A HIGH-ACCURACY SHEEP FACE RECOGNITION MODEL BASED ON IMPROVED ResNet50
/
一种基于改进 ResNet50 的高精度羊脸识别模型

Xiwen ZHANG[1,2], Chuanzhong XUAN*[2], Tao ZHANG [2], Quan SUN[1,3]
[1]Jiangsu Maritime Institute, College of Marine Electrical and Intelligent Engineering, Nanjing, China
[2]Inner Mongolia Agricultural University, College of Mechanical and Electrical Engineering, Inner Mongolia, China
[3]State grid Inner Mongolia Eastern Electric Power Co., Ltd. Ewenki autonomous banner power supply branch, Hulunbuir, China
*Tel: 0471-4309215; *)Corresponding author E-mail: xcz@imau.edu.cn*
DOI: https://doi.org/10.35633/inmateh-74-03

Keywords: *Sheep face recognition; Deep learning; High-accuracy model; ResNet50; Attention mechanism*

ABSTRACT
Accurate identification of sheep is of significant importance for modern, intensive sheep farming. Traditionally, herders have used conventional methods to identify individual sheep, which are time-consuming, labour-intensive, and prone to considerable errors. In recent years, researchers have developed sheep face recognition models based on deep learning techniques to identify sheep using facial images. However, existing models suffer from insufficient theoretical research and limited recognition accuracy. To address these issues, this study develops a high-accuracy sheep face recognition model named ResNet-SFR. The core innovation of this model is the deepening of the feature extraction network of the original ResNet50, which enhances the model's ability to capture various facial features in sheep images, as well as improving its generalization and stability. Additionally, the Convolutional Block Attention Module (CBAM) attention mechanism is embedded into the original model to further enhance the identification of key features, significantly increasing the accuracy of sheep face recognition. Transfer learning is employed to pre-train the sheep face recognition model, further boosting the accuracy of ResNet-SFR. Experimental results show that on a self-constructed sheep face image dataset, ResNet-SFR achieves a recognition accuracy of 96.6%, demonstrating its superior performance in sheep face recognition tasks. The proposed ResNet-SFR not only offers high recognition accuracy but also exhibits strong applicability, meeting the practical needs of farm identification and showcasing promising application prospects.

摘要
羊只身份准确识别对现代化、集约化养羊业有着重要的应用意义。在过去，牧民们用传统的羊只身份识别方法对个体羊只进行身份识别，然而，传统方法既费时又费力，还存在较大的识别误差。近年来，学者们基于深度学习技术开发了羊脸识别模型，通过羊脸图像识别其对应的身份。然而，目前现有的羊脸识别模型存在理论研究不足和识别精度不足问题。针对上述问题，本研究开发了一组高精度羊脸识别模型，名为ResNet-SFR。该模型的核心创新在于将原ResNet50 的特征提取网络进行加深，此举不仅增强了模型捕捉羊脸图像中不同脸部特征的能力，同时也提高了其泛化性与稳定性。此外，在原模型的基础上嵌入了CBAM注意力机制，进一步加强了模型对关键特征的识别，显著提高了羊脸识别的准确度。本研究采用了迁移学习对羊脸识别模型进行预训练，进一步提升了ResNet-SFR的识别精度。试验结果表明，在自制的羊脸图像数据集上，ResNet-SFR的识别精度达到了 96.6%，证明了其在应对羊脸识别任务上的优越表现。本研究提出的ResNet-SFR在羊脸识别方面不仅识别精度高，且具有较强的应用性，符合养殖场识别的实际需求，展现了较好的应用前景。

INTRODUCTION
With the continuous development of intelligent animal husbandry, intelligent and intensive breeding methods have gained widespread attention. In modern sheep farms, it is necessary to collect various information about the sheep, such as birth dates, weight, vaccination records, and pregnancy status. With this information, herders can formulate scientific management strategies and improve feeding practices, thereby effectively managing the pasture and further reducing breeding costs (*Billah et al., 2023*). Accurate identification of sheep is a prerequisite for collecting individual information (*Xue et al., 2024*).

In the breeding process, the accuracy of sheep identification directly affects the effectiveness of individualized management. For instance, by identifying each sheep, it is possible to record and track its health status and production performance, ensuring that each sheep receives appropriate care and management

(*Zhang et al., 2023*). Additionally, identifying sheep helps to promptly detect and isolate sick individuals, preventing the spread of diseases within the flock and thus reducing breeding risks and losses. Accurate identification also provides reliable data support for breeding programs, assisting farmers in selecting superior breeds and enhancing breeding efficiency. Ultimately, accurate sheep identification enables traceability of meat quality, meeting consumer demand for high-quality meat. Therefore, sheep identification has become an indispensable part of modern sheep farming (*Salama et al., 2019*).

Traditional methods of sheep identification include paint markings, manual observation, and ear tags. However, these traditional methods have significant limitations. Manual observation is inefficient and inaccurate, making it unsuitable for large flocks (*Hitelman et al., 2022*). Additionally, paint markings and ear tags require regular maintenance and cleaning by staff, further increasing breeding costs. Currently, some farms use RFID ear tags and readers for sheep identification. However, in complex environments, RFID ear tags often become damaged or lost, and the identification process is easily disrupted. In summary, relying solely on traditional sheep identification methods may be inconvenient for managing modern sheep farms (*Sharma et al., 2020*).

With the development of information technology, biometric image recognition has gained increasing attention and has become a trend in the field of animal face recognition (*Zhang et al., 2022; Deng et al., 2022*). Biometric image recognition utilizes intelligent monitoring equipment and computer vision to capture stable biometric features of individuals, such as facial images, and then performs identification based on these individual features. In recent years, researchers have developed sheep face recognition models using deep learning techniques, training on sheep facial images to achieve identification. However, existing sheep face recognition models suffer from insufficient theoretical research and limited recognition accuracy. To address these limitations, this study developed a high-accuracy sheep face recognition model named ResNet-SFR. By integrating image recognition models and various improvement strategies, the aim is to overcome the limitations of traditional sheep identification methods. The feature extraction layers of the original ResNet50 network were deepened. Additionally, the CBAM attention mechanism was introduced to enhance feature extraction capabilities. Finally, transfer learning was employed to pre-train ResNet-SFR, further improving the model's recognition accuracy. Experimental results show that ResNet-SFR achieved the highest recognition accuracy, meeting the practical needs of modern sheep farms and providing technical support for sheep face recognition technology.

MATERIALS AND METHODS

Dataset

This study used a group of Small Tail Han sheep as test subjects. The Small Tail Han sheep is a widely bred sheep breed in China, known for its high reproductive rate and superior meat quality. The coat colour of the Small Tail Han sheep is predominantly white, with some individuals having black or brown spots on their faces. These spots are mostly concentrated around the eyes, ears, cheeks, or mouth. The sheep face images used in this study were collected in August 2020 at Tianjin Aoqin Animal Husbandry Co., Ltd. The collection details are as follows: A total of 50 test sheep were used, aged between 1 to 3 years. The test sheep were numbered from 1 to 50, with the numbers corresponding to their identity information. Before collecting the facial images, the test sheep were enclosed in a pen and allowed to move freely. A photographer used a Canon EOS 600D DSLR camera (Canon, Tokyo, Japan) to capture the sheep face images, which were saved in JPG format with a resolution of 2736×1824. The collection distance was over one meter. The collection times were from 9:00 AM to 11:00 AM and from 2:00 PM to 5:00 PM. To increase the complexity and practicality of the dataset, facial images of the test sheep were taken from different angles, including left profile, right profile, and frontal views. A schematic diagram of the sheep face image collection is shown in Fig.1.

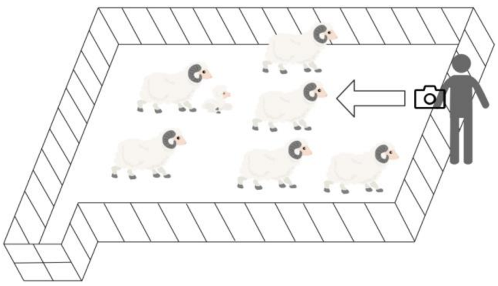

Fig. 1 - Schematic diagram of the sheep face image collection

Using the above methods, 50 facial images were collected for each test sheep. Sample images of the test sheep faces are shown in Fig.2.

Fig. 2 - Sample images of the test sheep faces

To enhance the applicability of the recognition model in practical scenarios, data augmentation was performed on the collected sheep face images. The specific data augmentation operations applied were: adjusting image brightness (-45% to 45%), adjusting image contrast (-45% to 45%), rotating the image 45 degrees to the left and right, and vertical flipping. Using these methods, 100 augmented images were generated for each test sheep to supplement the training of the recognition model. Ultimately, 150 facial images were retained for each test sheep, creating the sheep face image dataset. All images were randomly divided into training, validation, and test sets in a ratio of 8:1:1. The specific configuration of the sheep face image dataset is shown in Table 1.

Table 1

The sheep face image dataset

Dataset	Images	Size	Proportion
Training	6000	2736×1824	80%
Verification	750	2736×1824	10%
Testing	750	2736×1824	10%
Total	7500	2736×1824	100%

ResNet50

Convolutional Neural Networks (CNNs) achieve image recognition tasks by continuously extracting local features from images (*Deng et al., 2021*). Ideally, a deeper feature extraction network can capture richer and more complex target features. However, with increased network depth, CNNs are prone to issues such as gradient vanishing and network degradation, which can further lead to reduced learning efficiency or ineffective improvement in accuracy. To address these issues, residual networks were proposed (*He et al., 2016*). Residual networks successfully overcome these problems by introducing the concept of shortcut connections. The basic form of a residual network involves directly connecting the output of a layer to the input of a subsequent layer, which does not increase the model's parameter count or computational complexity. A schematic diagram of the residual network structure is shown in Fig.3. The input features x are processed through feature extraction to obtain output features $F(x)$. The residual network further adds the input features to the output features through a residual connection to obtain the final output features y. The specific calculation formula is as follows:

$$y = F(x) + x \tag{1}$$

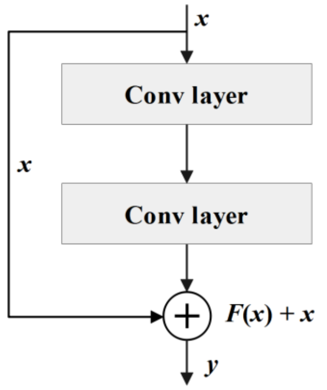

Fig. 3 - Schematic diagram of the residual network structure

ResNet50 is a residual network composed of forty-nine convolutional layers and one fully connected layer. It mainly consists of four basic modules: Block, Conv-Block, Identity-Block, and Block2 (*Zhang et al., 2023)*. The Block module is composed of a convolutional layer and a max-pooling layer. The convolutional layer is responsible for the initial feature extraction from the input image and forms the initial feature representation. Next, the model uses the max-pooling layer to downsample the extracted features, reducing the spatial dimensions of the feature map. At the same time, the max-pooling layer helps to retain the main feature information, reduce computational complexity, and improve the model's adaptability to translation invariance.

The Conv-Block is used to change the dimensions of the input vector. The Conv-Block can be divided into two parts: the left side is the main branch, which includes two sets of convolutional layers with batch normalization and ReLU activation functions, as well as an additional set of convolutional layers that only have batch normalization. The right side consists of the residual branch, which includes a set of convolutional layers with batch normalization. The Conv-Block captures abstract features at different levels and achieves nonlinear mapping through convolutional operations. Therefore, the Conv-Block plays a role in extracting high-order features throughout the ResNet architecture.

The Identity-Block is also composed of two parts. The left side is the same as the main branch of the Conv-Block, while the right side is the residual branch, specifically connecting the input features with the output features of the main branch. In ResNet, the Identity-Block is used to handle cases where the input and output have the same dimensions. The Identity-Block effectively prevents the problem of gradient vanishing, retains the original input information of the module, and enables iterative learning of features layer by layer in subsequent layers.

In Block2, a global average pooling layer and a fully connected layer are used. The global average pooling layer's function is to obtain a feature vector with the average value of each channel. Then, the fully connected layer maps the final feature vector to specific classes, outputting the recognized target's class through the Softmax activation function, thus completing the entire image recognition task.

ResNet50, with its unique network structure, demonstrates superior recognition performance in numerous recognition tasks. The overall structure of ResNet50 is shown in Fig.4.

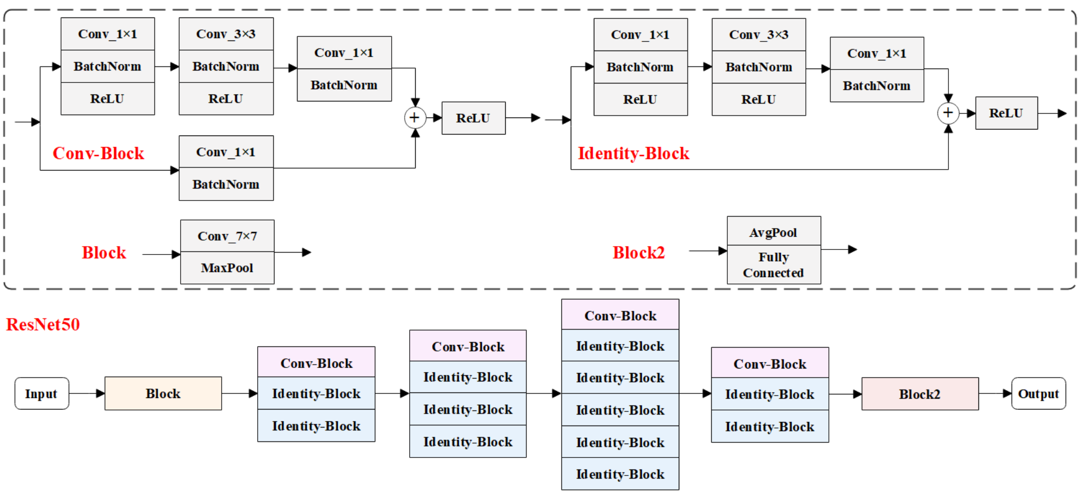

Fig. 4 - Overall structure of ResNet50

CBAM

The CBAM is an attention mechanism for CNNs that enhances feature representation by integrating both spatial and channel attention (*Ni et al., 2024*). The modular design of CBAM allows it to significantly improve model performance with relatively low computational cost (*Zhang et al., 2023*).

The CBAM module consists of two sub-modules: the channel attention module and the spatial attention module. The channel attention module first captures global information through global average pooling and global max pooling, and then generates channel attention weights using a shared multilayer perceptron (MLP). The input feature map is then multiplied by these channel attention weights to obtain the weighted feature map. The spatial attention module aggregates the input feature map along the channel dimension by performing global average pooling and max pooling, generating a two-dimensional feature map. This feature map then passes through a convolutional layer to generate spatial attention weights, which are multiplied by the channel-weighted feature map to obtain the final output (*Fang et al., 2024*).

The structure of CBAM is highly pluggable, making it easy to integrate into existing convolutional neural networks. Its workflow can be summarized as follows: first, channel dimension weighting is performed through the channel attention module, and then spatial dimension weighting is performed through the spatial attention module. This double weighting process enhances the representation ability of the input features. CBAM effectively increases the model's focus on important features, improving performance in the sheep face recognition task. The overall structure of the CBAM attention mechanism is shown in Fig.5.

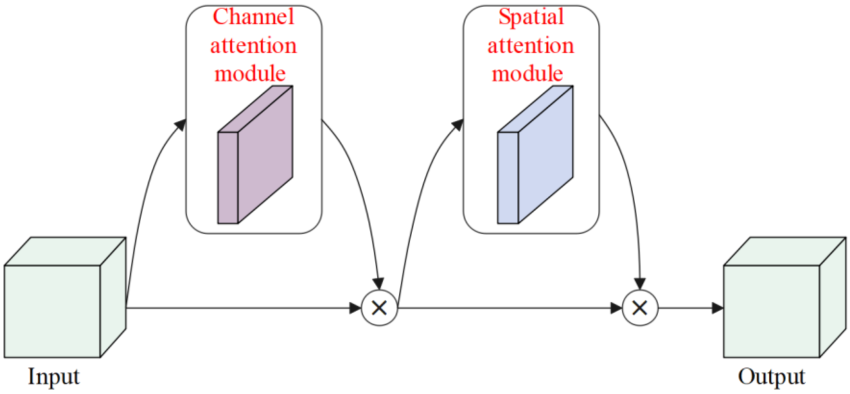

Fig. 5 - Overall structure of the CBAM attention mechanism

In the process of optimizing ResNet50, this study introduces the CBAM attention mechanism before the model's Block2 to further enhance performance. CBAM can adaptively adjust the responses at different positions in the feature map, making the model more focused on important regions in the image. Before Block2, ResNet50 completes the feature extraction of the final convolutional block and extracts the complete features of the target. Therefore, introducing the CBAM module before Block2 can effectively strengthen the sheep face-related features within the complete feature set and improve their representation. This design helps to enhance the model's focus on key features of the sheep face.

Deepening network

ResNet50 includes four sets of stacked convolutional block structures, each composed of Conv-Blocks and Identity-Blocks. Considering the high difficulty of the sheep face recognition task in this study, the network depth of ResNet50 was further increased to improve the model's feature extraction capability. The specific deepening method is as follows: an additional set of stacked convolutional blocks was added to the third and fourth sets of stacked convolutional blocks in ResNet50. The configuration of the added convolutional blocks is consistent with the second set of stacked convolutional blocks, including one Conv-Block and three Identity-Blocks. This design introduces deeper convolutional layers to extract more detailed sheep face features while maintaining the stability of the network structure. This approach aims to make the improved ResNet50 better suited for the sheep face recognition task, thereby enhancing the model's overall performance and generalization ability.

Transfer learning

Transfer learning is a model training method. Transfer learning is defined as using a model trained on task A as the initial model and retraining it for task B (*Bo et al., 2024*).

By leveraging the knowledge from the source domain data, transfer learning makes the model more robust and generalizable in the target domain. Additionally, transferring the knowledge of a complex model to a simplified model can achieve model lightweighting and accelerated inference (*Liu et al., 2024*). The application of transfer learning not only improves model performance but also provides feasibility for the practical application of deep learning tasks in resource-constrained or data-limited situations.

In this study, the following method is used to apply transfer learning to multiple recognition models to achieve pre-training. First, the sheep face image dataset is divided into Dataset A and Dataset B. The training, validation, and test sets are all evenly split. In Method 1, the recognition model is first trained on Dataset A, and a pre-trained model is obtained once the training results are stable. This pre-trained model represents a model that has already learned the sheep face recognition task and achieved a "pre-learning" effect. The resulting pre-trained model is then further trained on Dataset B, and the final training results for Dataset B are obtained, which are considered as the final training results for Method 1. In Method 2, the recognition model is first trained on Dataset B to obtain a pre-trained model. Then, this pre-trained model is trained on Dataset A, resulting in the final training results for Method 2. The two sets of results are averaged to provide the final training results of the recognition model after transfer learning.

ResNet-SFR

To further enhance the recognition performance of the sheep face recognition model, this study employs the three improvement strategies mentioned above based on ResNet50, ultimately constructing the sheep face recognition model ResNet-SFR. The overall structure of ResNet-SFR is shown in Fig.6.

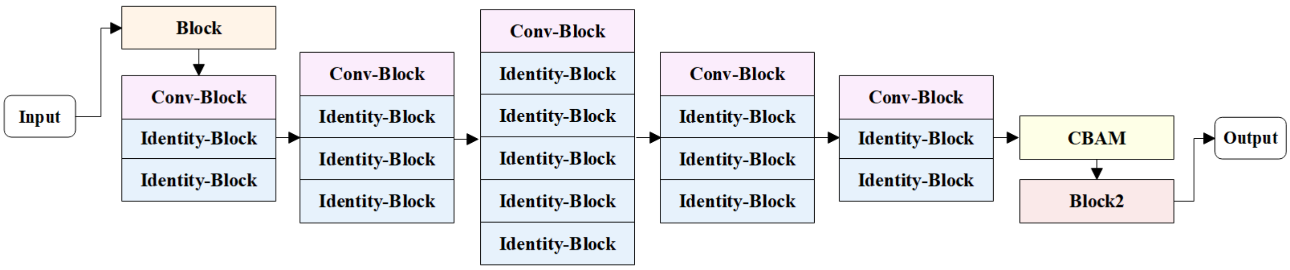

Fig. 6 - Overall structure of ResNet-SFR

Evaluation Metrics

The training platform used in this study is configured as follows: CPU is Intel Core i7-9700, GPU is NVidia GeForce RTX 2080Ti with 11 GB VRAM, memory is 16 GB, and the operating system is Windows 10. The software platform is PyCharm, with CUDA 11.3, PyTorch version 1.10.0, and Python version 3.8. During model training, the dynamic learning rate is set to 0.001, with 50 epochs and a batch size of 16.

In this study, the evaluation metrics for the sheep face recognition model include precision, recall, F1-score, and accuracy. Precision represents the percentage of correctly classified samples out of the total number of samples. Recall measures the ratio of correctly retrieved samples to the number of samples that should have been retrieved. Here, TP, FP, TN, and FN represent the counts of true positives, false positives, true negatives, and false negatives, respectively. The F1-score takes both precision and recall into account. Accuracy is one of the fundamental metrics for evaluating the performance of deep learning models, as it measures the proportion of correctly classified samples out of all samples. The ranges for F1-score, recall, precision, and accuracy are from 0% to 100%.

The formulas for these metrics are as follows:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1-score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

$$\text{Accuary} = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

RESULTS AND DISCUSSIONS

Training curve

To further investigate the training results of ResNet-SFR, an analysis was conducted on the training curves of ResNet50 and ResNet-SFR, specifically including accuracy curves and loss curves. The training curves for both models are shown in Fig.7. From the training curves, it can be observed that after the number of epochs exceeds 10, the recognition accuracy of ResNet-SFR is consistently higher than that of ResNet50. Additionally, the loss value for ResNet-SFR is consistently lower than that of ResNet50. Ultimately, the training accuracy of ResNet-SFR stabilizes at 96.6%. These results indicate that, compared to ResNet50, ResNet-SFR is more effective in learning the sheep face recognition task.

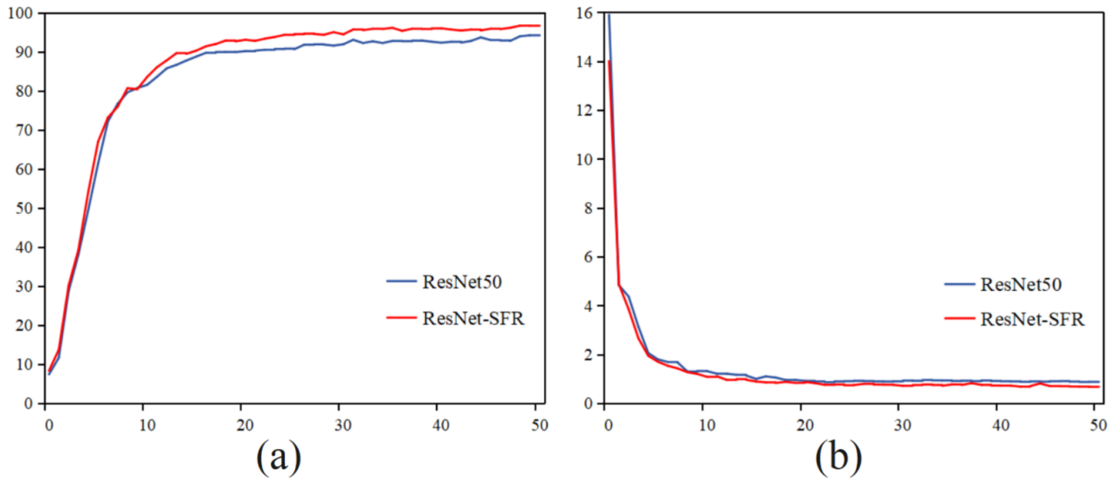

Fig. 7 - The training curves for both models
a. Accuracy curve; b. Loss curve

Comparison of Recognition Performance of Different Models

In this section, several classic deep learning models were selected for pre-training and their performance on the sheep face recognition task was evaluated. The chosen models include those previously used in sheep face recognition research, specifically: VGG16, ResNet18, MobileViT, AlexNet, RepVGG, YOLOv4, and ResNet50.

The pre-training results for each sheep face recognition model are shown in Table 2. From the results in Table 2, it can be seen that among the various recognition models, ResNet50 achieved the best performance on the sheep face image dataset, with an F1-score and accuracy of 93.9% and 94.3%, respectively. Based on these results, this study uses ResNet50 as the benchmark model for high-precision sheep face recognition and will proceed with targeted improvements to further enhance recognition accuracy.

Table 2

The training results of the sheep face recognition model

Model	F1-score (%)	Accuracy (%)
VGG16	87.5	88.0
ResNet18	88.2	87.9
MobileViT	89.8	89.9
AlexNet	91.0	90.5
RepVGG	92.3	92.5
YOLOv4	93.0	93.6
ResNet50	93.9	94.3

Ablation Experiment

In this section, ablation experiments were conducted on ResNet-SFR to investigate the specific performance of several improvement strategies. The results of the ablation experiments on ResNet-SFR are shown in Table 3. Based on the three improvement strategies, eight combinations of improved models were tested. Here, TL represents models pre-trained using transfer learning.

From the results in Table 3, it can be observed that compared to ResNet50, deepening the network layers alone improved the F1-score and accuracy of the modified model by 0.8% and 0.7%, respectively.

Introducing the CBAM module alone improved the F1-score and accuracy by 1.0% and 0.4%, respectively. Using TL pre-training alone improved the F1-score and accuracy by 0.6% and 0.5%, respectively. These results indicate that each of the three improvement strategies positively enhanced the recognition performance of ResNet50.

When using two improvement strategies simultaneously, compared to the results from ResNet50, deepening the network layers and introducing the CBAM module improved the F1-score and accuracy of the modified model by 1.8% and 1.6%, respectively. Deepening the network layers and using TL pre-training improved the F1-score and accuracy by 1.5% and 0.9%, respectively. Introducing the CBAM module and using TL pre-training improved the F1-score and accuracy by 1.7% and 1.1%, respectively. Finally, when all three improvement strategies were applied together, ResNet-SFR achieved an F1-score and accuracy of 96.3% and 96.6%, respectively. Compared to ResNet50, the F1-score increased by 2.4% and the accuracy increased by 2.3%. The experimental results indicate that the proposed improvement strategies effectively enhance the recognition performance of ResNet50.

Table 3

The ablation results of ResNet-SFR

ResNet50	Deepening network	CBAM	TL	F1-score (%)	Accuracy (%)
√				93.9	94.3
√	√			94.7	95.0
√		√		94.9	94.7
√			√	94.5	94.8
√	√	√		95.7	95.9
√	√		√	95.4	95.2
√		√	√	95.6	95.4
√	√	√	√	96.3	96.6

Comparison of Attention Mechanisms

In this section, different attention modules were introduced into ResNet50, and the training results were compared. The integration position of each attention module was consistent. The test results are shown in Table 4. In Table 4, the improved models with embedded attention modules all achieved performance enhancements. Specifically, when the CBAM module was added to ResNet50+Deepening network+TL, the improved model achieved higher recognition accuracy. In summary, the CBAM module was chosen as the attention mechanism improvement strategy for ResNet-SFR.

Table 4

The training results of ResNet-SFR introducing different attention modules

ResNet50+Deepening network+TL	F1-score (%)	Accuracy (%)
/	95.4	95.2
+SE	95.6	95.5
+ECA	95.8	95.9
+CA	96.1	96.3
+CBAM(Ours)	96.3	96.6

CONCLUSIONS

This study presents a high-precision sheep face recognition model named ResNet-SFR, which builds upon the ResNet50 architecture and incorporates several advanced techniques to enhance performance. By deepening the network layers, integrating the CBAM attention mechanism, and employing transfer learning, ResNet-SFR significantly improves the accuracy and robustness of sheep face recognition.

The experimental results demonstrate that ResNet-SFR outperforms traditional methods and other deep learning models, achieving a recognition accuracy of 96.6% on the sheep face image dataset.

Compared to ResNet50, the proposed model shows a 2.4% improvement in F1-score and a 2.3% improvement in accuracy. These enhancements validate the effectiveness of the introduced modifications in capturing and utilizing complex features of sheep faces.

In addition, the comparative experiments with various attention mechanisms highlighted the superiority of the CBAM module, which further contributed to the model's performance boost. The comprehensive evaluation and ablation studies confirm that the combination of deepened network layers, CBAM, and transfer learning is instrumental in achieving high precision in sheep face recognition tasks.

Overall, this research not only provides a robust and accurate solution for sheep face recognition but also contributes valuable insights into improving recognition models through advanced techniques. The proposed ResNet-SFR model holds significant promise for practical applications in modern sheep farming and animal management.

ACKNOWLEDGEMENTS

We acknowledge that this work was supported by the Fundamental Research Funds of Inner Mongolia Agricultural University (BR221314 and BR221032), the Science and Technology Planning Project of Inner Mongolia Autonomous Region (2021GG0111), and the Research and Innovation Project for Doctoral Candidates in Inner Mongolia Autonomous Region (B20231075Z).

REFERENCES

[1] Billah M., Wang X., Jiang Y., (2022), Real-time goat face recognition using convolutional neural network. *Computers and Electronics in Agriculture*, Vol 194, pp. 106730, China.

[2] Bo C., Liu S., Liu Y., Guo Z., Wang J., Xu J., (2024). Research on Isomorphic Task Transfer Algorithm Based on Knowledge Distillation in Multi-Agent Collaborative Systems, *Sensors*, Vol. 24, Issue 14, pp. 4741, China.

[3] Deng X., Zhang S., Shao Y., Yan X., (2022). A real-time sheep counting detection system based on machine learning. *INMATEH-Agricultural Engineering*, Vol 67, Issue 2, pp. 85-94, Romania.

[4] Deng X., Yan X., Hou Y., et al. (2021). Detection of behaviour and posture of sheep based on YOLOv3. *INMATEH-Agricultural Engineering*, Vol. 64, Issue 2, pp. 457-466, Romania.

[5] Fang S., Chen C., Li Z., Zhou M., Wei R., (2024). YOLO-ADual: A Lightweight Traffic Sign Detection Model for a Mobile Driving System, *World Electric Vehicle Journal*, Vol. 15, Issue 7, pp. 323, China.

[6] He K., Zhang X., Ren S., Sun J., (2016). Deep residual learning for image recognition. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770-778, China.

[7] Hitelman A., Edan Y., Godo A., Berenstein R., Lepar J., Halachmi I. (2022). Biometric identification of sheep via a machine-vision system. *Computers and Electronics in Agriculture*, Vol.194, pp. 106713, Israel.

[8] Liu J., Liu H., He Y., Tong S., (2024). An Improved Detail-Enhancement CycleGAN Using AdaLIN for Facial Style Transfer. *Applied Sciences*, Vol. 14, Issue 14, pp. 6311, China.

[9] Ni X., Wang F., Huang H., Wang L., Wen C., Chen D., (2024). A CNN- and Self-Attention-Based Maize Growth Stage Recognition Method and Platform from UAV Orthophoto Images, *Remote Sensing*, Vol.16, Issue 14, pp. 2672, China.

[10] Salama A., Hassanien AE., Fahmy AA., (2019). Sheep identification using a hybrid deep learning and Bayesian optimization approach. *IEEE Access*, Vol.7, pp. 31681-31687, Egypt.

[11] Sharma A., Jain A., Gupta P., (2020). Machine learning applications for precision agriculture: A comprehensive review. *IEEE Access*, Vol.9, pp. 4843-4873, India.

[12] Xue J., Hou Z., Xuan C., Ma Y., Sun Q., Zhang X., Zhong L., (2024), A Sheep Identification Method Based on Three-Dimensional Sheep Face Reconstruction and Feature Point Matching. *Animals*, Vol.14, Issue 13, pp. 1923, China.

[13] Zhang X., Hou Z., Xuan C., (2022). Design and experiment of recognition system for coated red clover seeds based on machine vision. *INMATEH-Agricultural Engineering*, Vol. 66 (1), pp. 62-72, Romania.

[14] Zhang X., Xuan C., Ma Y., Su H., (2023). A high-precision facial recognition method for small-tailed Han sheep based on an optimised Vision Transformer, *Animal*, Vol.17, pp. 100886, China.

[15] Zhang X., Xuan C., Xue J., Chen B., Ma Y., (2023), LSR-YOLO: A High-Precision, Lightweight Model for Sheep Face Recognition on the Mobile End. *Animals*, Vol 13, Issue 11, pp. 1824, China.

[16] Zhang Y., Wang H., Liu J., (2023). A lightweight winter wheat planting area extraction model based on improved DeepLabv3+ and CBAM. *Remote Sensing*, Vol. 15, Issue 17, pp. 4156, China.