Research on Dairy Cow Identification Methods in Dairy Farm

Liu Bo[1], Liu Yuefeng[1], Bao Xiang[1], Wang Yue[1], Liu Haofeng[1], Long Xuan[2]

ABSTRACT

Background: There is a large amount of occlusion in dairy farms and changes in daytime and nighttime lighting seriously affect the accuracy of traditional dairy cow identification.

Methods: This paper proposed a method to dairy cow identification in dairy farm. Firstly, Resnet50 was used to extract the pattern features of dairy cows and this paper fused the fourth and fifth scale features with semantic information. Secondly, on the basis of the triplet loss, it added label smoothing loss instead of softmax cross-entropy loss to further avoid over-fitting and reduced the intra-class distance with the center loss, which improved the mean average precision (mAP) by 8.9% compared with the initial Re-identification (Reid) and the joint optimized distance calculation formula was improved the single distance measurement in the triplet loss, using two distance measurement methods to effectively improve the recognition accuracy. Finally, it was combined with pruning operation to reduce the redundant parameters of the model.

Result: The experiment proves that the identification mAP of dairy cows reaches 81.2%, which is an improvement 10.9% compared to the baseline model mAP, providing a strong technical support for the subsequent dairy cow identification and target tracking system.

Key words: Feature fusion, Image recognition, Metric learning, Model compression, Object detection.

INTRODUCTION

The dairy cow identification and target tracking system is the basic module for scientific supervision of dairy cows using information technology and is a key technology for achieving quality benefits from dairy farming (Yu, 2016). Most dairy farms today still identify cows by wearing devices (Li *et al*., 2020; Guo *et al*., 2020; Zhao *et al*., 2020), which not only affects the health of cows and the quality of milk produced, but also increases the cost for dairy farmers. Therefore, choosing a no-contact and highly accurate artificial intelligence method to identify cows by computer instead of human is the trend.

The artificial intelligence method of cow identification task is divided into two approaches, one of which is a classification algorithm based on target detection. Zhao *et al*. (2015) proposed a method for individual cow identification which captured the side-view video of a cow walking in a straight line, calculated the rough outline of the cow using the inter-frame difference method and carried out segmental span analysis on its binary image to determine the structure. Wang *et al*. (2019) proposed a YOLOv3 detection algorithm for training and testing the cow datasets. They proposed a multi-scale target detection algorithm with a spatial pyramid pooling layer for cow identification. He *et al*. (2020) proposed an individual identification method for dairy cows based on the improved YOLOv3 depth convolution neural network. They used video frame decomposition technology to get the back image of cows and used pixel linear transformation method to enhance image brightness and contrast. Shen *et al*. (2020) collected images of cows' backs with a fixed camera and used a target detection model to locate cows. Then they constructed a Convolutional Neural Network to classify the target images to confirm dairy cow identity. Yang *et al*. (2021) constructed an improved YOLOv4 model based on

[1]School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China.

[2]School of Economics and Management, Southeast University, Nanjing 211189, China.

Corresponding Author: Liu Yuefeng, School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China. Email: 15947325205@163.com

fused coordinate information to identify cow faces and improved recognition speed. Zhang *et al*. (2021) implemented multi-objective tracking of cattle in a real-world scenario through an improved target tracking algorithm and with the LSRCEM-YOLO algorithm for cow identification to assess the health status of beef cattle and conduct behavior perception. Another approach to cow identification is through metric learning methods. They applied different features such as cows' lip lines, nose lines, retina (Andrew *et al*., 2020) and face features to obtain the final cow ID through feature distance metrics. Kumar *et al*. (2018) combined Convolutional Neural Network and Deep Belief Network methods to learn the extraction of cow nose lines texture features and introduced an automatic encoder framework for superimposed denoising to encode and decode cow individual nose region features for better feature representation. Gou *et al*. (2019) used Inceptionv2 to replace ZF network as the basic network of Faster R-CNN for the demand of multi cow detection scene accuracy and optimized the non-maximum suppression accordingly, which

greatly improved the recall rate of cow face recognition model. Cho *et al*. (2018) proposed a hybrid method in which rolling skew histogram and neural network techniques were fused to recognize patterns and identify cows in the milking rotary parlor of dairy farms. Zhao *et al*. (2022) reformulated the SoftMax as SoftMax-nB, proposed compact loss by combining Softmax-nB and distance metric learning, which highly enhanced the discriminative power of the features. Other scholars have also studied lightweight models for scientific cow feeding. Liu *et al*. (2022) proposed a detection method for cows feeding behaviour based on iterative magnitude pruning, using the YOLOv3 classification method to detect cow feeding behaviour and using iterative magnitude pruning to compress the model, effectively reducing the amount of model computation.

The above technologies are more aimed at the identification of a single cow, they only focus on the front face of the cow and require a short shooting distance. It is difficult to achieve automatic data collection in practical applications (Xu *et al*., 2020). Therefore, we proposed a method for dairy cow identification, in which the dairy cow target was detected by the YOLOv7 algorithm and extracted the targets' features using the imporved Resnet50 network (Zhuo *et al*., 2022; Wu *et al*., 2022; Xu *et al*., 2022). Then model obtain the final dairy cow ID after calculating the distance. The innovations of this paper are as follows: 1). Fusing the semantic features extracted from the last two scales in the Resnet50 network. 2). Using the label smoothing loss instead of the original softmax cross-entropy loss to make the triplet loss better train the complex cow IDs. In order to reduce the intra-class gap between dairy cows with the same ID and the same shooting angle, center loss was added. 3). Proposing a joint optimization distance calculation formula to calculate the similarity between features. The triplet loss calculation distance method was added to the cosine distance calculation from a single euclidean distance calculation. 4). Lighting the model based on iterative magnitude pruning method, which saved computational costs and further improved the recognition accuracy.

MATERIALS AND METHODS
Collect data materials and build datasets

(1) Construction of dairy cow target detection datasets. The datasets for this research method was derived from the video of 493 Holstein cows captured by a dairy farm camera in Baotou City, Inner Mongolia Autonomous Region in 2020. The video is 3840 segments of 45 minutes each, the video format is MPEG4, the height of the video frame is 1080 pixels, the width is 1920 pixels, the bit rate is 1639kb/s. We combined video images of dairy cows in dairy farm and in order to prevent the model training from over-fitting due to the similarity of dairy cow shapes in adjacent frames, converting the captured video into a picture by drawing one frame from eighty frames. The images that were not clear and had no dairy cows in the area to be recognized were eliminated. Finally, a total of 24210 data images were obtained after screening. Process all images through shuffling, randomly selecting 19730 as the training set, 2220 images as the validation set and 2260 images as the test set. (2) Construction of dairy cow identification datasets. Each dairy cow was cropped down and a total of 72020 dairy cow target images were obtained after screening. There were 493 cows in the images and the cow images were classified by ID. We randomly choice 293 cows as the training set with 38110 images and 200 cows as the test set with 33910 images.

The datasets of each cow in the training set and test set needed to be divided secondly. Each dairy cow was photographed at an angle of 360° and with 60° as one category, each dairy cow could be divided into six categories, representing six different angles of shooting. Finally, these 493 * 6= 2958 categories of pictures were renamed.

Dairy cow target detection methods

The initial input image was resized to 640*640*3. After the backbone layer, three different scales of feature maps were generated through the head. We choice the Adam as optimizer, the initial learning rate was 0.001, the momentum factor is 0.9, the weight decay was 0.0005 and the batch size was 32. Since this paper only needed to detect the dairy cow targets, one class of dairy cow targets was needed to be classified. Before training begins, 9 anchors arranged from small to large were obtained through k-means algorithm (Wang *et al*., 2022). Each ground-truth was matched with them and its width to width ratio and height to height ratio were calculated. The maximum ratio was taken and compared with the set threshold. If it was less than this threshold, it was considered a positive sample. Introducing the structure of REPVGG to improve the model performance by multi-way branching. An Aux-Head was added to work with Lead-head for model optimization. When the anchor was matched with ground-truth, three positive samples were assigned to Aux-Head and five positive samples were assigned to Lead-head with a loss weight ratio of 1:4. The final six-dimensional prediction values were obtained, which represented the border coordinates (x, y, w, h), the border confidence and the number of classes.

Cow identification methods

The Reid technique, or person Re-identification, is a sub-problem of image retrieval that uses the calculation of distance in metric learning to search for targets in an image gallery (Hermans *et al*., 2017). In this paper, the target-detected images were used as the Query and we randomly selected 200 cows as the Gallery which could be representative of the entire dataset, then the images of the Query are used for the metric calculation of features with the images in the gallery to find the IDs of the target-detected dairy cows in the gallery.

Backbone

The dairy cow identification model extracted features through the improved Resnet50 network and it did not require target

localization because the model was based on target detection. The high resolution feature information from the shallow network was not very useful for the recognition model, but might be used by the network as noise and affected the final recognition effect. Therefore, we only carried out feature extraction at different scales for the last two scales, setting the stride of the 1/32 scale to 1, so that the resolution of the last two scales could be kept consistent. Feature fusion refers to the fusion of feature information from different scales into a set of features through concat operation, which can better combine the features reflected by different scale features. The feature maps obtained after the last layer of convolution of the fourth scale and fifth scale were self-adaptively fused to make better use of the semantic information of the last two scale features, as shown in Fig 1.

The positive and negative sample pairs

The training method of positive and negative sample pairs seriously affects the accuracy of recognition. We set the batch size to 16 and the NUM_INSTANCE to 4, which meant randomly selecting 16 images from the prepared identity recognition datasets for processing in a mini-batch. These 16 images included 4 classes, with 4 images for each class of target. Randomly selected one from 16 images as a template sample, formed a positive sample pair with the same ID and a negative sample pair with different IDs. After data enhancement such as flip, rotation, crop, affine, calculated the distance between positive and negative samples, used the value of the maximum positive sample distance and the value of the minimum negative sample distance as hard example. Finally put it into the triplet loss for training.

Loss and distance calculation formula

This paper used a different triplet loss from the traditional Reid and proposed a joint training method with multiple loss. Center loss could reduce the intra-class gap, so adding it could reduce the gap between dairy cow features with the same ID and the same shooting angle. The data selected for the triple loss might not be uniformly distributed, which would lead to unstable performance of the model training process, slow convergence speed, easy over-fitting. We added label smoothing loss in the full connection layer, which could effectively solve the above problems, preventing the model from focusing only on the loss of correct label positions and also considering the loss of other incorrect label positions, increasing the generalization ability of the model. The triplet loss is as shown in formula (1). As described in (2) the positive and negative sample pairs, take the positive hard sample as d_p and the negative hard sample as d_n, set the margin to 0.25 and calculate the triplet loss. The center loss is as shown in formula (2). It simultaneously learned the deep feature center value of each cow class and reduce the distance between each cow class and its class center. The label smoothing loss is as shown in formula (3). The last layer of Resnet 50, which outputs the ID prediction logits of images, is a fully-connected layer with a

hidden size being equal to numbers of cows N (N= 493). Given an image, we denoted y as truth ID label and P_i as ID prediction logits of class i. The final loss is as shown in formula (4). After experimental verification, here β was 0.34.

$$L_{Tri} = [d_p - d_n + \alpha] + \quad \text{....(1)}$$

$$L_c = \frac{1}{2}\Sigma_{j=1}^{B}\left\|f_{tj} - c_{yj}\right\|_2^2 \quad \text{....(2)}$$

$$LID = \Sigma_{j=1}^{N} q_i \log(P_i) \begin{cases} q_i = 0, y \neq i \\ q_i = 1, y = i \end{cases} \quad \text{....(3)}$$

$$L = L_{ID} + L_{TRI} + \beta L_c \quad \text{....(4)}$$

d_p and d_n = Feature distances of positive pair and negative pair.
α = Set to 0.25.
y_j = Label of the th image in a mini-batch.
C_{yj} = Denotes the th class center of deep features.
f_{ti} = Denotes the feature of the th image.
B = Number of batch_size.
y = Truth ID label.
P_i = Indicates ID prediction logits of class i.

A joint optimized calculation of euclidean distance and cosine similarity distance was proposed. This is because the euclidean distance measures the absolute distance between features and it is directly related to the location coordinates between the points, whereas the cosine similarity distance measures the angle of the spatial vectors and it is more concerned with the difference in direction. Because we could not determine which measurement method was most effective, we propose the combined optimization distance calculation formula. After extracting features from the target and gallery images through backbone, the results passed in x_i and y_i. And by combining the two distances self-adaptively, it replaced the heavy and complex traversal method to find the optimal hyper-parameter value. For example, formula (5) is the euclidean distance formula and formula (6) is the cosine similarity formula. In formula (7), represent the weight of euclidean distance and cosine distance in the joint optimization distance formula, that is the impact of the two distance formulas on the recognition effect. First initialized α, β then normalized them and put them into the network as parameter to optimize. The final printout of the weights shows that α is 0.32 and β is 0.68.

$$d = \sqrt{\Sigma_{i=1}^{n}(X_i - y_i)^2} \quad \text{....(5)}$$

$$\cos\theta = \frac{\Sigma_{i=1}^{n}(X_i \times y_i)}{\sqrt{\Sigma_{i=1}^{n}(x_i)^2} \times \sqrt{\Sigma_{i=1}^{n}(y_i)^2}} \quad \text{....(6)}$$

$$Distance = \alpha d + \beta(1 - \cos\theta) \quad \text{....(7)}$$

d = Indicates euclidean distance between two vectors.
x_i = Represents detecting the ith dimensional feature vector of the dairy cow.
y_i = Represents the ith dimensional feature vector of the gallery dairy cow.

cosθ = Cosine similarity of two vectors. Distance represents the joint optimization distance calculation formula:
α = Euclidean distance weights.
β = Cosine distance weights.

Evaluation indicator

The metrics evaluated in this experiment are the Rank-k metric and the mAP metric. Take Rank-1 as an example, it means to count whether all query images are the same as their first returned result in the gallery, which means the accuracy of the first retrieval target of the image and the same calculation can obtain Rank5 and Rank10. mAP metric means average precision mean, which is a common evaluation metric in multi-target detection and multi-label image classification. It sums and averages the average precision AP (Average Precision) in multi category tasks, that is, it represents the accuracy of all retrieval results. The calculation formula is shown in formula (8):

$$AP = \frac{\Sigma \ Precision_c}{Images_c}$$

$$mAP = \frac{\Sigma_{k=0}^{c} AP_K}{c} \quad ....(8)$$

$Precision_c$ = Accuracy rate for a single category.
$Images_c$ = Number of images containing targets of this type.
AP_K = Average accuracy for a single category.
C = Total number of categories.
　　There are 493 dairy cows from 6 different angles, a total of 2958 categories.

Model lightweight methods

In this paper, the iterative magnitude pruning algorithm based on lottery hypothesis was used to filter the optimal sub-network from the dairy cow identification model. The experimental steps are as follows:

1). The network was initialized with pretrained weights obtained by training the Resnet50 network and the network was subjected to k times of gradient descent to obtain the weights W_0 and saved, where k was 0.1%-7% of the total training times.
2). Trained the dairy cow datasets and got the weight $W_T^{(1)}$ after convergence.
3). Created a mask $m^{(1)}$, set a fixed pruning rate and unstructuredly pruned the model $W_T^{(1)}$.
4). The weights W_0, obtained by performing k-gradient descent using the original network, reset the sparse network weights obtained by pruning.
5). Repeated step 2~4, pay attention to the final accuracy of each pruning model after convergence and ended pruning until the accuracy of the model begins to decline, then took the result with the highest pruning accuracy.

RESULTS AND DISCUSSION
Experimental system environment and parameter settings

The experimental operating system is Ubuntu 18.04, CPU is core i5-10300H at 3.6GHz, GPU is NVIDIA GeForce GTX 2080Ti × 1 and running memory is 11 GB. The dairy cow identification model is trained with 100 epochs, initial learning rate is 0.0035, warm-up cycle is 10 epochs, the weight_decay factor is 0.0005, the momentum factor is 0.9, the batch_size is 16 and the NUM_INSTANCE is 4.

Results of dairy cow target detection

The first stage of the dairy cow identification task is completed by the dairy cow target detection technique. The training samples were divided into only one class, the total period was set to 300 epochs, the early stop parameter was set to 75 epochs and the confidence size was set to 0.8. The experiment showed that the accuracy rate of the detection model is 96.1% and the recall rate is 94.2%. Because the detection task is only the first stage of the recognition task, the aim is to locate the dairy cow and detect the dairy cow target, the accuracy and recall of the target detection model only need to be maintained at 90% or more to fully satisfy the needs of the second stage of this study, so the detection model will not be further analyzed and discussed.

Results of dairy cow identification

Results of feature fusion. After the dairy cow target was detected using the dairy cow target detection technique, the cow features were extracted by the improved Resnet50 network and the extracted features of the fourth and fifth scales were self-adaptively fused. The experimental results are shown in Table 1, with S1 to S5 representing the five scales of Resnet50 network feature extraction. By comparing feature extraction methods with other models, it can be clearly found that our model has more obvious superiority in both Rank indicators and mAP indicators. When using the third scale of features jointly with the fourth and fifth scales for training, the mAP was only 61.9% and not as good as the traditional Reid using only the fifth scale of features (Li *et al.*, 2022), proving that the features extracted from the third scale of the network have a negative impact on the recognition results and the semantic information of the fourth scale of features has a positive impact on the recognition results.

　　Results of the joint training of loss and the joint optimized distance formula. We adopted a combined training approach using triplet loss, center loss and label smoothing loss. As ablation experiments, this paper also tested the experimental results of the model under the same hyper-parameter condition using only triplet loss, triplet loss + softmax cross-entropy loss, triplet loss + softmax cross-entropy loss + center loss and other loss respectively. Experiments proved that label smoothing loss is better than softmax loss, as the former allowed the model to focus slightly on the weights of the low probability distribution and some low probability positive samples were not ignored. The experiments also demonstrated that adding a center loss to reduce the distance between the same class helped to improve the recognition ability of the model slightly. We also proposed a joint optimization distance formula, which adapted the distance calculation formula in the triplet loss to the joint

optimization of the euclidean distance and the cosine distance and obtained the final total distance by combining the two distances in an self-adaptive weighted linear way. From Table 2 we can see the mAP is improved from 76.3% to 80.5%, proving the effectiveness of the self-adaptive weighted linear combination distance method.

Results of the iterative magnitude pruning. This paper compressed the dairy cow identification model based on the iterative magnitude pruning technique, setting the pruning rate to 20%, 40%, 60% and 80% respectively and pruning after each iteration, with each iteration containing 50 epochs. The experiment showed that the model after the 12th pruning was the best when the pruning rate is 20%. Similarly, the optimal

number of pruning were 4, 3 and 1 for the pruning rate of 40%, 60% and 80% respectively. The results of the model testing under different pruning rates are shown in Table 3. Among the 20% and 40% pruning models, the 20% pruning model has fewer remaining parameters and provides better lightweight compression of the network, reflecting the importance of recovery training after iterative magnitude pruning, which can effectively filter out the optimal sparse sub-network after a small number of pruning and recovery training.

Compared with other experiments. Because few studies deal with similar data as this paper, it is challenging to compare it with existing methods under the same conditions. Zhao *et al.* (2022) proposed a series of experiments to
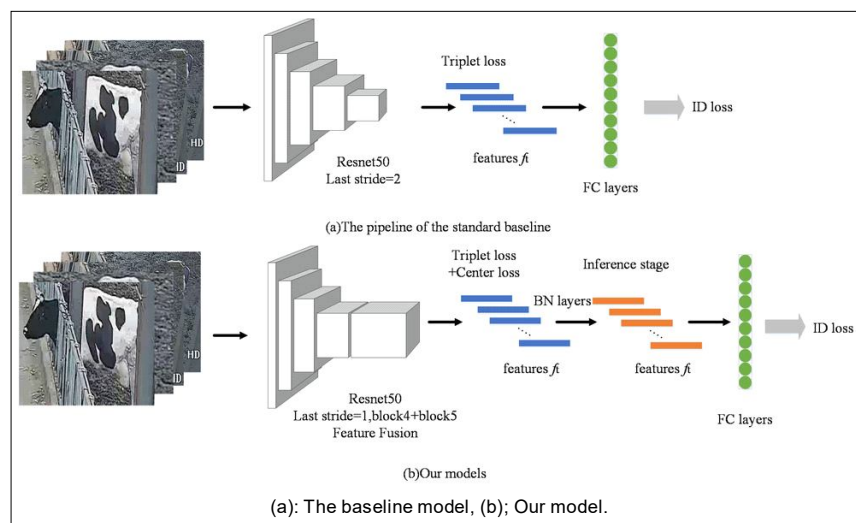

(a): The baseline model, (b); Our model.

Fig 1: Schematic diagram of the cow identification model.

Table 1: Experimental results of different feature extraction models.

Resnet 50 model used_layers					Evaluation indicators			
S1	S2	S3	S4	S5	Rank1	Rank5	Rank10	mAP
				√	90.9%	93.9%	945%	63.2%
			√	√	92.4%	94.6%	95.3%	68.1%
		√	√	√	90.8%	92.9%	93.5%	61.9%

Table 2: Joint optimization distance formula two formulas accounted for weighting test results.

Formula for calculating joint optimization distances	Rank1	Rank5	Rank10	mAP
Euclidean distance	95.5%	98.1%	98.4%	71.3%
0.9 euclidean distance + 0.1 cosine distance	95.5%	98.4%	98.8%	71.5%
0.8 euclidean distance + 0.2 cosine distance	95.3%	98.0%	98.5%	73.1%
0.7 euclidean distance + 0.3 cosine distance	94.1%	97.3%	98.0%	63.5%
0.6 euclidean distance + 0.4 cosine distance	94.7%	97.8%	98.6%	65.4%
0.5 euclidean distance + 0.5 cosine distance	96.5%	98.5%	98.8%	76.6%
0.4 euclidean distance + 0.6 cosine distance	95.6%	97.8%	98.2%	68.3%
0.3 euclidean distance + 0.7 cosine distance	96.3%	98.2%	98.8%	78.5%
0.2 euclidean distance + 0.8 cosine distance	97.3%	98.5%	98.7%	77.5%
0.1 euclidean distance + 0.9 cosine distance	97.4%	98.5%	98.7%	77.2%
Cosine distance	96.9%	98.3%	98.8%	70.9%
Idean distance + β * cosine di	96.5%	98.3%	98.9%	80.5%

Table 3: Model test results at different pruning rates.

Pruning rate%	Times of pruning	Number of parameters	Proportion of remaining parameters%	Rank1	Rank5	Rank10	mAP
0	0	$2.562*10^7$	100	96.6%	98.5%	98.7%	80.1%
20	12	$1.760*10^6$	6.87	97.5%	98.3%	98.7%	80.9%
40	4	$3.320*10^6$	12.96	97.5%	98.4%	98.9%	81.2%
60	3	$1.639*10^6$	6.40	97.3%	97.9%	98.6%	78.4%
80	1	$5.123*10^6$	20.00	97.2%	97.8%	98.5%	77.5%

Table 4: Results on different models or datasets.

Type of model and datasets	Average accuracy
Compact loss model on MVCAID100 and MVCAIDRE datasets.	92.53%
This paper's model, complete datasets	82.85%
This paper's model, small volume validation datasets	93.91%

evaluate the performance of compact loss on their MVCAID100 and MVCAIDRE datasets. Compared with other studies, the data scenario in Zhao's study was the most similar to ours, so it was analyzed as one of the comparative experiments. Because MVCAID100 and MVCAIDRE were private datasets, we could not obtain them and compare them intuitively on the same datasets. In order to compare the accuracy more intuitively, we made a small volume validate datasets closer to the MVCAIDRE datasets as a comparison. We randomly selected 400 no occlusion, daytime photography and the closest data for verification. The results are shown in Table 4. It can be seen that the accuracy of the complete datasets in this paper is relatively low, because our datasets has a large amount of occlusion data, including dark conditions and the shooting distance is relatively long. When this paper randomly selects 400 data similar to MVAIDRE datasets to evaluate the experimental results, the average accuracy reaches 93.91%, increased by 1.38%, which proves the effectiveness of this method.

From the above experiments, it can be seen that our model has stronger feature extraction ability compared to existing models. It is more reasonable to use the distance measurement method for ID matching and adds center loss and label smoothing loss based on data characteristics. It has better advantages in dealing with scenes with smaller recognition targets caused by long shooting distances, and the recognition accuracy achieves the best effect. Due to the addition of some adaptive weights and an increase in parameter size, the model is slightly more complex than the original method. However, the subsequent addition of pruning operations effectively improve the problem. Therefore, our method is effective. In the future, we will further research the identification of cows under extensive occlusion, such as adding tracking models (Liu *et al.,* 2017; Meng *et al*., 2019; Li *et al*., 2021).

CONCLUSION

We proposed a model for dairy cow identification, the final model obtained the highest mAP value of 81.2% and the average accurate value of 82.85% while saving computational resources. Experiments have shown that the identification method has a higher performance improvement than the baseline model in actual natural scenes and achieves higher recognition accuracy.

Conflict of interest

The authors declare no conflicts of Interest.

REFERENCES

Andrew, W., Greatwood, C., Burghardt, T. (2020). Fusing Animal Biometrics with Autonomous Robotics: Drone-based Search and Individual ID of Friesian Cattle (Extended Abstract). 2020 IEEE Winter Applications of Computer Vision Workshops (WACVW).

Cho, N.P., Thi, T.Z., Hiromitsu, H. (2018). A Hybrid Rolling Skew Histogram-Neural Network Approach to Dairy Cow Identification System. International Conference on Image and Vision Computing New Zealand (IVCNZ).

Gou, X.T., Huang, W., Liu, Q.F. (2019). A Cattle Face Detection Method Based on Improved NMS. Computer and Modernization. 40: 642-653.

Guo, Y., Chen, G.P., DING, J. (2020). Advances in cattle behavior monitoring techniques and classification methods. Acta Agriculture Jiangxi. 32(11): 99-105.

He, D.J., Liu, J.M., Xiong, H.T. (2020). Individual recognition method of dairy cows based on improved YOLO v3 model. Transactions of the Chinese Society of Agricultural Machinery. 51(4): 250-260.

Hermans, A., Beyer, L., Li, B.B. (2017). In defense of the triplet loss for person re-identification. arXiv preprint. https://doi.org/10.48550/arXiv.1703.07737.

Kumar, S., Pandey, A., Sai, R.S.K. (2018). Deep learning framework for recognition of cattle using muzzle point image pattern [J]. Measurement. 116: 1-17.

Li, F.J., Huang, Z.W.A. (2021). Target tracking method combining YOLOv3 detection and ReID. Journal of North China University of Science and Technology (Social Science Edition). 43(3): 110-118.

Li, Q., Shang, J., Li, B. (2021). Grassland Cattle Tracking System based on YOLOv3 and Deep SORT. Transducer and Microsystem Technologics. 40(6): 83-85.

Li, S.l., Yao, K., Cao, Z.J. (2020). Dairy industry technology development report 2019. Chinese Journal of Animal Science. 56(3): 136-144.

Li, F.J., Huang, Z.W. (2022). A Target Tracking Method Combining YOLOv3 Detection and ReI D. Journal of North China University of Science and Technology (Social Science Edition). 43(3): 110-118.

Liu, Z., He, S., Hu, W. (2017). Moving object detection based on background subtraction for video sequence. Journal of Computer Applications. 37(6): 1777-1781.

Liu, Y.F., Bian, H.D., He, Y.J. (2022). Detection method of multi-objective cows feeding behavior based on iterative magnitude pruning. Transactions of the Chinese Society for Agricultural Machinery. 53(2): 274-281.

Meng, L., Yang, X. (2019). A survey of object tracking algorithms. Acta Automatic Sinica. 45(7): 1244-1260.

Shen, W.Z., Hu, H.Q., Dai, B.S. (2020). Individual identification of dairy cows based on convolutional neural networks. Multimedia Tools and Applications. 79(21): 14711-14724.

Wang, C.Y., Bochkovskiy, A., Liao, H. (2022). Yolov7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. arXiv e-prints.

Wang, W.H. (2019). Research on cow identification method based on body feature image and depth learning. (Master's thesis).

Wu, C.X., He, X.X. (2022). Research on human posture estimation in earthquake rescue based on ResNet50. Cyber Security and Data Governance. 41(3): 50-58, 70.

Xu, B.B., Wang, W.S., Guo, L.F. (2020). A review and future prospects on cattle recognition based on non-contact identification. Journal of Agricultural Science and Technology. 7: 79-89.

Xu, X.Y., Liu, S.J., Qian, C. (2022). Study on the identification methods of typical cultured fish based on Res Net. Fishery Modernization. 3: 49. https://doi.org/10.3390/jmse10111598.

Yang, S.Q., Liu, Y.Q., Wang, Z. (2021). Improved YOLO V4 model for face recognition of diary cow by fusing coordinate information. Transactions of the Chinese Society of Agricultural Engineering. 37(15): 7. doi: 10.11975/j.issn.1002-6819.2021.15.016.

Yu, X. (2016). Research on precision feeding control system for dairy cattle. (Master's thesis).

Zhang, H.M., Wang, R., Dong, P.J. (2021). Beef cattle multi-target tracking based on DeepSORT algorithm. Transactions of the Chinese Society of Agricultural Machinery. 52(4): 248-256.

Zhao, J.M., Lian, Q.S. (2022). Compact loss for visual identification of cattle in the wild. Computers and Electronics in Agriculture 195(April): 106784.

Zhao, K.X., He, D.J. (2015). Recognition of individual dairy cattle based on convolutional neural networks. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE). 31(5): 181-187.

Zhao, W.J., Ding, L.Y., Li, Q.F. (2020). A method for identifying dairy cows' feeding behavior based on triaxial acceleration and artificial neural networks. Journal of Anhui Agricultural Sciences. 48(18): 231-234.

Zhuo, L., Yuan, S., Li, J.F. (2022). Pedestrian multi-attribute collaborative recognition method based on ResNet50 and channel attention mechanism. Measurement and Control Technology. 41(8): 1-8.