Open-Sheep-Face: A Comprehensive Application for Sheep Face Analysis and Pain Estimation

* Zejian Feng  
School of Computing Science  
University of Glasgow  
Glasgow, UK  
2535290F@student.gla.ac.uk

* Martina Karaskova  
School of Computing Science  
University of Glasgow  
Glasgow, UK  
2478724k@student.gla.ac.uk

Marwa Mahmoud  
School of Computing Science  
University of Glasgow  
Glasgow, UK  
marwa.mahmoud@glasgow.ac.uk

Abstract—Sheep, being vulnerable prey species, frequently conceal overt signs of pain, which can lead to neglect of their welfare. These signs of pain are usually indicators of diseases that can escalate or becomes contagious. Recent work have developed computational methods for automatic detection of pain from sheep faces for early estimation of diseases. This paper presents an open source tool ‘Open Sheep Face‘ that implements various techniques grounded in prior research as a comprehensive pipeline for sheep face detection, landmarks localisation, pose estimation and pain level estimation.

Index Terms—Sheep welfare, pain estimation, application

I. INTRODUCTION

This paper details a comprehensive, automated pipeline that leverages machine learning and computer vision techniques to detect pain in sheep by analyzing their faces. We present the step-by-step construction of this pipeline, from face detection, head pose estimation, facial landmark localization, and pain estimation. The innovation lies in the integration of these standalone models into a functional application designed to automate the process of sheep face analysis and pain detection, which can be instrumental for veterinarians and sheep handlers. With a specific focus on its application, this paper aims to provide an overview of how the models are developed, validated, and then combined into an end-to-end solution for identifying and analyzing pain in sheep from images or videos.

II. METHODOLOGY

A. Face Detection

YOLOv5 [1] was retrained to enable sheep face detection using the same sheep dataset as [2]. This dataset contains accurate bounding box labels for sheep faces, which were utilized during the training process. Initially, the dataset are pre-processed with several steps, such as auto-orientation, resizing to 640x640, and conversion to grayscale. Subsequently, additional data is generated through data augmentations. Image augmentations involved horizontal flipping, random rotation within the range of -15° to +15°, and shearing between -15° and +15°. To further enhance the dataset, the bounding boxes were horizontally flipped, and the brightness was adjusted by -25% to +25% during augmentation. The dataset are split into training set, validation set and test set respectively.

ERT [9] method was employed to identify landmarks on sheep, which required ground-truth values for the absolute yaw angle, bounding box, and facial landmark coordinates to train the model accurately. The ERT method is a decision tree used for regression tasks that integrates multiple regression tree algorithms into a single model for prediction. The cascade of regressors approach was chosen to implement ERT due to its high accuracy rate in detecting facial landmarks in humans. An existing Python implementation of the ERT algorithm for detecting human facial landmarks, created by Xiao (2019) [10], was used. The provided code creates a cascade forest by iteratively building a sequence of random forests, where each forest is trained to predict the error of the previous forest. The output of the final forest is then added to the initial predictions of the first forest to obtain the final prediction.

B. Head Pose Estimation

Transfer learning is utilized to create a model capable of estimating sheep head poses by leveraging a deep network originally designed for human head pose estimation. The Hopenet network [4] is selected for its specific focus on head pose estimation and its superior performance among the networks trained in [4]. As the base model, a pre-trained Hopenet model trained on the 300W-LP dataset [5] is employed. To establish the ground-truth head pose for images in the sheep facial landmarks in the wild (SFLW) dataset, a 3D base landmark model is created manually, featuring a neutral head pose (0 yaw, pitch, and roll) and an average head shape [6]. A RANSAC [7] based method implemented by OpenCV [8] is then used to solve the perspective-n-point (PnP) problem, recovering the approximate head pose with the 3D points of the manually generated landmark model and the 2D annotated landmarks for each image [6]. The base model is fine-tuned on the SFLW-NCA dataset [13] with additional augmentation. Here, SFLW-NCA dataset is the SFLW dataset applied with horizontal mirroring, TPS warping and rotation augmentations, with the number of times each original image is used weighted by negatively correlated augmentation (NCA) [2]. The augmentation includes randomly flipping the input images in the x-direction and translating the image in x and y directions randomly by no more than 7% [4].

C. Facial Landmark Localization

ERT [9] method was employed to identify landmarks on sheep, which required ground-truth values for the absolute yaw angle, bounding box, and facial landmark coordinates to train the model accurately. The ERT method is a decision tree used for regression tasks that integrates multiple regression tree algorithms into a single model for prediction. The cascade of regressors approach was chosen to implement ERT due to its high accuracy rate in detecting facial landmarks in humans. An existing Python implementation of the ERT algorithm for detecting human facial landmarks, created by Xiao (2019) [10], was used. The provided code creates a cascade forest by iteratively building a sequence of random forests, where each forest is trained to predict the error of the previous forest. The output of the final forest is then added to the initial predictions of the first forest to obtain the final prediction.
D. Pain Estimation

To assess pain levels in sheep, HOGs [11] feature descriptor was utilized. HOG generates a feature vector by calculating gradients’ magnitudes and angles within an image, where an input image is partitioned into small sections and pixels are grouped into small cells. This process helps to calculate gradients, and orientation bins are formed. To identify pain indicators, the prominent facial features of sheep, such as both ears, eyes, and mouth regions, were identified, and the images were cropped to isolate these regions [12]. Then, these regions were converted into numpy representation, and HOG was applied to compute feature vectors for these regions. The feature vectors were mapped to a single sheep and flattened to create a representation vector. In addition to these features, geometric features such as the global angle of each ear and the distance between the ear root landmarks were also extracted and included in the analysis.

III. Evaluation

A. Face Detection

The YOLOv5-based sheep face detector is trained using cross-validation. Evaluation metrics recommended by YOLOv5 [1], including precision, recall, F1 score, mAP@0.5, and mAP@0.5:0.95, are utilized [3]. With 5-fold cross-validation, the sheep face detector achieves an average precision of 0.980, recall of 0.979, F1 score of 0.980, mAP@0.5 of 0.993, and mAP@0.5:0.95 of 0.93.

B. Head Pose Estimation

Training is performed with 16-batch increments over 16 epochs using an initial learning rate of 0.0001. The model is fine-tuned, and the one with the lowest validation loss is chosen as the sheep head pose estimation model. Mean Absolute Error (MAE) [4], Pearson’s Correlation Coefficient (PCC), and Sign Agreement metric (SAGR) [16] are employed for evaluation [13]. On the SFLW dataset, the fine-tuned model achieves an average MAE of 6.58, PCC of 0.75, and SAGR of 0.80 for the three angles (yaw, pitch, roll).

C. Facial Landmark Localization

The PI-ERT model is optimized through grid search, utilizing Mean Squared Error (MSE) and Mean Normalized Error (MNE) metrics for evaluation. Augmentation techniques, including TPS warping, are applied to the sheep dataset [13]. The PI-ERT model, with 10 forests and 500 trees each, achieves a Mean Normalized Error of 0.044, comparable to the state-of-the-art model by Hewitt and Mahmoud (2019) [13].

D. Pain Estimation

A Support Vector Machine (SVM) [14] classifier is trained for pain estimation in sheep, evaluated using accuracy, precision, recall, and F1 score. Dataset augmentation techniques, including image and landmark mirroring, are applied. The SVM model achieves a precision score of 83%, matching the state-of-the-art model [15], with recall, F1, and accuracy scores reaching 80% for all metrics.

E. Open Sheep Face Application

The Open Sheep Face application is a comprehensive pipeline that automates the process of detecting sheep faces, analyzing sheep head pose, predicting landmarks, and providing pain estimation results with visual displays. Users can conveniently obtain pain estimation results by simply uploading an image from their local device. The application then automatically analyzes the input and generates results that include landmarks and bounding boxes drawn around the detected sheep faces, all of which are displayed within the application interface. In cases where an image contains multiple sheep faces, users have the option to examine each individual sheep within the image using the application. For image analysis, the application provides results including the confidence level of the sheep face detection, pain estimation and whether the sheep is probably in pain or not within seconds (in the GPU version). An illustrative example demonstrating the application’s functionality with an image containing multiple sheep faces is depicted in Figure 1.

Fig. 1. Image with multiple sheep faces processed by the application

IV. Conclusion

In conclusion, we have presented a comprehensive, automated pipeline for detecting and analyzing sheep faces through an integrated application ‘Open Sheep Face’. This application encapsulates multiple complex processes, including face detection, head pose estimation, facial landmark localization, and pain estimation, to deliver an accessible solution for veterinary care and animal behavioral studies.

Our robust models, trained and fine-tuned using various techniques, exhibited competitive performance in respective tasks. The combined application’s ability to automatically analyze images and videos, while providing quick, precise results, marks a significant step forward in the automation of animal pain detection. With this application, the process of detecting pain in sheep becomes more accurate, efficient, and user-friendly, highlighting the transformative potential of integrating computer vision and machine learning techniques into the veterinary and animal care sector. For future work, video analysis within this application will be added, which will output the processed videos with estimated results. The code and application are made available as open source at https://github.com/zejian1gla/Open_Sheep_Face.
REFERENCES


