Pain Recognition in Cows Using Deep Learning

Suresh Kumar Choudhary[1]
Alisa Sergeeva[1]
Kerstin E. Müller[2]
Vitaly Belik[1]

[1]System Modeling Group, Institute of Veterinary Epidemiology and Biostatistics
[2]Clinic for Ruminants
School of Veterinary Medicine, Freie Universität Berlin, Germany

Our team

Alisa Sergeeva

Suresh Choudhary

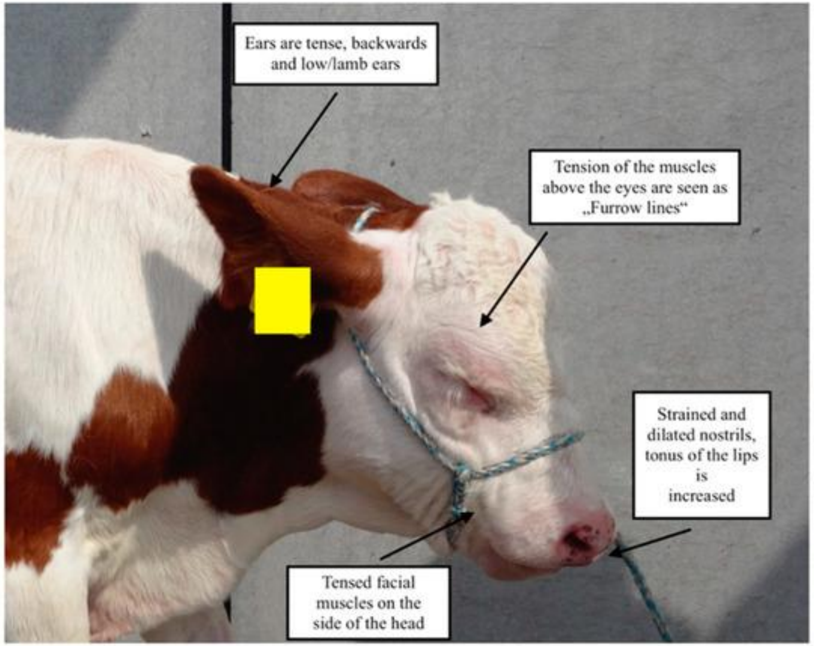
Prof. Kerstin E. Müller

Goal of the study

Source: Tschoner et al. [1], Gleerup et al. [2],

- To develop an automatic, effective and reliable method for pain detecting in cows using advanced deep learning techniques

- Pain indicators: posture changes, vocalizations, feeding behavior alterations, and locomotion patterns etc.

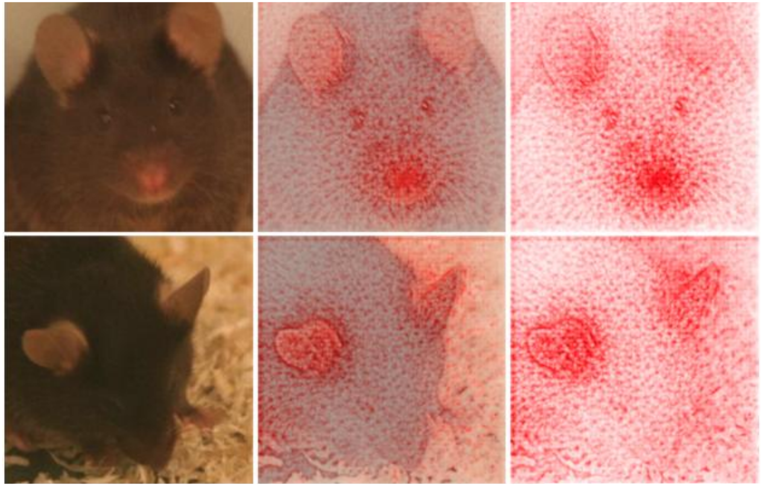


Motivation

- Productivity is affected by pain causing factors
- If pain can be detected early, a lot of diseases can be avoided with less efforts
- Human based pain scoring models require expert observer
- No direct communication and sometimes suppression of the pain indicators
- Traditional human-based pain detection have limitations: labor-intensive, time-consuming, and prone to bias
- Empower dairy farmers and veterinarians with an automated tool for early and accurate pain detection in dairy cows

Animal pain scoring

- No uniform way
- Single-dimensional scoring
- Multidimensional scoring: behavioral and physiological parameters
- Grimace Scale [3,4], Glasgo-CMPS [5] etc.
- Cow Pain Detection - Manual and Sensor Based
- Gleerup et al. [3] explained the possible various visual pain features in cows



Andresen et al, 2020
DOI:10.1371/journal.pone.0228059

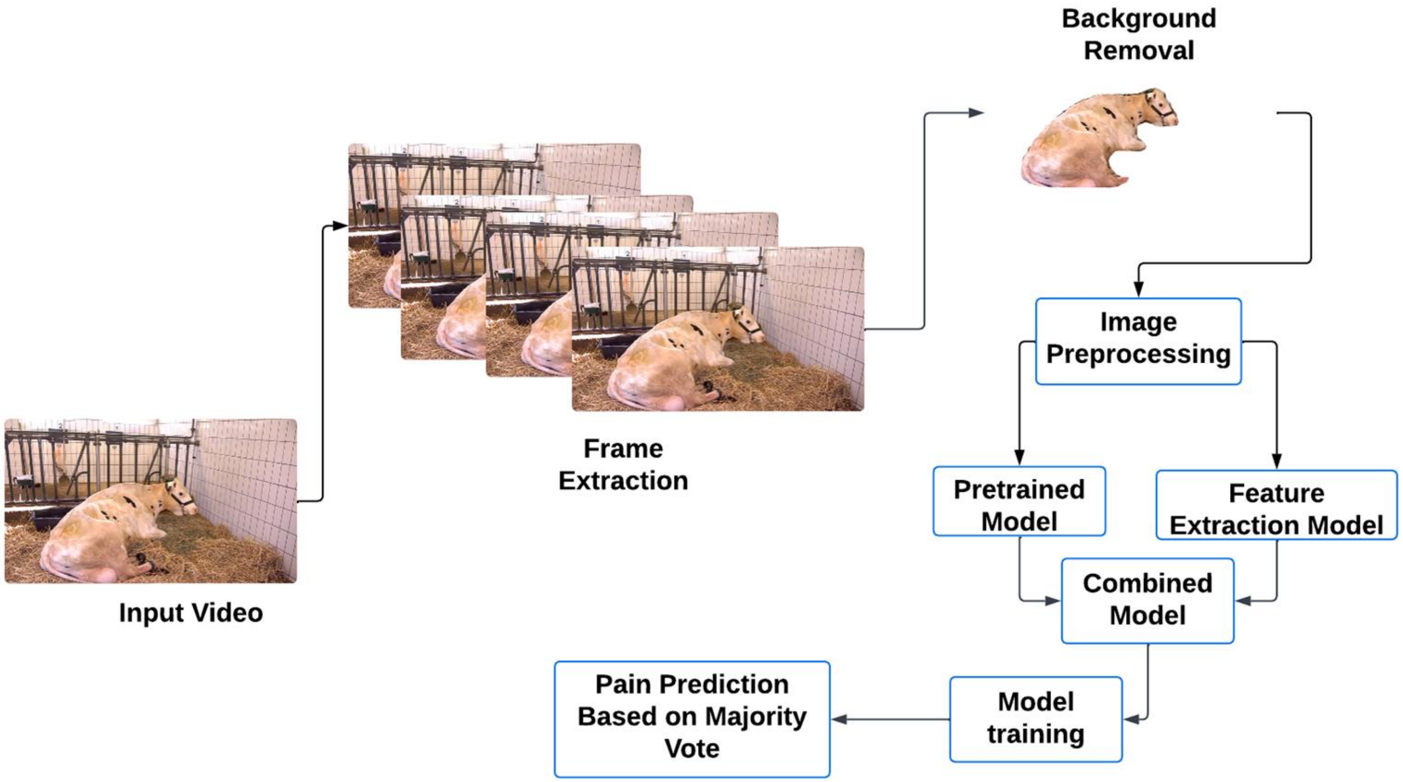
CattleCareDataset

- Created by a veterinary student (Alisa) in the clinic for ruminants
- Holstein cows
- 3 Cows, aged 4–5 years, and 4 calves, aged 1–2 months
- August 16 to August 30 in the clinic for ruminants at FU Berlin
- Attention, Head Position, Ear Position, Facial Expression, Reaction to Approach, and Back Line
- For the model we binarized the scores into pain vs. no pain labels
- 3 Cows, aged 4–5 years, and 4 calves, aged 1–2 months
- Manually scored 1022 images, 766 images with pain and 256 images without pain (after binarization)

Leveraging knowledge transfer

- ImageNet dataset: 14 million images, approximately 18000 cow images with variety of breeds, younger to older, and with many poses.

Proposed methodology (video input)



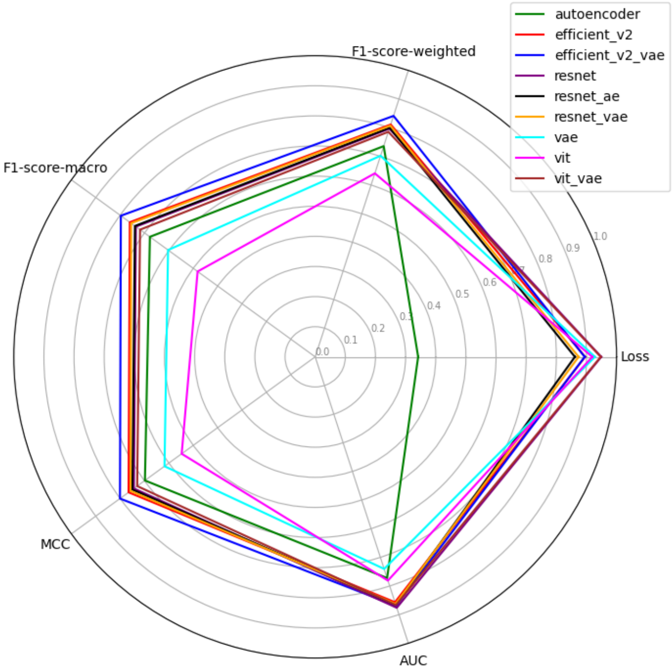
Results for test set

Model Name	**Loss**	**F1-score-weighted**	**F1-score-macro**	**MCC**	**AUC**
autoencoder	0.3416 ± 0.1907	0.7356 ± 0.0590	0.6779 ± 0.0551	0.6982 ± 0.0349	0.7723 ± 0.0480
efficient_v2	0.8953 ± 0.0227	0.8111 ± 0.0262	0.7602 ± 0.0353	0.7652 ± 0.0324	0.8567 ± 0.0254
efficient_v2_vae	**0.8942 ± 0.0134**	**0.8411 ± 0.0152**	**0.7970 ± 0.0209**	**0.8005 ± 0.0192**	**0.8693 ± 0.0272**
resnet	0.9489 ± 0.0012	0.7966 ± 0.0139	0.7353 ± 0.0175	0.7435 ± 0.0189	0.8760 ± 0.0076
resnet_ae	0.8617 ± 0.0304	0.7998 ± 0.0450	0.7392 ± 0.0694	0.7496 ± 0.0530	0.8656 ± 0.0144
resnet_vae	0.8739 ± 0.0242	0.8071 ± 0.0147	0.7547 ± 0.0189	0.7573 ± 0.0188	0.8627 ± 0.0127
vae	0.9280 ± 0.0012	0.7019 ± 0.0128	0.6033 ± 0.0206	0.6179 ± 0.0170	0.7404 ± 0.0070
vit	0.9194 ± 0.0164	0.6406 ± 0.0743	0.4821 ± 0.1327	0.5475 ± 0.1007	0.7810 ± 0.0577
vit_vae	0.9475 ± 0.0014	0.7856 ± 0.0151	0.7183 ± 0.0216	0.7300 ± 0.0194	0.8666 ± 0.0091

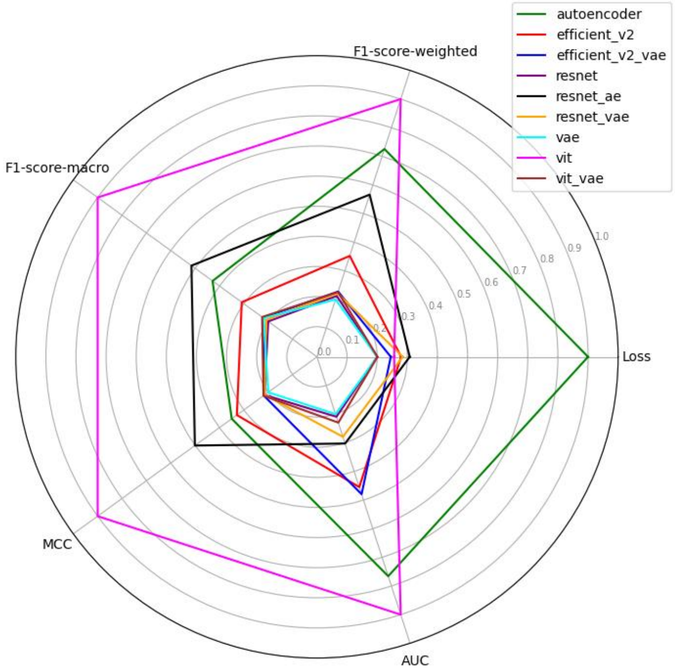
Higher values indicate better performance across all metrics

Loss: 1 – [cross entropy loss / max(cross entropy loss)]

Empirical results (test set)

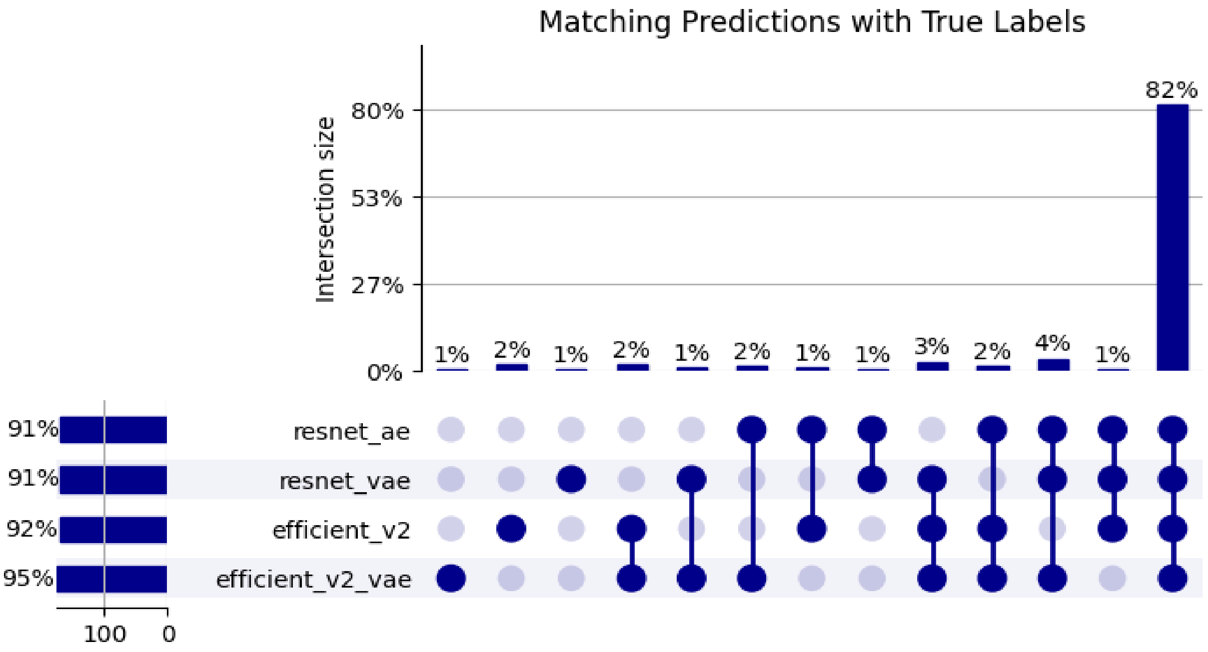


Average

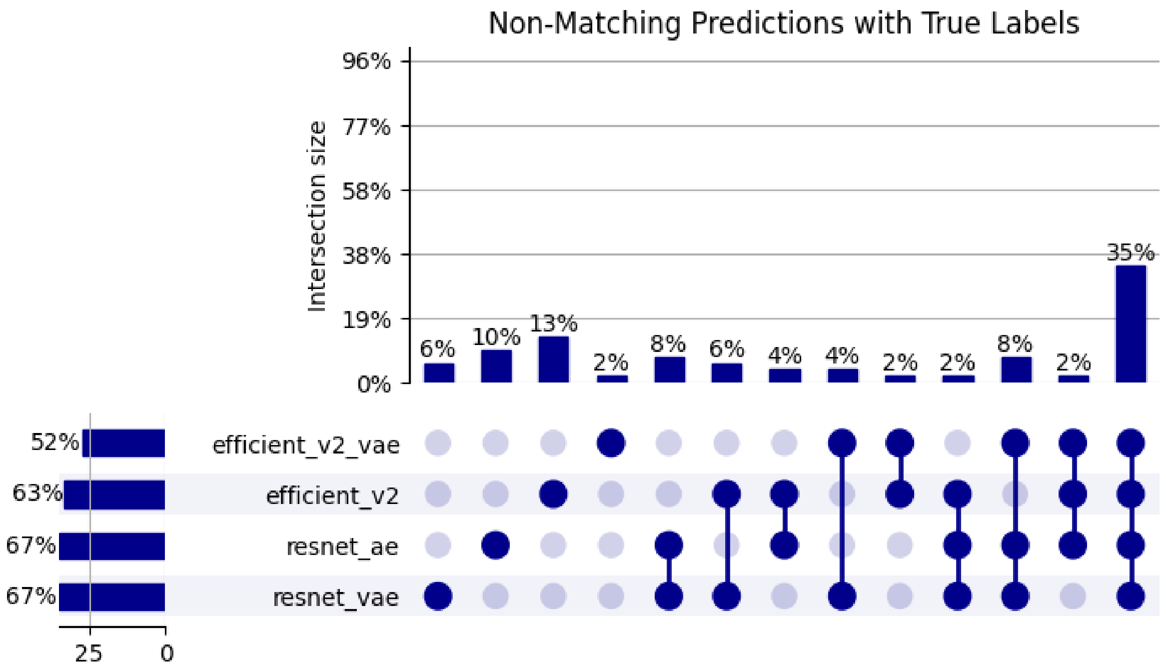


Standard Deviation

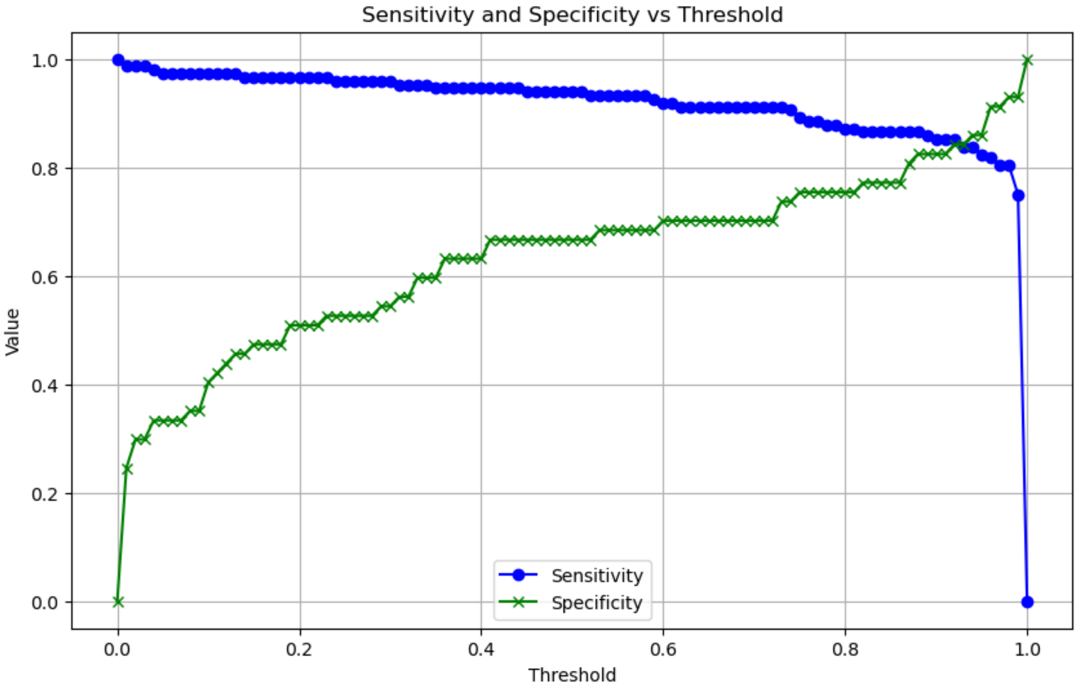
Discussion (matching prediction)

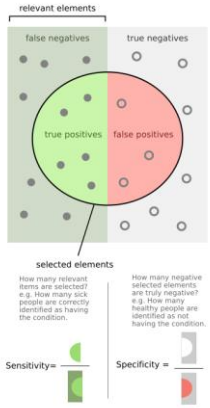


Discussion (non-matching prediction)

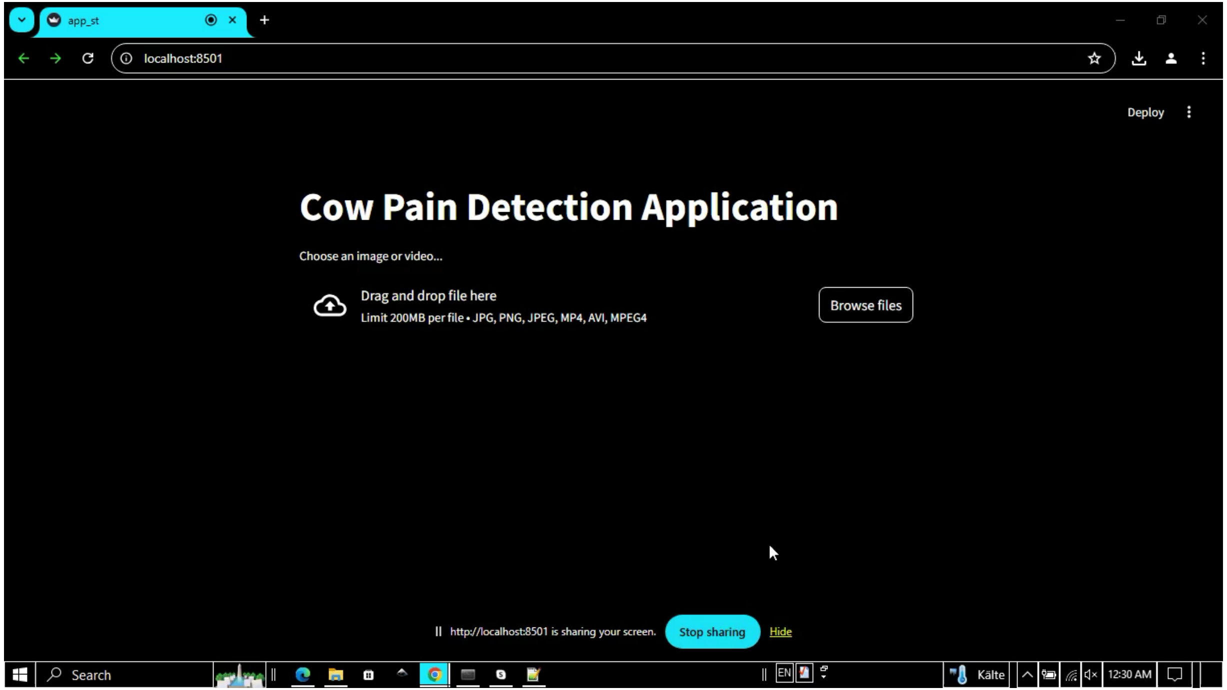


Sensitivity vs. Specificity





App demo


app_st
localhost:8501
Deploy
Cow Pain Detection Application
Choose an image or video...
Drag and drop file here
Limit 200MB per file • JPG, PNG, JPEG, MP4, AVI, MPEG4
Browse files
http://localhost:8501 is sharing your screen.
Stop sharing
Hide
Search
EN
Kälte
12:30 AM

Conclusion and Outlook

- *Efficientnet_v2_vae* model consistently outperforms others
- *vae* component of model captures the pain feature effectively from the visuals of cow.
- For a robust model selection, we conducted a Bayesian statistical comparison. While no pair of models was found to be practically equivalent across all metrics, vit_vae and resnet_ae showed equivalence in AUC.
- In contrast to transformer model, convolutional-based models deliver both accuracy and stability
- Misclassifications, segmentation, improving image quality, and expanding the dataset.
- Extending this work to video-based datasets could leverage spatio-temporal information for real-time pain detection at dairy farms.

References

1. Theresa Tschoner. “Methods for pain assessment in calves and their use for the evaluation of pain during different procedures—A review”. In: Animals 11.5 (2021), p. 1235.
2. Karina Bech Gleerup, Pia Haubro Andersen, Lene Munksgaard, and Björn Forkman. “Pain evaluation in dairy cattle”. In: Applied Animal Behaviour Science 171 (2015), pp. 25–32.
3. Niek Andresen, Manuel Wöllhaf, Katharina Hohlbaum, Lars Lewejohann, Olaf Hellwich, Christa Thöne-Reineke, Vitaly Belik, “Towards a fully automated surveillance of well-being status in laboratory mice using deep learning”, PLoS ONE (2020) 15(4): e0228059.
4. Shari Cohen and Thierry Beths. “Grimace scores: tools to support the identification of pain in mammals used in research”. In: Animals 10.10 (2020), p. 1726.
5. J Reid, AM Nolan, JML Hughes, D Lascelles, P Pawson, and EM Scott. “Development of the short-form Glasgow Composite Measure Pain Scale (CMPSSF) and derivation of an analgesic intervention score”. In: Animal welfare 16.S1 (2007), pp. 97–104.

Pain detection in cows using deep learning analysis of images

Alisa Sergeeva[1,+], Suresh Kumar Choudhary[1,+], Kerstin E. Müller[3,×], and Vitaly Belik[1,×,*]

[1]System Modeling Group, Institute of Veterinary Epidemiology and Biostatistics, School of Veterinary Medicine, Freie Universität Berlin, Berlin, 14163, Germany
[2]Clinic for Ruminants, Freie Universität Berlin
[*]Corresponding author: vitaly.belik@fu-berlin.de
[+]these authors contributed equally to this work
[×]these authors contributed equally to this work

ABSTRACT

A comprehensive manual scoring system for pain assessment in bovines was recently established. On the other hand, the utility of deep learning technology for pain recognition in animals has been proven across various species. In this study, based on a custom bovine dataset from a clinic including approximately 1000 manually scored images named as (*CattleCareDataset*), we demonstrate that deep learning algorithms can be successfully employed for pain detection. We proposed a combined approach involving an EfficientNet-based model (`efficient_v2`) and a feature embedding model to tackle this problem. Nine models were trained, including `resnet`, `efficient_v2`, and Vision Transformer (`vit`), both as standalone models and in combination with feature embedding components such as autoencoders and Variational Autoencoders (vae). Among these, the EfficientNet-based model with`vae` (`efficient_v2_vae`) demonstrated superior performance across most metrics, including Loss, F1-score-weighted, F1-score-macro, Matthews Correlation Coefficient (MCC), and Area Under the Curve (AUC), offering stability and reliability for the pain detection task.
On the test dataset, `efficient_v2_vae` achieved an F1-score (weighted) of 0.8411 ± 0.0152 and an AUC of 0.8693 ± 0.0272. Models incorporating the `vae` component, such as `efficient_v2_vae` and `resnet_vae`, outperformed their standalone counterparts, highlighting the `vae`'s effectiveness in capturing pain-related visual features. Convolutional-based models proved more suitable for this classification task compared to transformer-based models like `vit`, which showed limitations when applied to this smaller dataset.
To further validate model performance, we conducted a Bayesian statistical comparison using correlated t-tests, Mixture Bayesian Model (MBM), and Multivariate Bayesian Model (MvBM). While no pair of models was found to be practically equivalent across all metrics, `vit_vae` and `resnet_ae` showed equivalence in a single metric (AUC). These insights underscore the importance of robust model selection for this domain.
The study's practical implications include assisting veterinary hospitals in pain prediction, significantly reducing veterinarian time for diagnosis, and enabling early intervention to improve recovery outcomes. We engineered a web-based application that allows stakeholders to upload cow images or videos and receive pain predictions with explanatory insights. This tool enhances veterinary diagnostics, supports veterinarians by automating assessments, promotes animal welfare, and broadens AI accessibility for dairy farmers.
Future research can build on this foundation by addressing misclassifications, improving image quality, and expanding datasets. Extending this work to video-based datasets could leverage spatio-temporal information for real-time pain detection at dairy farms. This integration of technology and veterinary medicine offers transformative potential for improving livestock care and management.