Automating Body Condition Scoring of Dairy Cows Using Machine Learning on Time-of-Flight Data

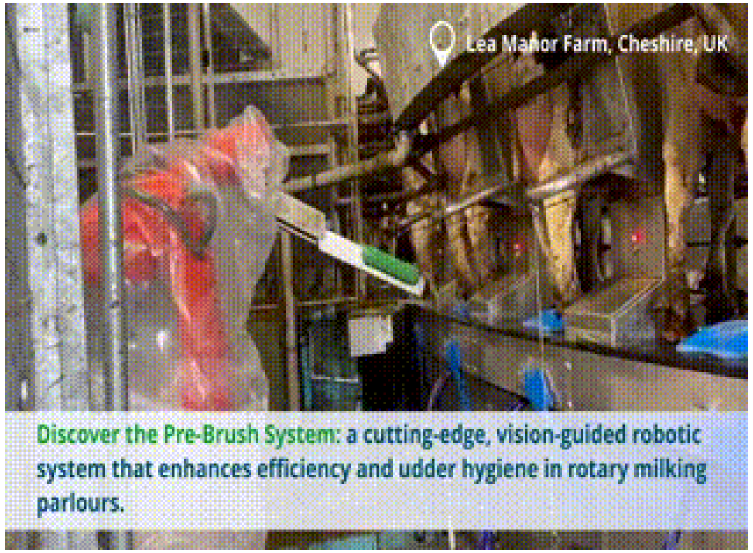
Nicolas Martinez-Baquero, G. Wager-Jones, M. Fujiwara, A. Peacock

AI4AS – Session F

06/06/2025

What we do

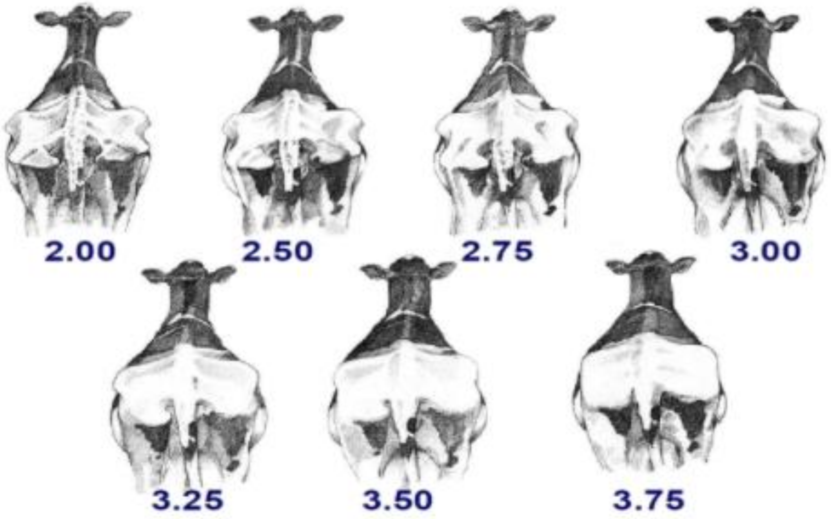
- Robotics and automation company with a focus on AgriTech.
- Based in Stirling, the heart of Scotland.
- Specialised in wearables/pedometers and computer vision for accurate monitoring of cow behaviour and wellbeing.



Peacock TECHNOLOGY

Body Condition Scoring (BCS)

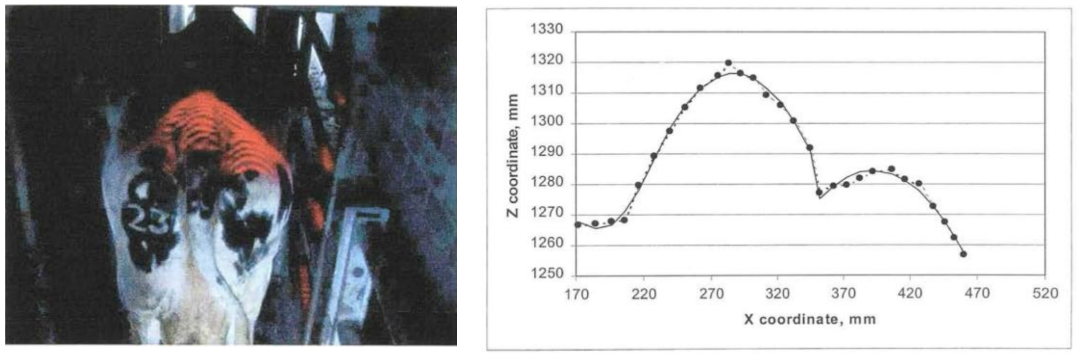
- Proxy measurement of fat reserves in dairy cow's body and other ruminants.
- Multiple scoring systems available, the most common one in the UK is the one developed by the Agriculture and Horticulture Development Board (AHDB) in the 1980s.
- 1-5 scores in increments of 0.25.
- A score of 1 indicates an emaciated cow, a score of 5 represents an obese one.
- Done by trained scorers (farmers, nutritionists, veterinarians).
- Used to assess nutritional status and aid on-farm feeding decisions.
- It's time consuming and subjective.



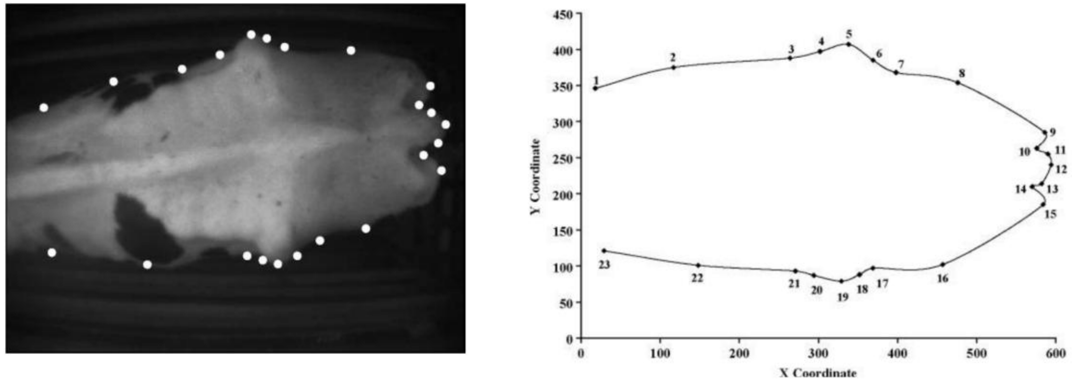
Courtesy of AHDB

Peacock TECHNOLOGY

Previous work on automating BCS



Manual extraction of features from laser lines,
Coffey et al., 2003



Manual extraction of 23 anatomical keypoints,
Bewley et al., 2008

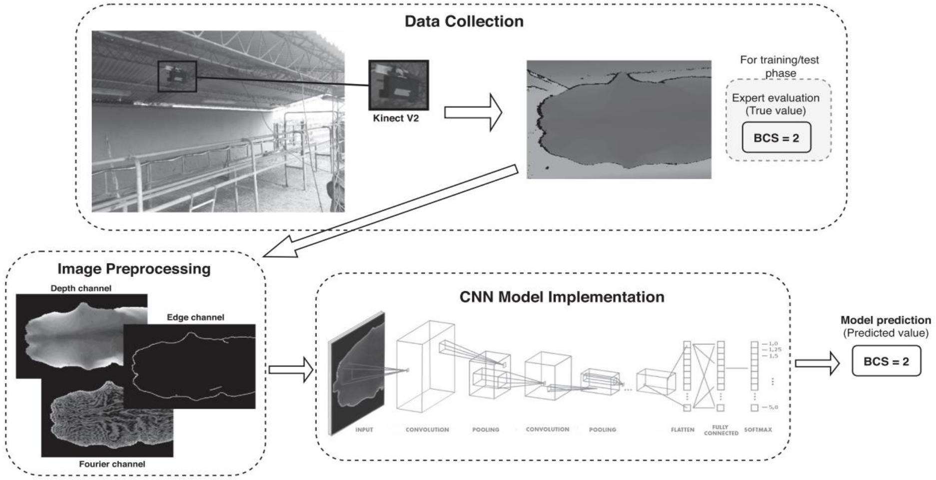


Fig. 1. Overview of developed BCS estimation system.

CNN trained on Kinect TOF data
Rodriguez-Alvarez et al. (2018)

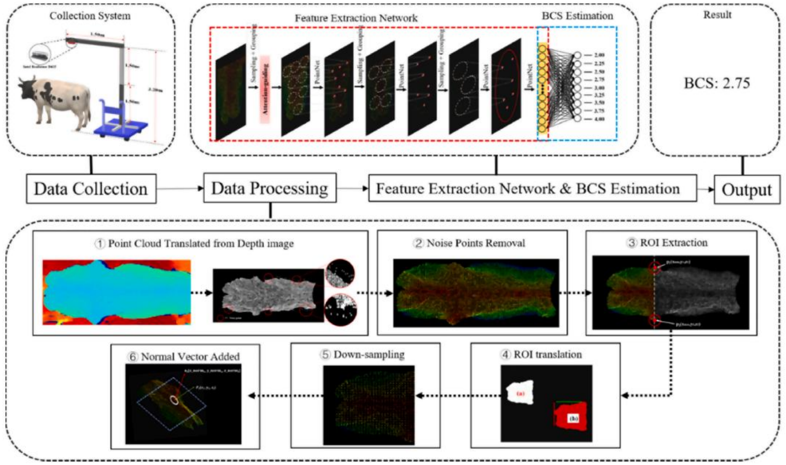


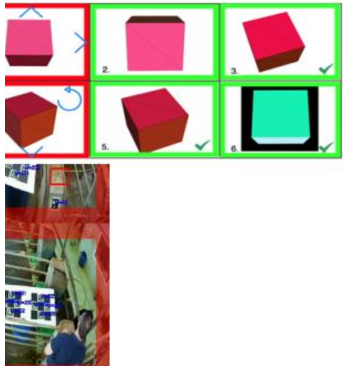
Fig. 1. The workflow of this study.

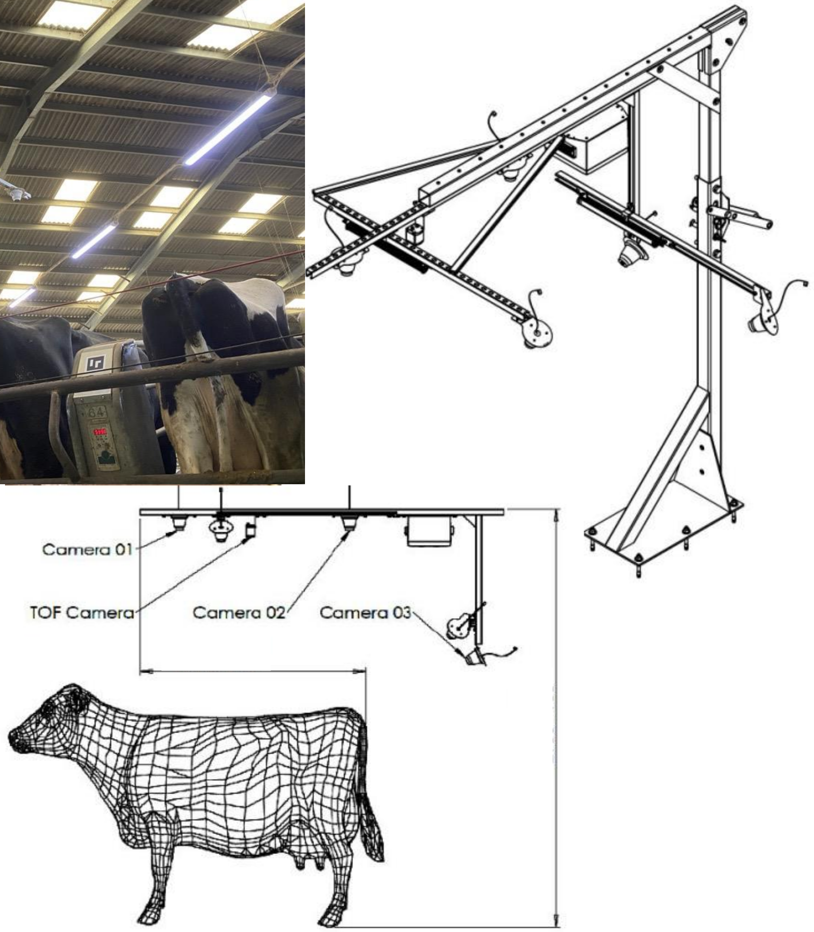
CNN trained on point clouds from depth
images, Shi et al. (2023)

Peacock
TECHNOLOGY
4

Data Collection

- Peacock Technology designed and installed rig comprised of 4 cameras (IP + Time-of-Flight).
- Installed in a rotary parlour at a farm in the U.S.A.
- ~4K cows, 2 milkings/day.
- Data captured between August-December 2024.
- Images are captured and automatically uploaded to the Cloud.

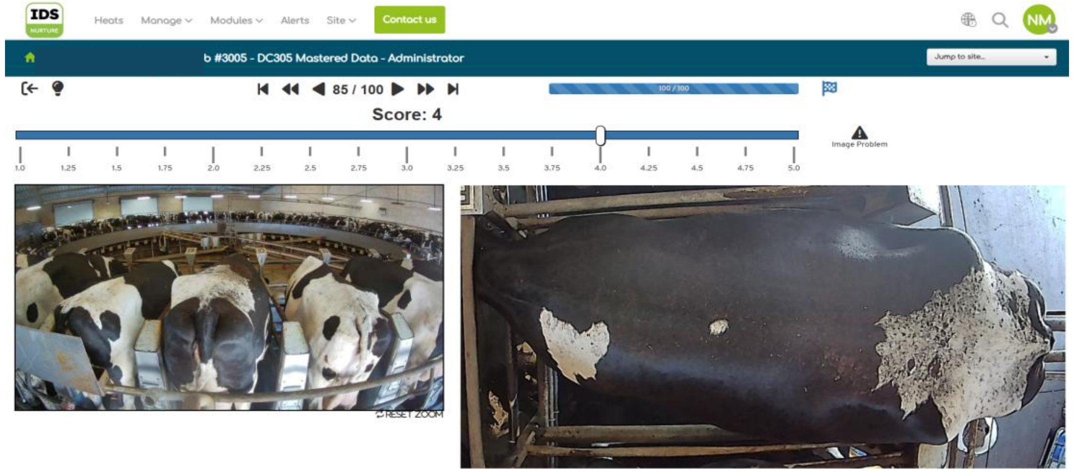




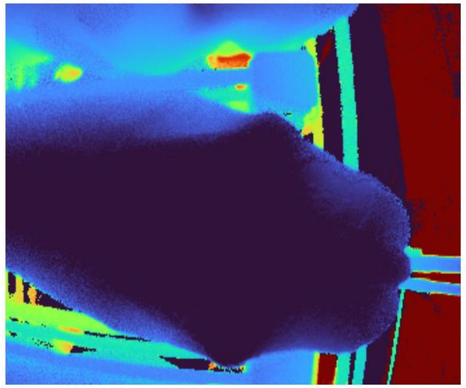
Camera Rig

Data Collection

- Body condition scoring (BCS) performed by an expert veterinarian (30+ years of experience) and high intra- scorer agreement (quadratically weighted Cohen's kappa = 0.79)
- Each cow was scored 3 times, independently, by the expert scorer.
    - Outlier score dropped if considerable disagreement between a score and the other two values.
    - If there was significant disagreement among the 3 scores, the image/scores were discarded.
- Scoring done remotely.



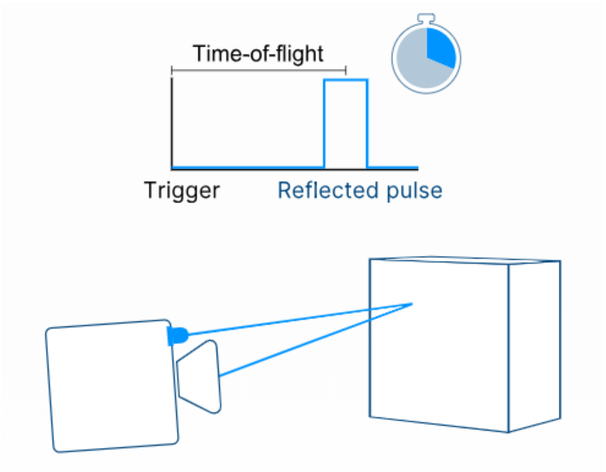
Scoring dashboard

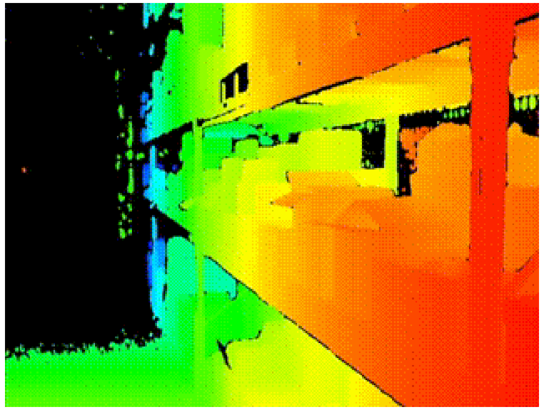


TOF image (false colour)

Time-of-Flight data

- TOF cameras have a single source that emits light.
- Measuring the time elapsed (or phase shift) of the reflected light to calculate the distance to objects in the scene.
- Each pixel represents distance information.
- Bypasses the need of complex sensors/and or algorithms to estimate depth.
- Cost-effective 3D imaging solution.





TOF pulsed light measurement, courtesy of Basler

Dataset curation

- The curated dataset consists of TOF images from **Farm A:**
  - 17,700 used for training/validation.
  - 7,300 used for testing.
- This corresponds to 3,800 and 820 Holstein-Friesian cows in the training and test sets, respectively.
- Capture of consecutive milkings allow significant increase in the size of the dataset.
- Assumption: BCS of the cow won't vary significantly ± 3 days of the scoring capture.

-2 milkings

-1 milking

Scored capture

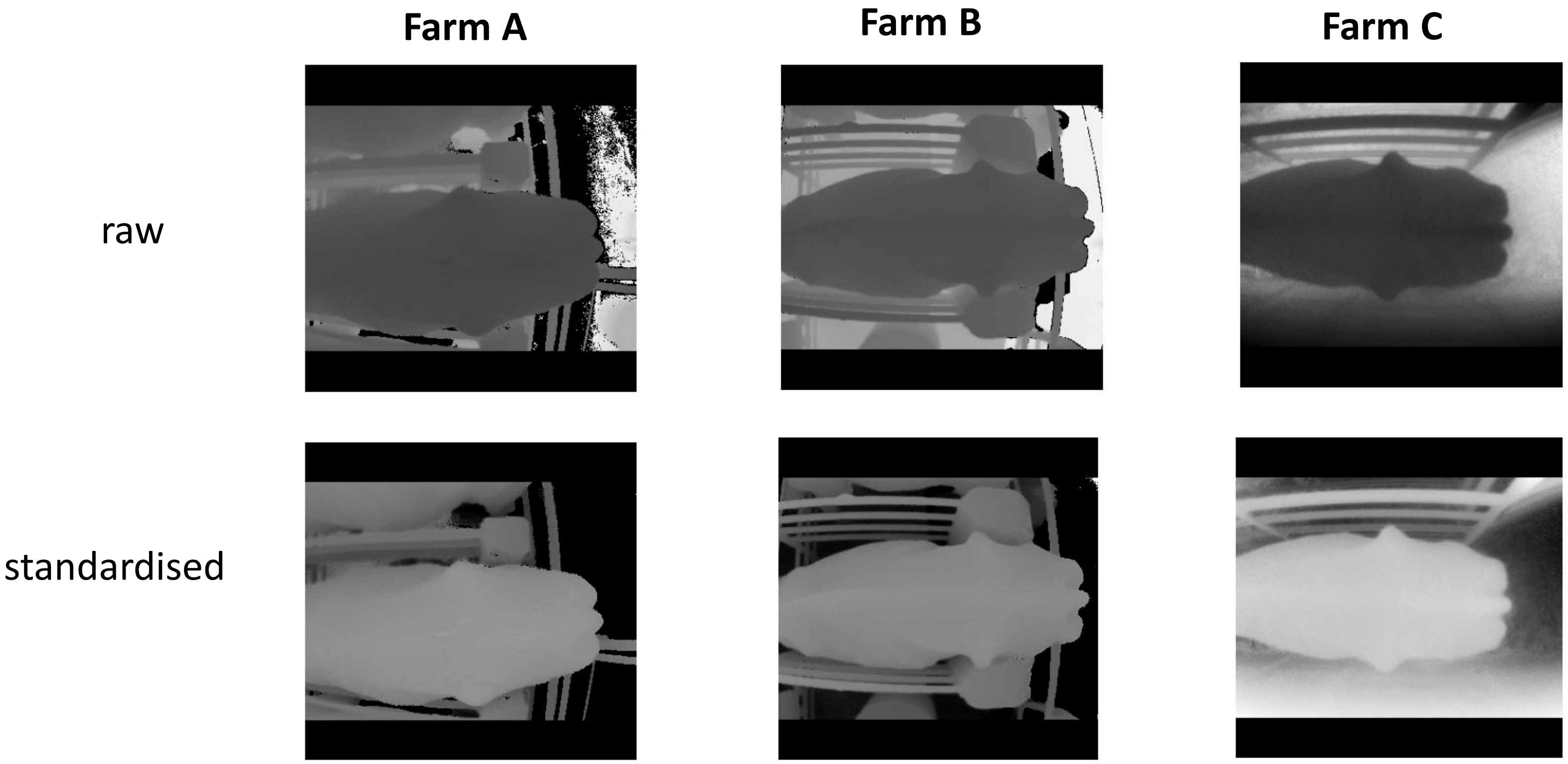
+1 milking

+2 milkings

Peacock TECHNOLOGY

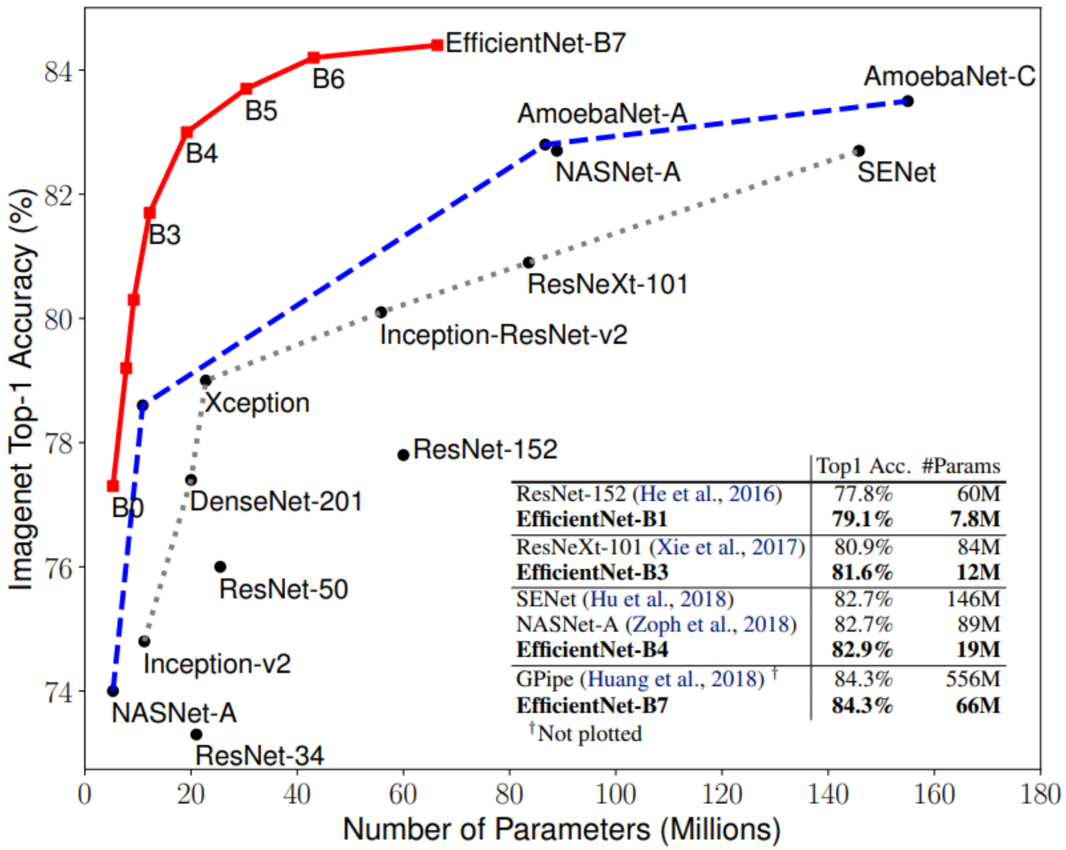
Data pre-processing

- "Standardising" the images led to smoother training and better generalisation when the model was applied to data from other farms.



Machine Learning Model

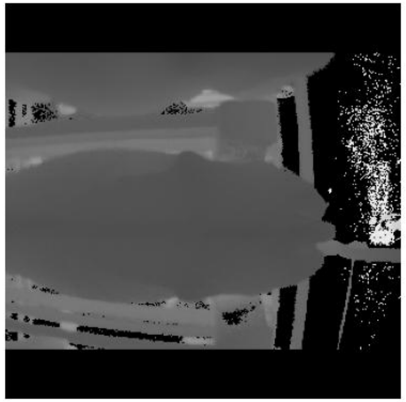
- Trained an EfficientNetB0 CNN with a regression head.
- Finetuned a model pre-trained on ImageNet.
- Efficient scaling method to achieve state-of-the-art performance with faster and smaller architectures (Tan et al., 2020):
  - EfficientNetB0: **5.3M** parameters.
  - 224 x 224 x3 input size.
- Larger model --> more training data.

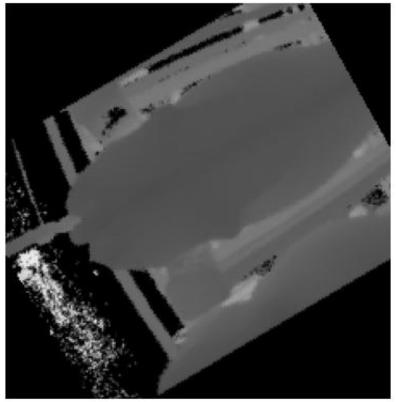


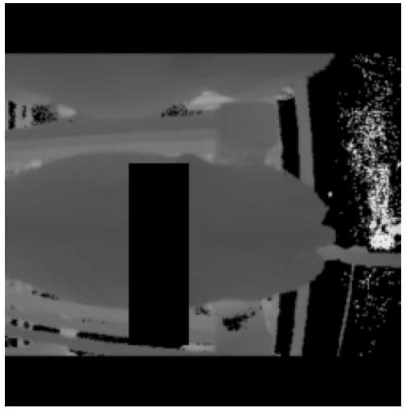
Model size vs ImageNet performance (Tan et al., 2020)

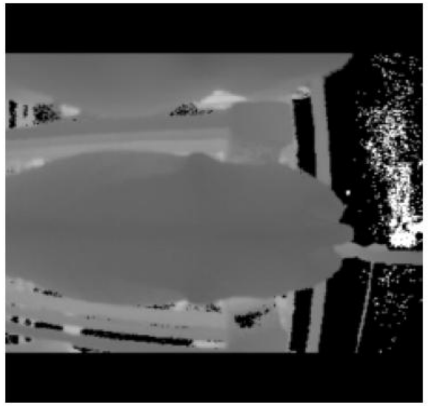
Training Procedure

- Trained the CNN for 300 epochs, using an AdamW optimizer and Cosine Annealing with Warm Restarts scheduler.
- Mean squared error (MSE) loss function.
- Learning rate of 1e-3.
- Batch size of 128.
- Trained on an Nvidia RTX 3090 GPU (24GB).
- Moderate levels of data augmentation during training (random flipping, rotation, colour jitter, erasing).
- Hyperparameter tuning through Bayesian search.


Original/padded image


Random flipping/rotation

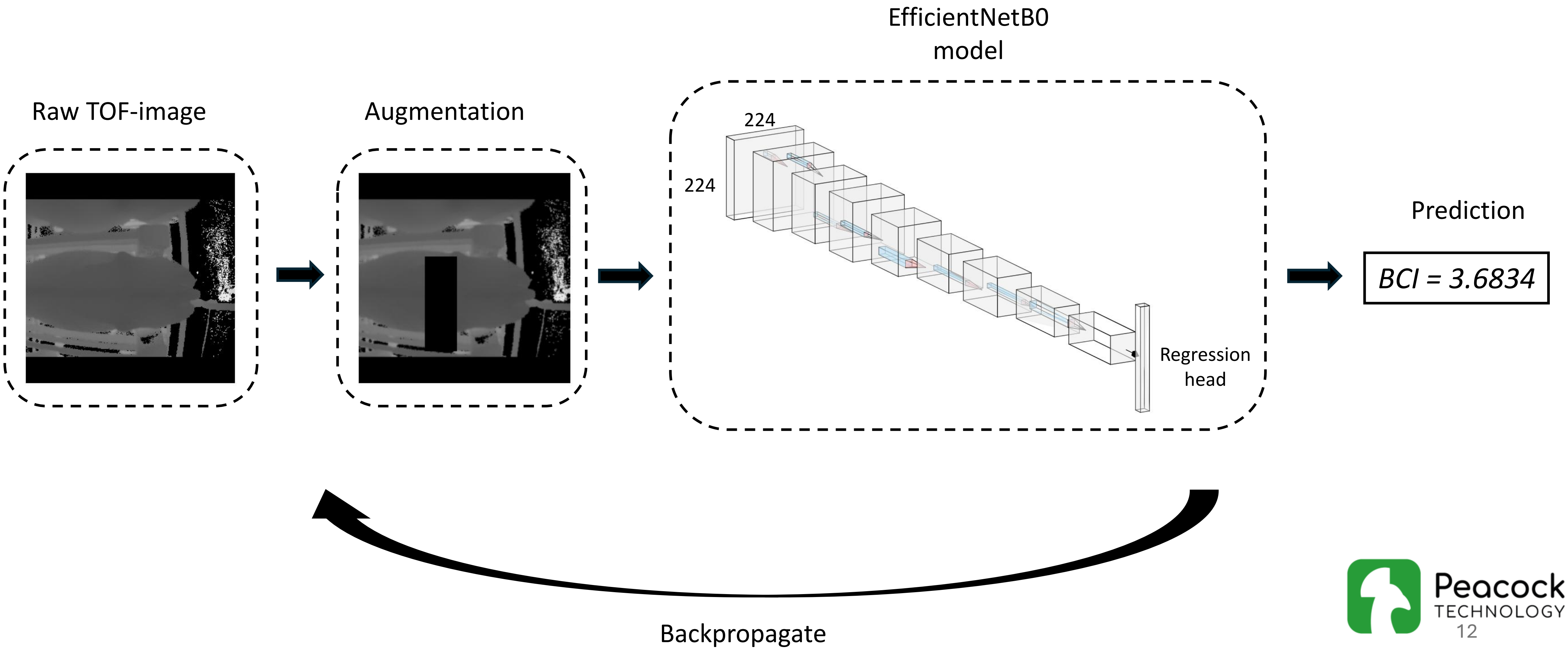

Random erasing


Color jitter

Peacock TECHNOLOGY

Pipeline

- The prediction is a Body Conditioning Index (BCI), which should be correlated with BCS.



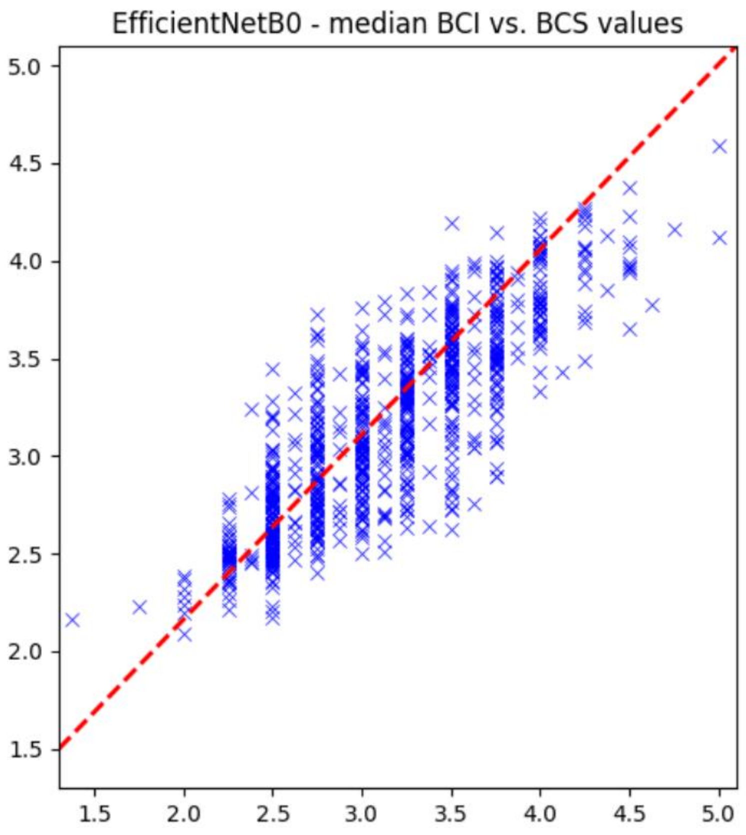
Peacock TECHNOLOGY

Raw data - test set results

- Metrics obtained by taking the median of ± 3 days.
- Number of cows: 823.

Metric	value
R2	0.7061
Pearson correlation	0.8411 (p < 0.001)
RMSE (global)	0.3087
RMSE (BCS ≤ 2)	0.3972
RMSE (2 < BCS ≤ 3)	0.3001
RMSE (3 < BCS ≤ 4)	0.2973
RMSE (4 < BCS)	0.4664

median BCI



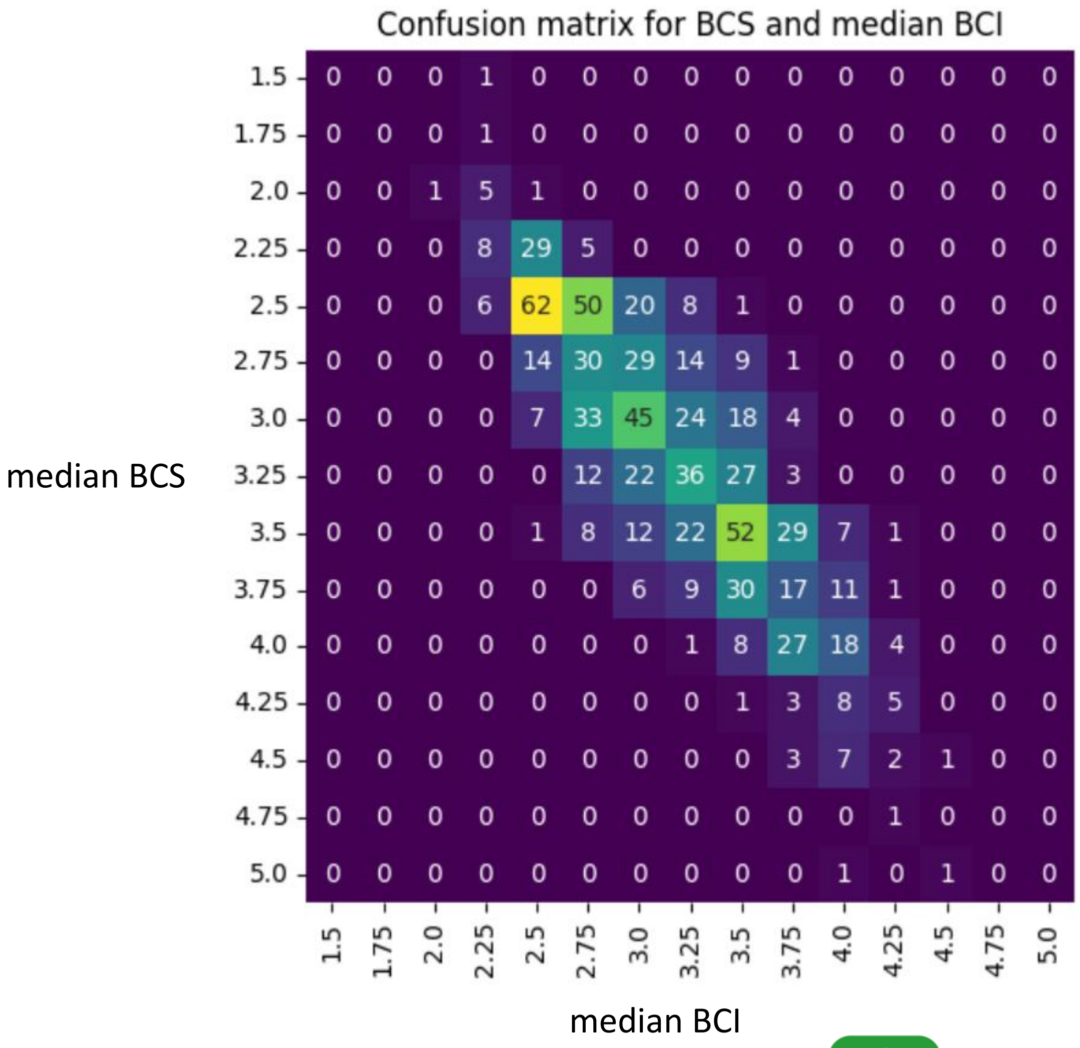
median BCS

Model-scorer agreement and repeatability

- Cohen's kappa: observed agreement between raters, considering the agreement that would be expected due to chance.
- BCI rounded to the closest 0.25 step.
- Model-scorer agreement:
  - Quadratically weighted Cohen's kappa: **0.8219**
- Intra-model agreement:
  - Krippendorff's alpha: **0.8861**

Value of Kappa	Level of Agreement
0–.20	None
.21–.39	Minimal
.40–.59	Weak
.60–.79	Moderate
.80–.90	Strong
Above.90	Almost Perfect

Interrater reliability: the kappa statistic, courtesy of the NIH



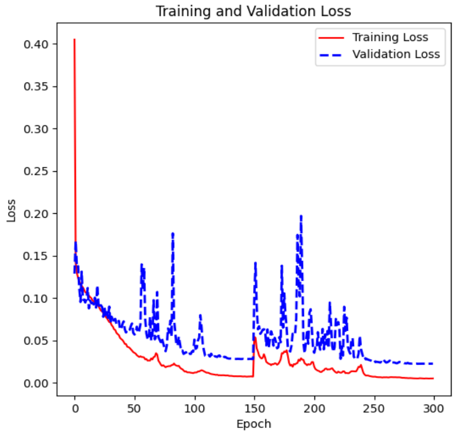
Peacock TECHNOLOGY

Model generalisation

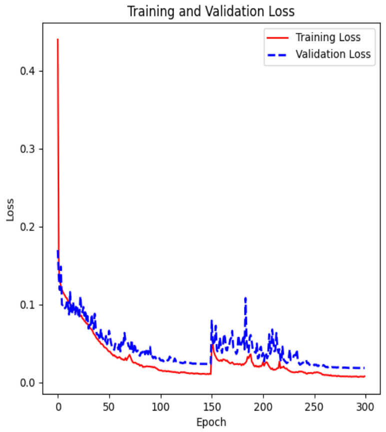
- Better generalisation when trained on standardised data, especially in farms with a completely different set up.

Weighted Cohen's Kappa	**Farm A** (rotary, n = 823)	**Farm B** (rotary, n = 285)	**Farm C** (race exit, n = 123)
Raw images	0.8219	0.7220	0.0032
Standardised images	0.8224	0.7343	0.5060

- Smoother validation loss and better generalisation gap during training:



Training with raw data



Training with standardised data

Future work

- Gather data from additional farms.
- Train larger models/add capacity (EfficientNet B1/B2, EfficientNetV2-S, etc.).
- Commercial deployment and use model-in-the-loop to do active learning.
- Related work: "*Interspecies Crossover of Highly Pathogenic Avian Influenza into Dairy Cattle*" will be presented at the 2025 ADSA Annual Meeting by G.A. Wager-Jones.

Thank you!

Peacock TECHNOLOGY