Cow Identification Network Trained with Similarity Learning

Alexander Ulrichsen[1], Brian Lee[2], Paul Murray[1], Stephen Marshall[1], Mark Rutter[3]
[1]University of Strathclyde, Glasgow, Scotland
[2]Peacock Technology Ltd, Stirling, Scotland,
[3]Harper Adams University, Newport, UK.
alexander.ulrichsen.2015@uni.strath.ac.uk

Abstract

Cow identification is a key phase in automated processing of cow video footage for behavioural analysis. Previous cow identification works have achieved up to 97.01% accuracy on 45 cows and 94.7% on datasets containing up to 200 cows. This paper presents new results from applying similarity learning to a cow identification Convolutional Neural Network on a group of 537 cows. Our method achieves identification accuracy of up to 99.3% and generalizes well to new cows, eliminating the need for retraining every time a new cow is added to the heard.

Introduction

Some work has been done to perform cow identification (Kobayashi, 2018; Yao et al., 2019; Zin et al., 2020) which take various approaches. Of this related work, (Kobayashi, 2018) achieves the highest accuracy of 97.01% identifying cows using a Convolutional Neural Network (CNN) classifier. The drawback of classification CNN approaches is that resulting network representations do not necessarily generalize well to new classes, meaning that if new cows are to be introduced, the entire network will require to be retrained - not just the final classification layer. This work focuses on training cow pattern identification CNNs with similarity learning to overcome these limitations.

Similarity learning is an alternative training approach to the standard classification training, first introduced in Google's Facenet paper (Schroff, Kalenichenko, & Philbin, 2015). Similarity learning trains a network to produce embeddings that are close in terms of Euclidean distance if images are of the same class and further apart if they are of different classes. Classification is done on the embeddings by using a more basic classifier such as a K Nearest Neighbours (KNN).

Dataset

Training data was acquired by positioning an IP camera above the radio frequency identification (RFID) scanner in the rotary milking parlour within the target farm. Each new RFID detected by the scanner trigged the acquisition of three images to capture the cow at different perspectives. The acquired images were passed through a YOLOv3 model (Redmon & Farhadi, 2018) which generated bounding boxes on the cows present in the image. The central localised cow bounding box from each image was then saved in folders corresponding to its RFID value. Human verification was required to ensure all cow images were in the correct folder locations.

All image pixel values were min max normalized between 0 and 1. The images were cropped to remove the RFID antennae and resized to 224x100. Of the 537 cows in the dataset, the first 500 were used for training and the final 37 were reserved "unseen" testing. Each individual cow had 15 images and of the 500 cows selected for training, 10 images from each were used for training, 3 were used for validation and 2 were reserved for testing.

This is a peer-reviewed, accepted author manuscript of the following paper: Ulrichsen, A., Murray, P., Marshall, S., Lee, B., & Rutter, M. (Accepted/In press). Cow identification network trained with similarity learning. Paper presented at European Conference on Precision Livestock Farming, Vienna, Austria.

The model used was ResNeXt50_32x4d (Xie, Girshick, Dollár, Tu, & He, 2017) due to its cardinality property which improves performance without additional parameters. The model was trained with an initial learning rate of 0.00001 and the learning rate was reduced by a factor of 0.1 every 200 epochs. A hard triplet selection strategy was employed on batches of 40 with 10 unique cows. The dataset was augmented with random: rotation +/-10 degrees; erasing; gaussian blur; colour jitter; auto contrast; perspective. The model was trained for 349 epochs.

Results

All images from the training set were embedded and used to fit a KNN classifier with K=5. Various combinations of "seen" cows (first 500 cows used to train the embedding model) and "unseen" cows (final 37 cows not used to train the embedding model) were used for testing to examine the performance on trained and untrained classes. Each of the KNN sets utilise the first 10 images from each relevant cow and each of the test sets utilise the final two images from each relevant cow. Results can be seen in Table 1.

Table 1 - Cow Identification Accuracy Results.

KNN set (first 10 images per cow)	**Test set** (final 2 images per cow)	**Accuracy (%)**
Seen cows	Seen cows	99.3
Seen + unseen cows	Seen + unseen cows	99.16
Seen + unseen cows	Unseen cows	97.3

Conclusion

This work has shown that similarity learning can provide cow identification performance above 99% on a data set of 537 cows. Our results show that our model generalises well to new classes, it is therefore possible to add new classes to the KNN classifier without the need to retrain the embedding model. Although not directly comparable to other works, our model boasts significantly higher identification accuracy on a dataset with significantly more classes.

References

Kobayashi, T. T. Z. M. I. C. N. P. P. T. H. H. I. (2018). Image Technology based Cow Identification System Using Deep Learning. *IAENG*. Retrieved from http://www.iaeng.org/publication/IMECS2018/IMECS2018_pp320-323.pdf

Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. *arXiv e-prints*, arXiv:1804.02767. Retrieved from https://ui.adsabs.harvard.edu/abs/2018arXiv180402767R

Schroff, F., Kalenichenko, D., & Philbin, J. (2015, 7-12 June 2015). *FaceNet: A unified embedding for face recognition and clustering.* Paper presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Xie, S., Girshick, R., Dollár, P., Tu, Z., & He, K. (2017, 21-26 July 2017). *Aggregated Residual Transformations for Deep Neural Networks.* Paper presented at the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Yao, L., Hu, Z., Liu, C., Liu, H., Kuang, Y., & Gao, Y. (2019). *Cow face detection and recognition based on automatic feature extraction algorithm*.

Zin, T. T., Pwint, M. Z., Seint, P. T., Thant, S., Misawa, S., Sumi, K., & Yoshida, K. (2020). Automatic Cow Location Tracking System Using Ear Tag Visual Analysis. *Sensors, 20*(12). doi:10.3390/s20123564