Research Article

Facial Recognition of Cattle Based on SK-ResNet

He Gong,[1,2,3] Haohong Pan,[1] Lin Chen,[1] TianLi Hu [ID],[1,2,3] Shijun Li,[4] Yu Sun,[1] Ye Mu,[1,2] and Ying Guo[1]

[1]*College of Information Technology, Jilin Agricultural University, Changchun 130118, China*
[2]*Jilin Province Agricultural Internet of Things Technology Collaborative Innovation Center, Changchun 130118, China*
[3]*Jilin Province Intelligent Environmental Engineering Research Center, Changchun 130118, China*
[4]*College of Electronic and Information Engineering, Wuzhou University, Wuzhou 543003, China*

Correspondence should be addressed to TianLi Hu; hutianli@jlau.edu.cn

With the gradual increase of the scale of the breeding industry in recent years, the intelligence level of livestock breeding is also improving. Intelligent breeding is of great significance to the identification of livestock individuals. In this paper, the cattle face images are obtained from different angles to generate the cow face dataset, and a cow face recognition model based on SK_ResNet is proposed. Based on ResNet-50 and using a different number of sk_Bottleneck, this model integrates multiple receptive fields of information to extract facial features at multiple scales. The shortcut connection part connects to the maximum pooling layer to reduce information loss; the ELU activation function is used to reduce the vanishing gradient, prevent overfitting, accelerate the convergence speed, and improve the generalization ability of the model. The constructed bovine face dataset was used to train the SK-ResNet-based bovine face recognition model, and the accuracy rate was 98.42%. The method was tested on the public dataset and the self-built dataset. The accuracy rate of the model was 98.57% on the self-built pig face dataset and the public sheep face dataset. The accuracy rate was 97.02%. The experimental results verify the superiority of this method in practical application, which is helpful for the practical application of animal facial recognition technology in livestock.

1. Introduction

With the development of animal husbandry in the direction of scale, informatization, and refinement, intelligent cattle farms will gradually replace the traditional farming mode of small scale, such as retail farming. In large cattle farms, in order to realize the automatic and information-based daily fine management of individual cattle and to realize the health status tracking of each cow and the traceability of dairy and meat products, it is necessary to realize the construction and improvement of the quality traceability system, and the key is the identification of individual cattle [1]. The traditional methods of individual identification of cattle include ear engraving, the external marking method with an ear tag, and the external equipment marking method with RFID [2]. Ear-cutting is a painful and time-consuming incision in the animal's ear [3]; the external marking of ear tags is often lost or damaged during breeding. Ear tags are easily lost, so they cannot be worn for long periods of time [4]. If the RFID tagging method is used for a long time, it will cause security problems such as ear tag falling off, tag content tampering, system crash, and server intrusion attack, and the cost is high [5, 6].

In recent years, driven by deep learning, the use of machine vision technology to supervise the identification of dairy cows has become a trend. As a popular technology for intelligent and precise breeding, machine vision has the advantages of low cost, noncontact and avoiding animal stress, long continuous monitoring time, and so on. Noncontact identification is a new trend in livestock and poultry identification, which is based on biological characteristics and is unique, invariable, low cost, easy operation, and high animal welfare. It is a new trend in livestock and poultry identification. Noncontact identification methods use computer vision to extract biometrics for individual identification. In biometrics, facial recognition has strong anti-

interference and scalability. In 2015, Moreira et al. [7] applied deep convolutional neural networks to recognize lost dogs in dog facial recognition, but this study has low recognition accuracy in dogs of the same breed and high similarity. In 2018, Hansen et al. [8] proposed a pig face recognition algorithm based on convolutional neural networks. By collecting information from pigs' facial features, pigs with black spots have a better recognition effect, but in this study, artificial paint to create artificial features is difficult to achieve in practice. In 2019, Yao et al. [9] proposed a cattle face recognition framework that combines Faster R–CNN detection and the PANSnet-5 recognition model. First, an image was input through the cattle face detection model, and the cattle face region in the image was detected and cropped. Then, the cropped image of the cattle face region was sent to the recognition model to confirm its specific number. However, the characteristics of the facial pattern of dairy cows are obvious, so there is less research value. In 2020, Mathieu Marsot et al. [10] proposed a new framework consisting of computer vision algorithms and machine learning and deep learning techniques. First, two cascaded classifiers based on Haar features and a shallow convolutional neural network automatically detect high-quality images of a pig's face and eyes. Second, deep convolutional neural networks are used for facial recognition. However, due to the black spots on pig faces, in order to train the network, the output images of the Haar cascade eye detector need to be manually classified and then input into the neural network. This study is very heavy and often difficult to achieve in practical applications. In 2021, Bello et al. [11] proposed a deep belief network to learn the texture features of bullnose images and use bullnose image patterns for recognition. There will be certain limitations when using this method to extract bullnose texture from free-range cattle farms.

Discrete orthogonal polynomials have attracted a lot of attention from researchers in many scientific fields, especially in speech and image analysis, due to their robustness to noise. The basic principle is to use orthogonal polynomials (OPs) to form matrices and to use the basis functions of the orthogonal polynomials as approximate solutions of differential equations. In recent years, orthogonal polynomials have been widely used in face recognition, edge detection, and other related fields. In 2020, Abdul-Hadi et al. [12] proposed a new recursive algorithm to generate Meixner polynomials (CHPs) for higher-order polynomials, which is 44 times faster than existing recursive algorithms but still has room for improvement in speed. In 2021, Abdulhussain et al. [13] proposed a new recursive algorithm for solving higher-order Charlier polynomials(MNPs) coefficients. Feature extraction tools are computationally expensive but not used for boundary detection. In 2022, Mahmmod et al. [14] proposed an operation method for calculating Hahn orthonormal basis and applied it to the calculation of high-order orthonormal basis. This method uses two adaptive threshold recursion algorithms to stabilize the generation of Discrete Hahn polynomials (DHP) coefficients. The algorithm has better performance in the case of wider parameter value ranges α and β and polynomial size.

Although the traditional recognition method has achieved good results, the recognition process is complicated and often requires manual intervention. The image features extracted by artificial design are usually shallow features of the image, with limited expressive ability and insufficient effective feature information. In addition, the artificial design method has poor robustness and is greatly affected by external conditions. With the development of deep learning technology and discrete orthogonal polynomials and the improvement of the hardware environment, the method of action image recognition based on deep learning has become a research hotspot. The cattle face dataset is generated by using cattle face images obtained from different angles, and an improved recognition model based on SK-ResNet is proposed to extract cattle face features. The model uses ResNet-50 as the basic model, uses different numbers of SK-Bottlenecks, and fuses the information of multiple receptive fields to extract facial features at multiple scales; the maximum pooling layer is connected in the model shortcut connection to reduce information loss. The ELU activation function is used in the network, and its linear part on the right makes the ELU more robust to input changes or noise. The average value of the left curve and ELU output is close to 0, which makes the model converge faster and can solve the problem of neuron death. The recognition model based on SK-ResNet was trained with the bovine face dataset, and it was proved that the cattle could be accurately identified. The method is tested on public datasets and self-built datasets. Compared with the existing recognition methods, the experimental results verify the advanced nature of the method.

The contributions of this paper include the following:

(1) We built up two large datasets for cattle identification and model robustness testing. The first one consists of cattle facial images. The second one consists of long white pig facial images. Those images were captured with different angles and backgrounds.

(2) We propose a cattle face recognition model based on SK-ResNet. Based on ResNet-50, this model uses different amounts of SK-Bottleneck to fuse multiple perceptual fields of information and extract cattle face features at multiple scales.

(3) We tested the model in this paper on the self-built pig face and public sheep face datasets and compared it with other models. The experimental results show that the proposed model is robust and can be well applied to livestock face recognition.

2. Materials and Methods

2.1. Data Collection and Processing. The data were collected at Dongfeng Dairy Cattle Farm, Liaoyuan City, Jilin Province, China, and the shooting time was July 2021. Through the camera equipment deployed on the farm, the images of eight solid-color cattle were intercepted from different angles. The solid-color cattle samples are shown in Figure 1, and the cattle face dataset with a resolution of 1298×1196 pixels and a format of Joint Photographic Experts Group

FIGURE 1: Part of solid-color cattle face data.

(JPG) was obtained. We aimed to prevent image saturation, avoid direct sunlight on the face in the images, and remove complex backgrounds [15] by extracting the face region of the image A total of 5677 facial images of cattle were collected, which were randomly divided into training and verification sets at a ratio of 7 : 3 (3974 training images and 1703 verification images). The purpose is that when the feature space dimension of the sample is larger than the number of training samples, the model is prone to overfitting. In order to enhance the robustness and generalization ability of the network, the number of training samples is increased by enlarging the limited number of training samples. The expansion method of translation, rotation, and cropping is used to increase the sample size to four times the original, avoiding the problem of overfitting. After data enhancement, 15,896 training sets were obtained. The size of the training dataset has a significant impact on the performance of the training network.

2.2. Cattle Individual Identification Process.

The identification process used in this paper is shown in Figure 2. The acquired images are preprocessed with the method in Section "Data Collection and Processing." Then, the dataset is divided into training set and a verification set with a ratio of 7 : 3. In this paper, ResNet is used as the skeleton network to construct the recognition model of the individual cattle's faces. The training set is used to train the model, and the validation set is used to verify the accuracy and robustness of the model, so as to realize the rapid and effective identification of cattle.

2.3. Methods

2.3.1. Convolutional Neural Network.

In the ResNet network [16], with the deepening of the network, problems such as gradient disappearance and gradient explosion will occur, which makes the training of convolutional neural networks difficult and the model performance will also decline. In

order to alleviate this effect, a residual block can be constructed to hop connections between different network layers in order to improve network performance. Therefore, the residual network has been widely used in plant disease spot classification [17], pathological image classification [18], remote sensing image classification [19], and face recognition [20] due to its superior performance. The residual module structure is shown in Figure 3.

For the multilayer stacked network structure, when the input data is X, the learning feature is denoted as $H(X)$. It is stipulated that when $H(X)$ is obtained, the residual can be obtained by linear transformation and activation function as follows:

$$F(X) = H(X) - X. \qquad (1)$$

The actual learned feature is as follows:

$$H(X) = F(X) + X. \qquad (2)$$

In the extreme cases, the convolutional layer implements the identity mapping even if $F(X) = 0$. The performance and characteristic parameters of the network remain unchanged. In general, $F(X) > 0$. The network can always learn new features, thus ensuring gradient transmission in backpropagation and eliminating the problems of network degradation and gradient disappearance.

2.3.2. SKNet Network.

SKNet [21] is an upgraded version of SENet [22], which is one of the visual attention mechanisms in the attention mechanisms. Convolution kernels of different sizes have different effects on targets of different scales. SKNet proposed a mechanism that not only takes into account the relationship between channels but also takes into account the importance of convolution kernels, that is, different images can obtain convolution kernels of different importance so that the network can obtain information of different receptive fields.

The SKNet network is formed by stacking multiple SK convolution units. The SK convolution operation consists of three modules, Split, Fuse, and Select, which contain
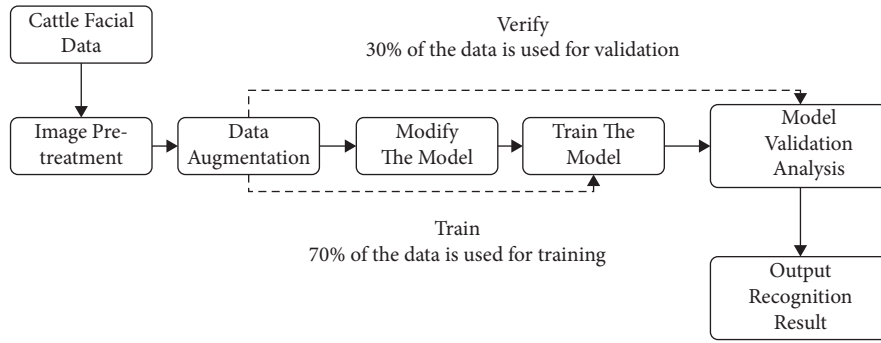
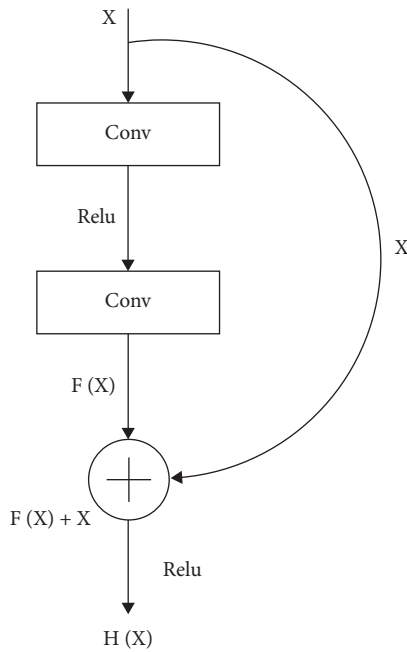FIGURE 2: Facial recognition process of cattle.

FIGURE 3: Residual block structure.

multiple branches. Take the two-branch SKNet network in Figure 4 as an example. First, the feature map X of size $c \times w \times h$ is subjected to group convolution and atrous convolution through the (3×3) and (5×5) size SK convolution kernels, respectively, through the spilt operation, output $\widetilde{U}$ and $\widehat{U}$. The Fuse operation fuses the two feature maps with element-wise summation and then generates a $c \times 1 \times 1$ feature vector S (c is the number of channels) through global average pooling. Feature maps S forms a vector Z after two full connection layers of dimensionality reduction and dimensionality enhancement. The select module regresses the vector Z to the weight information matrix a and matrix b between channels through 2 Softmax functions and uses a and b to weight the two feature maps $\widetilde{U}$ and $\widehat{U}$ and then sum to get the final output vector V. SKNet mechanism can not only make the network automatically learn the weight of the channel but also take into account the weight and importance of the two convolutions (convolution kernel). SK convolution unit not only uses the attention mechanism but also uses multibranch convolution, group convolution, and atrous convolution.

2.3.3. Model Building

(1) Model Improvements. The input image first passes through a 7×7 convolution layer and uses a large convolution kernel to retain the original image features; then, it uses a 3×3 maxpool with a stride of 2 to extract the feature map and compress the image. Then, enter four layers in turn; each layer includes a different number of SK-Bottleneck because the cattle's facial features are limited, so the model in this paper mainly extracts facial features through a shallow network, so the number of SK-Bottlenecks for the four layers is set to 3, 4, 1, and 1, where the branch of the SK module is set to 3, as shown in Figure 5(a). Facial features are extracted at multiple scales by integrating information of multiple receptive fields. Maxpool is used for fast connection in SK-Bottleneck, as shown in Figure 6(b). Atrous convolution is connected after the first layer to expand the receptive field without introducing parameters and accurately locate the target features. Each convolutional layer is connected to a BN layer. In the training phase, the cattle face dataset is small. In order to prevent overfitting, a dropout layer is added before the fully connected layer, and the global average pool is used to optimize the network structure and increase the generalization and antioverfitting ability of the model. Finally, use the Softmax classification layer for classification. Replace the ReLU activation function in the entire network with the ELU activation function, which is more robust to the vanishing gradient problem. Through the above methods to improve the recognition accuracy of network training, the improved network structure is shown in Figure 6.

(2) Activation Function. Traditional ResNet networks use rectified linear unit (ReLU) activation functions [23]. ReLU is simple, linear, and unsaturated. The algorithm can effectively alleviate gradient descent and provide sparse representation. The ReLU activation function is shown in the following equation:

$$ReLu(x) = \begin{cases} x, & x > 0, \\ 0, & x \leq 0. \end{cases} \tag{3}$$

It can be seen from formula (3) that when the value of x is 1, the gradient will disappear if it is too small. When the value of x is less than or equal to 0, as the training progresses, neurons will undergo apoptosis, resulting in the failure to update the weights.
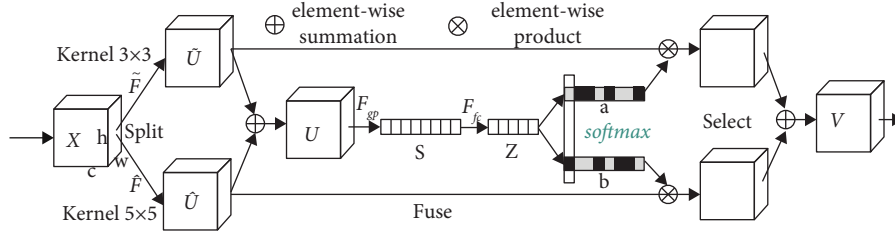
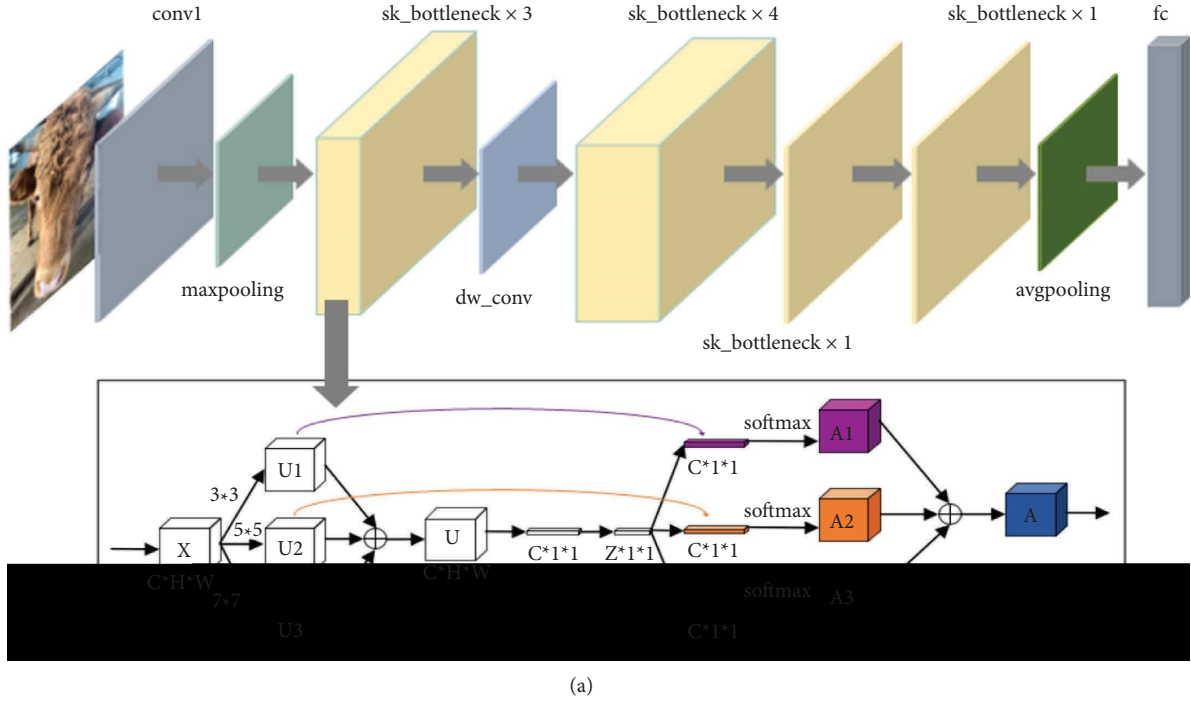FIGURE 4: Schematic diagram of SK convolution operation.

(a)

FIGURE 5: SK-ResNet network structure.

The ELU activation function [24] combines Sigmod and ReLU, with soft saturation on the left and desaturation on the right. The linear part on the right side makes the ELU more robust to input variations or noise. The output mean of the ELU is close to 0 and converges faster to solve the neuron death problem. The ELU activation function is shown in the following equation:

$$ELU = \begin{cases} x, & x > 0, \\ \alpha(e^x - 1), & x \le 0. \end{cases} \quad (4)$$

(3) Improved Quick Connect. In the original ResNet structure, when the dimension of x does not match the output dimension of $F(x)$, a shortcut connection is applied to x [25], and then, x is added to $F(x)$. Figure 6(a) is the default shortcut connection used in the original ResNet. The original shortcut connection uses a 1×1 convolutional layer; when the spatial size is reduced by a factor of two, a 1×1 convolutional layer with stride 2 skips 75% of the feature map activations, resulting in a significant loss of information. In addition, inputting 25% of the feature mapped activation obtained from the 1×1 convolutional layer to the next ResBlock introduces noise and information loss, which

negatively interferes with the main information flow of the network. The improved shortcut connection is shown in Figure 6(b), using spatial projection and channel projection; spatial projection uses a 3×3 max pooling layer with stride 2, and channel projection applies 1×1 convolution with stride 1 layer. Activation criteria for 1×1 convolutional layers are introduced via max pooling layers. Spatial projection not only guarantees all the information from the feature maps but also extracts the main features. The convolution kernel of the max pooling layer is consistent with the intermediate convolution kernel of ResBlock to ensure that element-wise addition is performed between elements in the same space, and the improved shortcut connection reduces information loss. A ResNet requires four shortcut connections. The shortcut connections used in this paper do not add any parameters to the model. The structure is shown in Figure 6.

3. Results and Discussion

3.1. Experimental Configuration. The computer used in this experiment had an Intel® Core™ i7-8700 CPU @3.29 GHz, 64 GB memory, and an NVIDIA GeForce GTX 1080Ti graphics card. A network training platform for deep learning
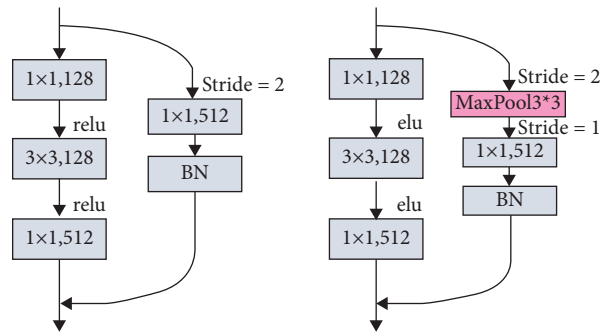
FIGURE 6: (a) Original shortcut link. (b) Improved shortcut link.

algorithms was built based on the Windows operating system and the PyTorch gpu1.8.0 framework, including Python version 3.8.3, integrated development environment PyCharm2020, Torch version 1.8.0, and cuDNN version 11.0.

In the experiment, to better evaluate the differences between real and predicted values, the batch training method was adopted. The other settings were as follows: loss function = cross-entropy loss, weight initialization method = Xavier, initialization deviation = 0, initial learning rate of the model = 0.001, batch size = 16, and momentum = 0.9; the model used the stochastic gradient descent (SGD) optimizer and Softmax classifier; the model was optimized by stochastic gradient descent; and the model was reduced by 0.1 every 10 iterations. When training and testing the model, the input image size was normalized to 224 × 224, a total of 51 epochs were trained, and, finally, the converged model was unified as the final saved model.

3.2. Experimental Results and Analysis. Based on the cattle facial recognition model and process of SK-ResNet, the SK module, the maximum pooling layer (maxpool) of the shortcut connection, and the ELU activation function are, respectively, explored to build the model. The influence of different modules on the model is shown in Table 1. Base represents the basic ResNet model with (3, 4, 1, 1) Bottle-Neck, and from Table 1, it can be seen that the use of the SK module in the base model improves the recognition accuracy by 1.69%, and there is a significant reduction in the model size and the number of model parameters. By adding maxpool to the shortcut part of the above model, the number of model parameters remains unchanged, while the accuracy is improved; the final model uses ELU, and the results show that the model accuracy has further improved and the growth rate of the number of model parameters is within our acceptable range. The final recognition accuracy of this model reaches 98.42%, which is 3% higher than the recognition accuracy of the base model. The model curve of this paper is shown in Figure 7.

In order to verify the effectiveness of the model, the model in this paper was compared with classic ResNet-50, SKNet, DenseNet, and GoogleNet on the constructed cattle face dataset constructed. The experimental results are shown in Table 2. Although the size, number of parameters, and

FLOPs of the GoogleNet model are slightly higher than those of this paper, the training state of the model is unstable, and the loss value of the model is higher than that of this paper, and the model in this paper has relatively higher recognition accuracy and stability. The model size, number of parameters, FLOPs, and loss values of ResNet-50, SKNet, and DenseNet models are all higher than those of this paper, and the recognition accuracy is also much lower than that of this paper's model. The final average accuracy of this paper's model is 98.42%, which shows that the SK-ResNet cattle facial recognition model constructed in this paper can guarantee the recognition accuracy while reducing the number of model parameters and can identify individual cattle faster. The result curves of the model in this paper and the comparison model are shown in Figure 8.

In order to observe and reach the purpose of correct classification, the observation network which focuses more on the area, we have adopted a class activation mapping (CAM) to determine whether a high-response area falls under our concerns. The principle is that for a CNN model, global average pooling (GAP) is performed on the last feature map to calculate the mean value of each channel and then mapped to the class score through the fully connected (FC) layer to find the argmax, and calculate the output of the largest class relative to the last one. The gradient of a feature map, and then visualize the gradient on the original image. Intuitively, it is to see which part of the high-level features extracted by the network has the greatest impact on the final classifier [26]. Figure 9 is a partial heat image of the face of cattle, and it can be seen that the dark color is concentrated in the facial area of cattle, which proves that the model in this study is trained to recognize the facial features of cattle, and this result verifies the accuracy of the model for facial recognition proposed in this study.

3.3. Model Generalizability. Apply the model in this paper to facial recognition of other animals and compare the results with similar model results. The proposed model was applied to the facial recognition of other animals, and its results were compared with those of similar models. The experimental implementation process included data preparation, network configuration, network model training, model evaluation, and model prediction. There were three tests, as follows. The proposed models and the ResNet, SENet, GoogleNet, and

TABLE 1: Results of the effect of different structures on the mode.

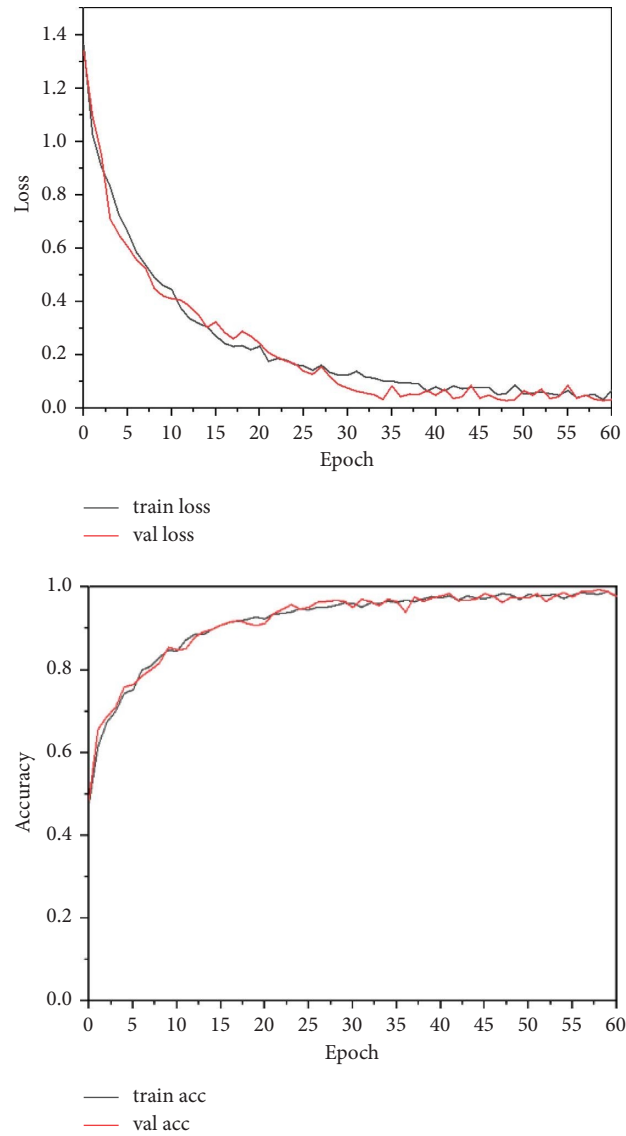| Model	GFLOPs	Params size (MB)	Params	Accuracy (%)
Base	2.57	34.37	9,009,222	95.42
Base_4SK	1.61	21.87	5,732,614	97.11
Base_4SK_maxpool	1.61	21.87	5,732,614	97.89
Base_4SK_maxpool_ELU	1.65	22.56	5,913,611	98.42

FIGURE 7: The model in this paper identifying the accuracy curve and loss curve.

TABLE 2: Cattle facial recognition results.

Model	FLOPs	Size (MB)	Params	Loss (%)	Accuracy (%)
DenseNet	2.86	26.58	6,960,006	7.91	95.84
GoogleNet	1.5	21.39	5,606,054	7.72	95.95
ResNet-50	4.11	89.72	23,520326	17.01	93.53
SKNet	2.38	50.66	13,281,478	10.27	94.26
SK-ResNet	1.65	22.56	5,913,611	3.38	98.42

DenseNet network models were trained on the Long White Pig facial dataset and their results compared. The experimental results show that (Table 3) the accuracy of the proposed model is 98.57% and the loss is only 9.6 on the self-built Long White Pig facial dataset. Compared with other models, the results show that the accuracy of the proposed model is higher than that of the other four models, and the loss of the model is much lower than that of the other four models. The model in this paper is compared with the classic
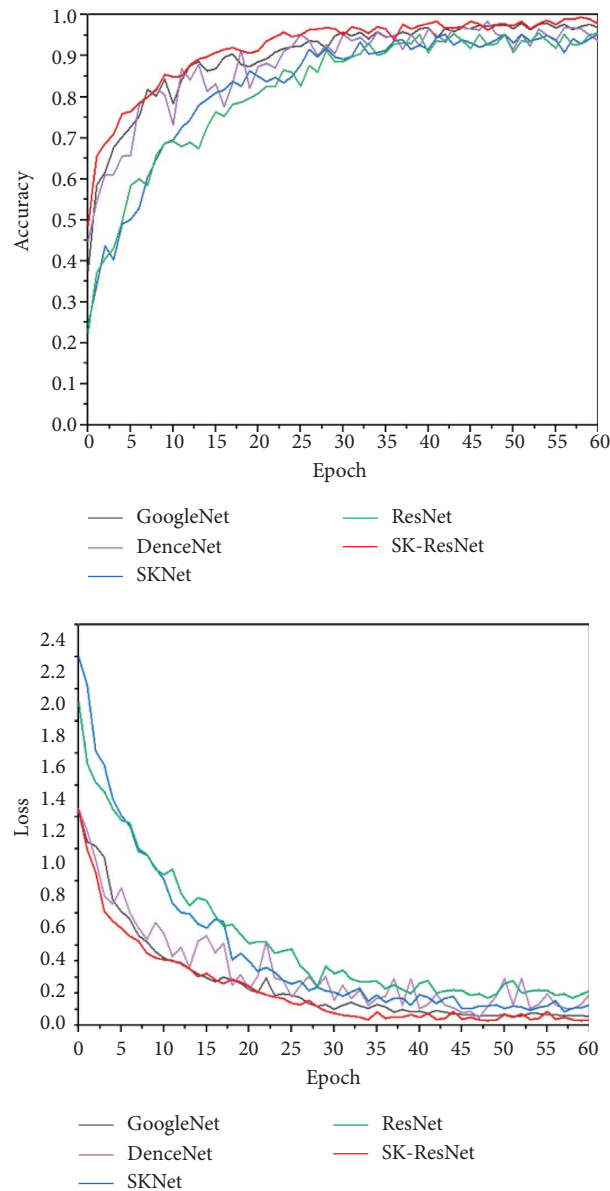
FIGURE 8: Compare the recognition accuracy curve and loss curve of the model.

FIGURE 9: Facial heat map of cattle.

TABLE 3: Long white pig facial recognition results.

Methods	GoogleNet	DenseNet	ResNet	SKNet	SK-ResNet
Accuracy (%)	96.79	95.73	94.31	96.42	98.57
Loss (%)	13.47	25.12	25.76	17.96	9.6

TABLE 4: Identification results of sheep face breeds.

Methods	GoogleNet	DenseNet	ResNet	SKNet	Abu Jwade et al.	SK-ResNet
Accuracy (%)	96.62	95.24	91.96	94.94	95	97.02
Loss (%)	12.15	15.84	34.49	14.5	16.42	10.42

TABLE 5: Recognition results of a pure-colour cattle facial dataset in different animal facial recognition models.

Methods	Proposed AlexNet	Proposed VGG-16	SK-ResNet
Accuracy (%)	93.95	94.21	98.42
Loss (%)	22.70	14.78	3.38
Size (MB)	69.09	512.32	22.56

models such as ResNet and the sheep face breed recognition method proposed by Abu Jwade et al. [27] on the sheep face datasets of four breeds collected. Table 4 shows that the accuracy rate of this model on this data set is 97.02. Compared with the model proposed by the original author, the recognition effect of this model is the best, which is 2% higher than that of the model proposed by the original author and 6% lower than that of the model proposed by the original author. From Table 4, it can be seen that the recognition effect of this model is better than other models, and the loss is the smallest.

To verify the effectiveness of the model proposed in this study, the solid-color cow face dataset was trained on the pig face recognition model proposed by Yan et al. [28] and the sheep face recognition model proposed by Abu Jwade et al. [27]. The experimental results are shown in Table 5. The recognition accuracy of the model proposed in this study is much higher than that of the other two models, and the loss value and model size of the model are smaller than those of the other models, so the model proposed in this study has a better recognition effect than the other models and is suitable for animal face recognition.

4. Conclusions

In recent years, with increase in the amount of cattle breeding, the monitoring of individual information and health management has become extremely important. Therefore, there is an urgent need for an intelligent, intensive, and standardized method of managing individual cattle on farms. The paper aims to improve the accuracy, stability, and speed of cattle identification to promote the development of intelligent cattle breeding. In this study, this paper generates a cattle face dataset using images of cattle faces obtained from different angles and proposes an improved ResNet-based recognition model to extract cattle facial features. The model uses ResNet-50 as the base model, in which the SK-Bottleneck module is used to fuse information from multiple sensory fields to extract facial features at multiple scales. Second, a max-pooling layer is used in the shortcut connection to reduce information loss. Finally, an ELU activation function is used in the network to reduce vanishing gradients, prevent overfitting, speed up convergence, and improve the generalization ability of the model. The SK-ResNet-based recognition model is

trained with the cattle face dataset, which proves that the individual cattle can be accurately identified. The improved method is compared with the existing models ResNet-50, SKNet, DenseNet, and GoogleNet and found to be more accurate in recognition while having fewer parameters and faster computation. The results show that the method achieves an average recognition accuracy of 98.42% on a dataset of 5677 images. To verify the generalizability of the proposed model, a Long White Pig facial dataset (with fewer facial features) and a sheep face dataset were used; accuracies of 98.57% and 97.02% were obtained, respectively. Hence, the recognition accuracy of the proposed model is higher than that of the other models. The experimental results show that the proposed model can accurately identify individual cattle, giving it great potential for application to cattle breeding and the facial recognition of other animals with high facial similarity.

Data Availability

The [cow face image] data used to support the results of this study were created by the authors themselves through video intercepts and can be obtained from the authors upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

Acknowledgments

References

[1] B. B. Xu, W. S. Wang, and L. F. Guo, "Progress and prospects of contactless-based cattle identification research," *China Agricultural Science and Technology Heald*, vol. 22, no. 07, pp. 79–89, 2020.

[2] D. J. He, D. Liu, and K. X. Zhao, "Research progress on intelligent sensing and behavior detection of animal information in precision animal husbandry," *Journal of Agricultural Machin-ery*, vol. 47, no. 5, pp. 231–244, 2016.

[3] Y. K. Sun, Y. J. Wang, and P. J. Huo, "Research progress on individual identification methods of dairy cattle and their applications," *Journal of China Agricultural Uni-versity*, vol. 24, no. 12, pp. 62–70, 2019.

[4] G. T. Fosgate, A. A. Adesiyun, and D. W. Hird, "Ear-tag retention and identification methods for extensively managed

water buffalo (Bubalus bubalis) in Trinidad," *Preventive Veterinary Medicine*, vol. 73, no. 4, pp. 287–296, 2006.

[5] G. Q. Li, M. Zhou, and F. Y. Chen, "Research on breeding management system of small and medium-scale cattle farm based on RFID handheld termi-nal," *Jiangsu Agricultural Sciences*, vol. 49, no. 22, pp. 192–197, 2021.

[6] D. J. Cook, J. C. Augusto, and V. R. Jakkula, "Ambient in-telligence: t," *Pervasive and Mobile Computing*, vol. 5, no. 4, pp. 277–298, 2009.

[7] T. P. Moreira, M. L. Perez, R. d O. Werneck, and E. Valle, "Where is my puppy? retrieving lost dogs by facial features," *Multimedia Tools and Applications*, vol. 76, no. 14, pp. 15325–15340, 2017.

[8] M. F. Hansen, M. L. Smith, L. N. Smith et al., "Towards on-farm pig face recognition using convolutional neural net-works," *Computers in Industry*, vol. 98, pp. 145–152, 2018.

[9] L. Yao, Z. Hu, and C. Liu, "Cow face detection and recognition based on automatic feature extraction algorithm," in *Proceedings of the ACM Turing Celebration Conference*, pp. 1–5, China, 2019.

[10] M. Marsot, J. Mei, X. Shan et al., "An adaptive pig face recognition approach using convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 173, pp. 105386–386, 2020.

[11] R. W. Bello, A. Z. H. Talib, and A. S. A. B. Mohamed, "Deep belief network approach for recognition of cow using cow nose image pattern," *Walailak Journal of Science and Tech-nology*, vol. 18, no. 5, p. 14, 2021.

[12] A. M. Abdul-Hadi, S. H. Abdulhussain, and B. M. Mahmmod, "On the computational aspects of Charlier polynomials," *Cogent Engineering*, vol. 7, no. 1, Article ID 1763553, 2020.

[13] S. H. Abdulhussain and B. M. Mahmmod, "Fast and efficient recursive algorithm of Meixner polynomials," *Journal of Real-Time Image Processing*, vol. 18, no. 6, pp. 2225–2237, 2021.

[14] B. M. Mahmmod, S. H. Abdulhussain, T. Suk, and A. Hussain, "Fast computation of hahn polynomials for high order mo-ments," *IEEE Access*, vol. 10, pp. 48719–48732, 2022.

[15] S. Yuheng, M. Ye, F. Qin et al., "Deer body adaptive threshold segmentation algorithm based on colour space," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 1317–1328, 2020.

[16] K. He, X. Zhang, and S. Ren, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770–778, Washington, June 2016.

[17] L. Meng, X. Y. Guo, and J. J. Du, "A lightweight CNN crop disease image recognition model," *Jiangsu Agri-cultural Journal*, vol. 37, no. 5, pp. 1143–1150, 2021.

[18] M. Gour, S. Jain, and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image clas-sification," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 621–635, 2020.

[19] S. Q. Yang, J. C. Liu, and K. K. Xu, "Improved CenterNet-based remote sensing image recognition of maize stamens by UAV," *Journal of Agricultural Ma-chinery*, vol. 52, no. 9, pp. 206–212, 2021.

[20] Y. Cheng and H. Wang, "A modified contrastive loss method for face recognition," *Pattern Recognition Letters*, vol. 125, pp. 785–790, 2019.

[21] X. Li, W. Wang, and X. Hu, "Selective kernel networks," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 510–519, 2019.

[22] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation net-works," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, Washington, June 2018.

[23] A. F. Agarap, "Deep learning using rectified linear units (relu)," vol. 1803, arXiv preprint arXiv, Article ID 08375, 2018.

[24] D. A. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," vol. 1511, 2015 arXiv preprint arXiv, Article ID 07289.

[25] I. C. Duta, L. Liu, and F. Zhu, "Improved residual networks for image and video recognition," in *Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR)*, pp. 9415–9422, IEEE, Milan, Italy, January 2021.

[26] B. Zhou, A. Khosla, and A. Lapedriza, "Learning deep features for discriminative localization," *IEEE Computer Society*, 2016.

[27] S. Abu Jwade, A. Guzzomi, and A. Mian, "On farm automatic sheep breed classification using deep learning," *Computers and Electronics in Agriculture*, vol. 167, Article ID 105055, 2019.

[28] H. Yan, Q. Cui, and Z. Liu, "Pig face identification based on improved AlexNet model," *INMATEH*, vol. 61, no. 2, pp. 97–104, 2020.