



**HAL**  
open science

# MobiCFNet: A Lightweight Model for Cattle Face Recognition in Nature

Laituan Qiao, Yaojun Geng, Yuxuan Zhang, Shuyin Zhang, Chao Xu

► **To cite this version:**

Laituan Qiao, Yaojun Geng, Yuxuan Zhang, Shuyin Zhang, Chao Xu. MobiCFNet: A Lightweight Model for Cattle Face Recognition in Nature. 5th International Conference on Intelligence Science (ICIS), Oct 2022, Xi'an, China. pp.386-394, 10.1007/978-3-031-14903-0\_41 . hal-04666448

**HAL Id: hal-04666448**

**<https://hal.science/hal-04666448v1>**

Submitted on 1 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

# MobiCFNet: A Lightweight Model for Cattle Face Recognition in Nature

Laituan Qiao, Yaojun Geng<sup>✉</sup>, Yuxuan Zhang, Shuyin Zhang, and Chao Xu

College of Information Engineering, Northwest A&F University, Shaanxi, China  
heiheiheit@foxmail.com

**Abstract.** In smart livestock, precision livestock systems require efficient and safe non-contact cattle identification methods in daily operation and management. In this paper, we focus on lightweight Convolutional Neural Network (CNN) based cattle face identification in natural background. Particularly, we first construct a fine-grained cattle recognition dataset with natural background. Then, we propose a lightweight CNN model MobiCFNet, containing a two-stage method that can realize one-shot cattle recognition. Finally, a series of experiments are conducted to validate the effectiveness of our proposed network.

**Keywords:** Cattle recognition · Lightweight Convolutional Neural Network · Cattle dataset.

## 1 Introduction

Intelligent agriculture has become a better option for growing agricultural production, improving resource management, and saving labor costs in recent years [4]. Precision livestock and smart livestock are a significant aspects of intelligent agriculture that can realize intelligent management of animal farming, improve the quality of meat products and reduce the amount of labor [7]. Accurate identification of individual cattle in intelligent pasture management is an essential problem to be solved in food traceability [14], false insurance claims [9], vaccination [11], and disease prevention [5].

Research on traditional cattle identification methods, for example, ear tags, nameplates, and tattoos began in the 1960s[10]. However, these methods may cause permanent damage to cattle and reduce accuracy due to loss of flag (i.e., ear tags, nameplates, and tattoos)[2]. In the 1980s, the emergence of radio frequency identification(RFID) technology improved the accuracy and enhanced the intelligence of identification, but also revealed high-cost and low-security weaknesses. With the rapid development of Deep Learning (DL) research, especially CNN, cattle identification methods based on DL alleviated mentioned problems existing in the previous recognition system [12].

At present, there are some difficulties in DL-methods based cattle identification research. Firstly the complex environment of the pasture adds uncertainty to the collected images. Then, collecting a rich dataset is a labor-intensive and time-consuming task due to the limitation of pasture size and cattle population.

Finally, the similarity in appearance between different cattle makes accurate identification difficult.

To settle these difficulties, we use a modified MobileNet V2 to complete the individual recognition of cattle based on cattle facial features. We introduce the Global Deep Convolutional (GDConv) layer that enables the network to extract cattle face features more fully. The introduction of Large Margin Cosine Loss (LMCL) makes the difference between classes increase and reduces the probability of model misjudgment. Our contribution could be summarized as follow: first, we construct a Qinchuan cattle and cows natural background dataset based on Northwest A&F University posture, containing 35000 images of 82 cattle. And then we propose a lightweight CNN called MobiCFNet and introduce a two-stage method that can realize one-shot cattle recognition. In the end, we compare conventional methods with our proposed methods on our dataset and show the effectiveness and robustness of our methods.

The rest of this paper is structured as follows: Section. 2 introduces the work related to cattle recognition, Section. 3 introduces the constructed dataset, Section. 4 describe the proposed method to solve the problem of cattle face recognition, Section. 5 summarizes the specific experimental results and discusses them, and Section. 6 is the conclusion of this study.

## 2 Related Work

Cattle identification has become a global issue involving scientific management of postures, food safety, false claims insurance, and more.

Since the 21st century, computer vision-related methods have been gradually applied in cattle recognition. Authors in [1] proposed a cattle face representation model based on local binary pattern texture features and trained Thirty cattle’s normalized gray face images separately. Experiments show that the LBP texture descriptor has good accuracy but low robustness.

In recent years, CNN has made a series of breakthroughs in computer vision, such as face recognition, target detection, and image classification. Authors in [8] proposed a CNN method to identify individual cattle, focusing on the pattern features of the cattle muzzle. In several test sets, the recognition accuracy of this method reaches 98.99%. The authors of [12] has designed a network framework that combines the advantages of CNN and Long Short-Term Memory to identify individual cattle. The dataset uses a fixed Angle rear view of the cattle. However, the dataset of the above two identification methods are not collected in natural scenes.

The continuous advancement of computer performance means that we can train CNN with larger dataset, thus significantly enhancing the recognition accuracy. Authors in [6] proposed a model of individual identification of dairy cows using CNN in the actual cattle farm environment. The model was trained using 60,000 trunk images of 30 cattle, and 90.55% of the 21,730 images tested were accurately identified. But this method is not applicable in pasture with a high number of cattle.

### 3 Fine-Grained Face Individual Recognition Dataset

In this section, we describe the workflow of constructing our Cattle Face Recognition Dataset.

The individual cattle breeds we identified are mainly Qinchuan cattle and dairy cows. The body of Qinchuan cattle is yellow and red, so it is challenging to identify the cattle from its body. However, Qinchuan cattle have more facial features compared with the body, including hair color, pattern, face shape, and other characteristics, as shown in Fig. 1. Therefore we choose to construct the individual recognition dataset of cattle face features and study the individual recognition based on cattle face features.



Fig. 1. The cattle face are diversified



Fig. 2. Large intra-class differences and small inter-class differences in the dataset

We visit cattle farms several times to collect data in different weather conditions and at different periods. Abundant cattle front or side face images under different backgrounds were taken to ensure a more comprehensive background of images in the dataset and a more diverse angle of cattle face.

A dataset of 82 Qinchuan cattle and dairy cows with 35,000 cattle face images was constructed, which are RGB images of 224x224. Then, it was divided into the training set, verification set, and test set in a ratio of 7:2:1. We found that different images of the same cow differ greatly. The first row of Fig. 2 shows three different images of a specific cow, with significantly different posture, background, and angles. Meanwhile, images of different cows at the same angle may

vary slightly. The second row of Fig. 2 shows three images of different cows, which are relatively similar in appearance from the same angle. The constructed dataset has significant differences within classes and slight differences between classes, which brings great challenges to the cattle recognition work.

In conclusion, we constructed a fine-grained cattle recognition dataset which consists of dairy cows and Qinchuan cattle face images with complex and diverse natural backgrounds that collect from various angles.

## 4 Proposed Method

This section describes our proposed cattle face identification network MobiCFNet. First, we introduce CNN and Lightweight CNN. Then, the Large Margin Cosine Loss (LMCL) in our work is described. Finally, we give the entire network architecture and describe the proposed feature dictionary based recognition method.

### 4.1 CNN and Lightweight CNN

Since AlexNet unexpectedly won the ImageNet competition in 2012, there has been a growing amount of innovative work in this area. In particular, ResNet, proposed by Kaiming He et al. in 2015, reduced the recognition error rate on the ImageNet dataset to 3.57% by introducing residual blocks and deepening the design of network layers.

However, the model structure of CNN becomes more complex to pursue high accuracy, which causes difficulty in transplantation to mobile devices or Internet of Things (IoT) devices. To this end, Google developed MobileNet series CNN models, which use deep separable convolution to decompose standard convolution into depthwise convolution and pointwise convolution to reduce computation and amount of parameters [3]. MobileNet V2 model adopts the Inverted Residuals structure, firstly elevates the features' dimension, then carries out convolution operation, and finally reduces the dimension. Besides, Linear Bottleneck is used instead of non-linear activation function to avoid information loss. Experiments show that the precision and portability of the model are well balanced [13].

### 4.2 Large Margin Cosine Loss

LMCL is a loss function that rethinks softmax loss from the view of cosine perspective, firstly proposed for face recognition in [15]. In the output layer of a CNN, we have

$$f_j = W_j^T x + B_j, \quad (1)$$

where  $W_j$  is the weight and  $B_j$  is the bias, set  $B_j$  to zero, we can derive that

$$f_j = W_j^T x = \|W_j\| \|x\| \cos \theta_j, \quad (2)$$

where  $\theta_j$  is the angle between  $W_j$  and  $x$ . Face recognition is realized by cosine similarity between feature vectors. Set  $\|x\|$  to  $s$  and  $\|W_j\|$  to 1 as

$$L_{ns} = \frac{1}{N} \sum_i -\log \frac{e^{s \cos(\theta_{y_i, i})}}{\sum_j e^{s \cos(\theta_{j, i})}}. \quad (3)$$

In this way, LMCL forces classifier to learn separable features in angle space by normalizing feature vectors and removing radioactive variables. The second contribution of LMCL lies in the incorporation of cosine margin  $m$  into cosine formula as

$$L_{lmc} = \frac{1}{N} \sum_i -\log \frac{e^{s(\cos(\theta_{y_i, i})-m)}}{e^{s(\cos(\theta_{y_i, i})-m)} + \sum_{j \neq y_i} e^{s \cos(\theta_{j, i})}}, \quad (4)$$

which reduces cosine similarity between feature vectors of different classes, forcibly increases the distance between classes, and increases the accuracy of recognition.

LMCL is a classifier designed for human face recognition. In LMCL, maximizing the variance between classes and minimizing the variance within classes is adopted. Cattle face also contains rich feature information. The facial contour, facial features, and hair texture of cattle are significant characteristics in cattle face identification. So migrating LMCL to cattle face identification based on cattle face features still works.

### 4.3 MobiCFNet

We propose a lightweight CNN model for cattle face recognition called MobiCFNet. Fig. 3 shows the architecture of MobiCFNet, we replace the last global average pooling layer and the fully connected layer of MobileNet V2 with GDConv layer and an LMCL classifier.

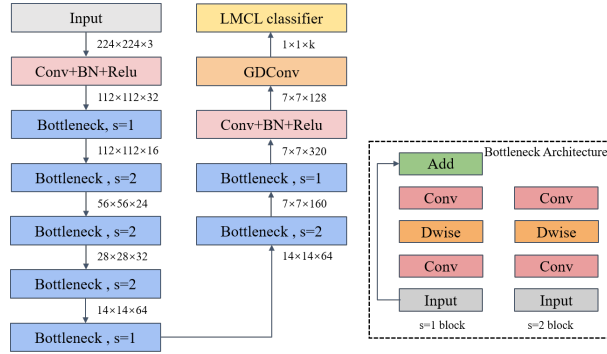
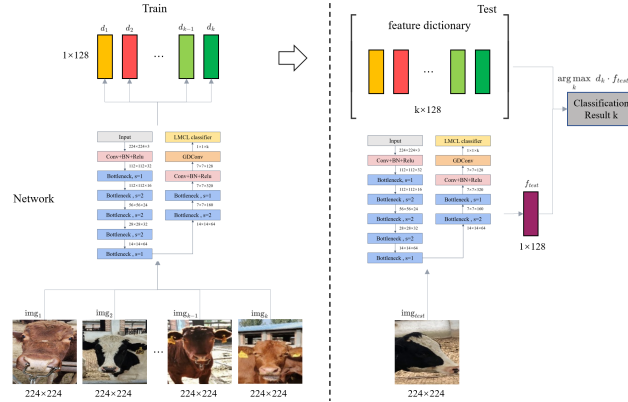


Fig. 3. MobiCFNet architecture

Compared with pooling layers, the advantage is that GDConv layer introduces attention mechanisms, which assign different regions in the input feature

map with different importance. Therefore, GDCConv layer contributes to extracting the features of essential regions in the image. The LMCL classifier in our work uses 128-dimensional features to distinguish 82 cattle in the dataset. Compared with traditional softmax loss, LMCL is more in line with the fine-grained classification tasks, forcing the interval between classes of the classifier to be greater than or equal to a certain hyperparameter  $m$ .

Based on MobiCFNet, we propose a flexible way of cattle individuals recognition, the workflow of which is plot in Fig. 4. First, a characteristic dictionary of cattle is constructed after training by combining feature vectors corresponding to 82 cattles in the dataset. The feature vector is the 128-dimensional output before LMCL classifier of MobiCFNet that use each cattle’s face image with rich features as input. When we need to classify an image, we pass the image through the feature extraction network to get a 128-dimensional feature vector. Then, normalized inner product vector between the feature vector the feature of each cow in the dictionary and then the cow with the largest inner product result is selected as the predicted result.



**Fig. 4.** Workflow of feature dictionary method in MobiCFNet

The principle is that LMCL migrates softmax to the angle space during network training and deliberately increases the interval between classes. The feature vectors of the same cattle image are clustered together in the hypersphere, and their normalized inner product results are close to 1. Besides, the most significant advantage of the feature dictionary construction method is that it enables the network to add or delete cattle after the training procedure while keeping the accuracy of cattle recognition at the same level. When we need to delete a cattle, we delete the characteristics of this cattle from the feature dictionary. If the data used in the training model is rich enough, we can assume that the network can robustly extract individual characteristics of cattle. Under this condition, we can easily add new cattle. We input a representative image of the new cattle through the feature extraction network and then add the feature to



the feature dictionary. Deleting cattle from the trained model is simple. We only need to delete the corresponding feature in the feature dictionary.

## 5 Experiment

We accomplish experiments on the server with Tesla V100 GPU and Core i7 8-core CPU. We implement MobiCFNet using PaddlePaddle DL framework. We verify cattle recognition ability of MobiCFNet in two cases: traditional end-to-end method and feature dictionary method. We compare MobileNet V2 with other proposed MobiCFNet on our cattle recognition dataset. In order to improve the reliability of comparison, we set batch size as 64, learning rate as 0.001, optimizer as Adam, and epoch numbers as 500. In order to verify the practicality of MobiCFNet, we ported it to an Android device and verified the experimental results in the pasture.

**Table 1.** Accuracy and Inference Time on Android Device.

Network	Accuracy(%)	Inference Time on Android Device
MobileNet V2 + softmax	84.57	28 ms
MobileNet V2 + LMCL	85.62	28 ms
MobiCFNet + softmax	86.38	24 ms
MobiCFNet + LMCL	87.44	24 ms

In the experiment, we take the cattle images as the input of MobiCFNet and connect the feature extraction network with the softmax classifier directly. In this case, 86.38% accuracy was achieved, which verified that MobiCFNet has strong feature extraction ability. In the feature dictionary experiment, the recognition accuracy of MobiCFNet reached 87.44%, performing better than MobileNet V2. Besides, we compare infer time of MobiCFNet on an Android device (Redmi K30 Pro) with MobileNet V2. All results are available in Table. 1.

In conclusion, the MobiCFNet proposed in our work can achieve higher accuracy and faster infer time than MobileNet V2 in the classification task of cattle face identification. In particular, our method with feature dictionary can improve the recognition accuracy and easily add and delete cattle in the pasture, which meets the actual need in the pasture.

## 6 Conclusion and Future Scopes

In this paper, we construct a fine-grained cattle recognition dataset and propose a lightweight convolutional neural network named MobiCFNet. Our proposed method can solve the problem of large intra-class differences and small inter-class differences in the dataset that can robustly recognize individual cattle images from different angles. In addition, we proposed a recognition method based on feature dictionary in which cattle can be added and deleted after the network is training. Feature dictionary based cattle recognition meets the needs of pasture

management. In the future, the proposed method can be combined with the target detection method to realize multi-cattle identification to realize end-to-end automatic pasture monitoring.

## References

1. Cai, C., Li, J.: Cattle face recognition using local binary pattern descriptor. In: 2013 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference. pp. 1–4. IEEE (2013)
2. Fosgate, G., Adesiyun, A., Hird, D.: Ear-tag retention and identification methods for extensively managed water buffalo (*bubalus bubalis*) in trinidad. *Preventive veterinary medicine* **73**(4), 287–296 (2006)
3. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., Adam, H.: Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861 (2017)
4. Hu, X., Sun, L., Zhou, Y., Ruan, J.: Review of operational management in intelligent agriculture based on the internet of things. *Frontiers of Engineering Management* **7**(3), 309–322 (2020)
5. Ismail, W.N., Hassan, M.M., Alsalamah, H.A., Fortino, G.: Cnn-based health model for regular health factors analysis in internet-of-medical things environment. *IEEE Access* **8**, 52541–52549 (2020)
6. Kaixuan, Z., Dongjian, H.: Recognition of individual dairy cattle based on convolutional neural networks. *Transactions of the Chinese Society of Agricultural Engineering* **31**(5) (2015)
7. Kim, W., Cho, Y.B., Lee, S.: Thermal sensor-based multiple object tracking for intelligent livestock breeding. *IEEE Access* **5**, 27453–27463 (2017)
8. Kumar, S., Pandey, A., Satwik, K.S.R., Kumar, S., Singh, S.K., Singh, A.K., Mohan, A.: Deep learning framework for recognition of cattle using muzzle point image pattern. *Measurement* **116**, 1–17 (2018)
9. Kumar, S., Singh, S.K., Dutta, T., Gupta, H.P.: A fast cattle recognition system using smart devices. In: *Proceedings of the 24th ACM international conference on Multimedia*. pp. 742–743 (2016)
10. Kumar, S., Singh, S.K., Singh, R., Singh, A.K.: *Animal Biometrics*. Springer (2017)
11. Manoj, S., Rakshith, S., Kanchana, V.: Identification of cattle breed using the convolutional neural network. In: *2021 3rd International Conference on Signal Processing and Communication (ICPSC)*. pp. 503–507. IEEE (2021)
12. Qiao, Y., Su, D., Kong, H., Sukkariéh, S., Lomax, S., Clark, C.: Individual cattle identification using a deep learning based framework. *IFAC-PapersOnLine* **52**(30), 318–323 (2019)
13. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 4510–4520 (2018)
14. Tang, Q., Li, J., Sun, M., Lv, J., Gai, R., Mei, L., Xu, L.: Food traceability systems in china: the current status of and future perspectives on food supply chain databases, legal support, and technological research and support for food safety regulation. *Bioscience trends* **9**(1), 7–15 (2015)
15. Wang, H., Wang, Y., Zhou, Z., Ji, X., Gong, D., Zhou, J., Li, Z., Liu, W.: Cosface: Large margin cosine loss for deep face recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. pp. 5265–5274 (2018)