.778	0.833	0.730
Back arch	0.639	0.854	0.511
Pawing	0.634	0.704	0.576
Tremble	0.443	0.383	0.524
Macro Average	0.723	0.763	0.702

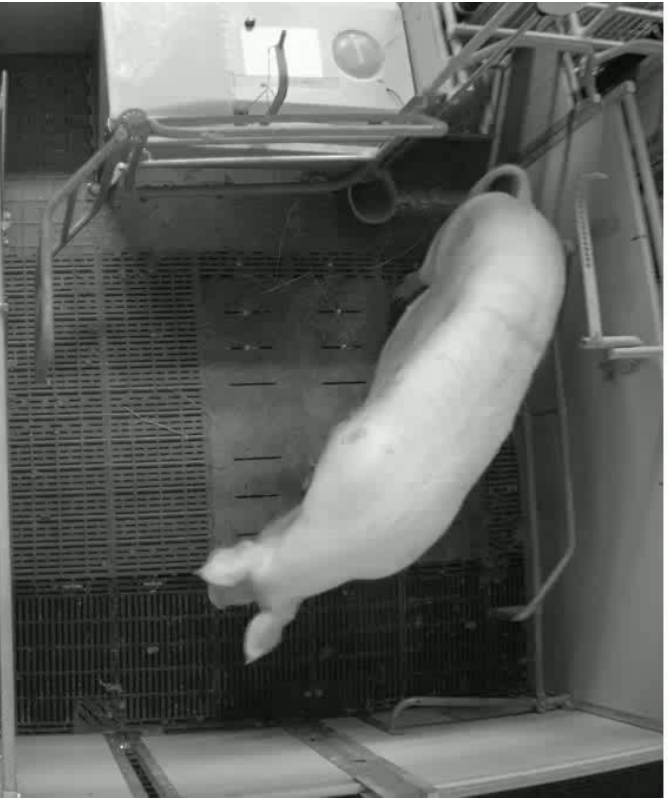
Nest-building   Pain-associated

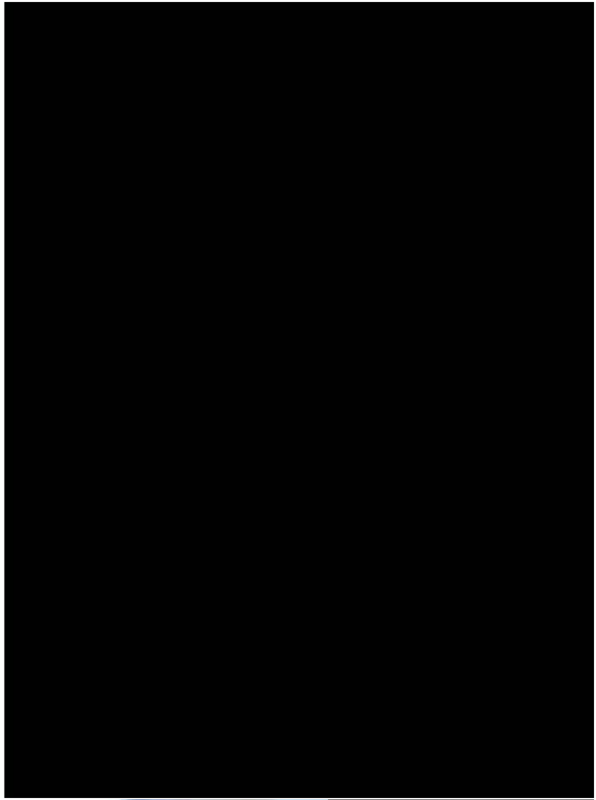
University of Veterinary Medicine Vienna

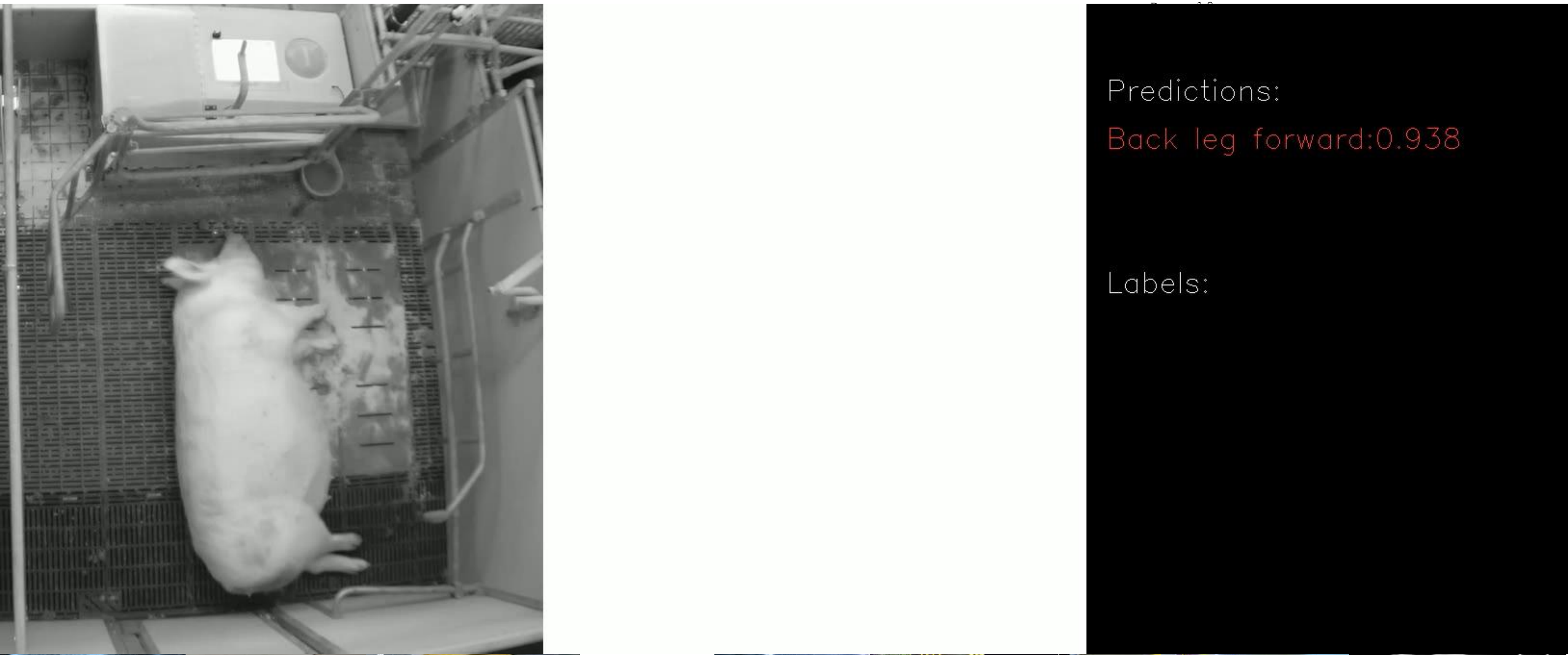
PLF-Hub










Predictions:
Back leg forward:0.938
Labels:

Limitations

Limited diversity

- Single farm, breed (Large White), and pen type, limits visual diversity and generalizability

Class imbalance

- Despite downsampling and focal loss, residual imbalance may still bias decision boundaries

Single camera perspective

- Top-view camera only, side view might be beneficial for some behaviors

Annotation noise

- Substantial annotator agreement (κ=0.724) but inter-annotator variability can still affect results

Future work

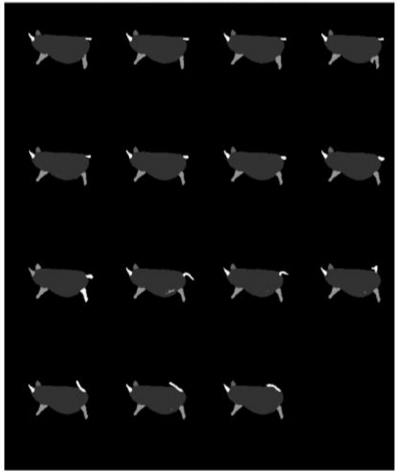
Use of segmentation or keypoints as additional input

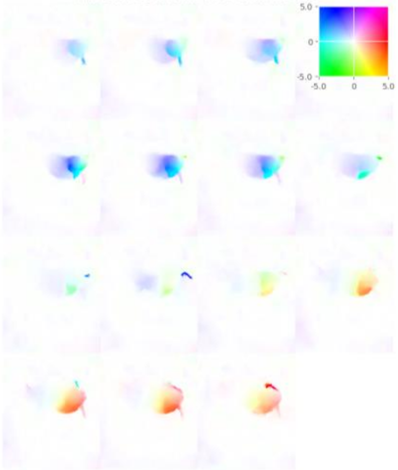
More data labelling to expand the dataset

Record more farrowings

Test different clip lengths/frame rates

Further modification of the neural network architecture





Thank you for your attention!

Peter Helf

PLF-Hub / Clinical Department for Farm Animals
and Food System Transformation
University of Veterinary Medicine, Vienna

peter.helf@vetmeduni.ac.at
Veterinärplatz 1, 1210 Vienna
Austria

Precision Livestock Farming DOC

www.vetmeduni.ac.at/plf-doc

PLF hub Precision Livestock Farming

www.vetmeduni.ac.at/plf-hub

FWF Austrian Science Fund