

Sika Deer Facial Recognition Model Based on SE-ResNet

He Gong^{1,3,4}, Lin Chen¹, Haohong Pan¹, Shijun Li^{2,5}, Yin Guo¹, Lili Fu¹, Tianli Hu^{1,3,4,*}, Ye Mu^{1,3} and Thobela Louis Tyasi⁶

¹College of Information Technology, Jilin Agricultural University, Changchun, 130118, China

²College of Electronic and Information Engineering, Wuzhou University, Wuzhou, 543003, China

³Jilin Province Agricultural Internet of Things Technology Collaborative Innovation Center, Changchun, 130118, China

⁴Jilin Province Intelligent Environmental Engineering Research Center, Changchun, 130118, China

⁵Guangxi Key Laboratory of Machine Vision and Intelligent Control, Wuzhou, 543003, China

⁶Department of Agricultural Economics and Animal Production, University of Limpopo, 0727, Polokwane, South Africa

*Corresponding Author: Tianli Hu. Email: hutianli@jlau.edu.cn

Received: 11 January 2022; Accepted: 08 March 2022

Abstract: The scale of deer breeding has gradually increased in recent years and better information management is necessary, which requires the identification of individual deer. In this paper, a deer face dataset is produced using face images obtained from different angles, and an improved residual neural network (ResNet)-based recognition model is proposed to extract the features of deer faces, which have high similarity. The model is based on ResNet-50, which reduces the depth of the model, and the network depth is only 29 layers; the model connects Squeeze-and-Excitation (SE) modules at each of the four layers where the channel changes to improve the quality of features by compressing the feature information extracted through the entire layer. A maximum pooling layer is used in the ResBlock shortcut connection to reduce the information loss caused by messages passing through the ResBlock. The Rectified Linear Unit (ReLU) activation function in the network is replaced by the Exponential Linear Unit (ELU) activation function to reduce information loss during forward propagation of the network. The preprocessed 6864 sika deer face dataset was used to train the recognition model based on SE-Resnet, which is demonstrated to identify individuals accurately. By setting up comparative experiments under different structures, the model reduces the amount of parameters, ensures the accuracy of the model, and improves the calculation speed of the model. Using the improved method in this paper to compare with the classical model and facial recognition models of different animals, the results show that the recognition effect of this research method is the best, with an average recognition accuracy of 97.48%. The sika deer face recognition model proposed in this study is effective. The results contribute to the practical application of animal facial recognition technology in the breeding of sika deer and other animals with few distinct facial features.

Keywords: Sika deer facial recognition model; ResNet-50; se module; shortcut connection; ELU



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Sika deer have great economic value as food and medicine [1]. In recent years, with the large increase in the amount of sika deer breeding, there is an urgent need for an automated, information-based and refined identification method for individual sika deer. This would allow the recording of individual information and tracking of the health status of sika deer, so as to facilitate the construction and improvement of a deer product quality tracking system. Traditional individual identification methods include external marking (e.g., ear tags) and the use of externally attached devices such as radio frequency identification (RFID) tags. These are prone to problems such as tag loss, tampering of tag content, system crashes, server intrusion attacks and high cost [2]. With the rapid progress in deep learning, facial recognition technology based on computer vision has become a focus of research [3]. Unlike traditional individual recognition methods, facial recognition methods can effectively avoid animal stress while tracking breeding. Facial recognition technology to monitor sika deer not only saves labour costs but also aids their breeding. The facial information of sika deer can be easily captured when they are feeding and drinking water. Therefore, sika deer individuals can be well recognized using facial recognition technology.

At present, in the animal husbandry industry, facial recognition technology is mostly used on animals such as pigs and cattle that have large differences in facial characteristics [4–6]. It has not been widely applied to animals with high facial similarity. In 2021, Otani et al. developed an individual recognition system for Japanese macaques that enhanced a small sample of macaque data and trained four individual recognition models. Meanwhile, a simple shallow model was used to increase the accuracy of individual recognition under limited conditions. The identification system aims to reduce the space, time and capacity constraints of traditional individual identification methods [7]. In 2020, Hou et al. established a giant panda individual recognition model using a convolutional neural network based on deep learning. An improved Visual Geometry Group (VGG) network was trained with pictures of captive giant panda faces to recognize individuals, saving labour and time. However, due to their facial similarity, the final recognition accuracy was low [8]. In 2020, Guo et al. used a system developed for automatic face detection and individual recognition based on video and still images and deep learning, which can be used for a variety of species. The system uses a data set containing carnivores and primates for training and testing, and can identify primates at an accuracy rate is 94% [9]. In 2020, Yan, H., Cui, Q. et al. proposed a pig face recognition method based on an improved AlexNet model for pig faces, introducing a spatial attention module (SAM) into the AlexNet model, which eventually had better recognition results on the pig face dataset [10]. In 2019, Chen et al. used a Convolution Neural Network (CNN)-based pure-colour cow face detection model combined with a PnasNet-5 recognition model to improve recognition performance. A face recognition experiment was carried out using face detection data of about 50,000 annotated pure-colour cows. The accuracy of recognizing pure-colour cow faces reached 94.1%. Facial recognition technology can not only recognize individual animals but can also achieve animal species classification [11]. In 2019, Jwade et al. Established a prototype computer vision system and database at a sheep farm. Face images of four sheep breeds were collected and their breeds were marked by experts. By fine-tuning the last six layers of a VGG-16 network and training the sheep breed classifier using machine learning and computer vision, a final classification accuracy of 95% was obtained [12].

This paper builds a sika deer facial recognition model based on ResNet-50, which can extract facial information more effectively, reduce the number of parameters and improve the accuracy rate. High performance is necessary due to the small number of facial features of sika deer. The model was trained using the sika deer facial dataset constructed in this paper, which was collected under different angles at a sika deer farm. As shown in Fig. 1, the model was validated by training it using facial

datasets with complex backgrounds and the backgrounds removed, which proved the effectiveness of the model.

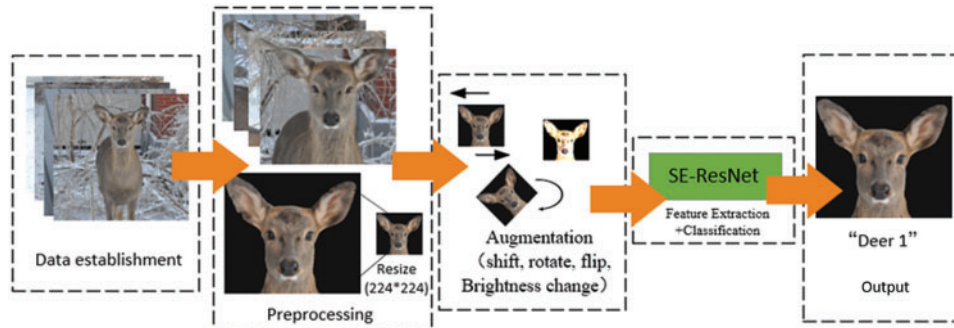


Figure 1: Sika deer facial recognition process

2 Model Improvement

He et al. [13] confirmed that, compared with other deep networks, the ResNet network will get better results with residual mapping and shortcut connections, and its training is easier. Therefore, ResNet is often used for plant disease spot classification [14,15], pathology image classification [16–19], remote sensing image classification [20,21], facial recognition [22,23] and other fields. Compared with other models, ResNet has the advantage of being able to increase the network depth to obtain more rich and complex features, while solving the problems of network gradient disappearance and training difficulties. The ResNet model is implemented by skipping connections on two to three layers and includes ReLU [24] and batch normalization in its architecture. Compared with other models, ResNet performs better in image classification and can extract image features well [25,26], so was considered suitable for this research. In this paper, ResNet is adopted as the main structure of feature extraction, so that the ResBlock of the model can obtain more key feature values by recognizing the superficial facial features of sika deer.

ResNet also has some shortcomings. When recognizing animals with similar faces, ResNet has low accuracy due to the deer's lack of facial features. In the original ResNet architecture, when the input ResBlock does not match the output ResBlock, a shortcut connection is used to match the input and output dimensions. The original shortcut connection uses a 1×1 convolutional layer with a step size of 2 to align the space size between the input ResBlock and the output ResBlock. The 1×1 convolutional layer reduces the space size by two times and skips part of the activated feature map. This operation causes serious loss of information and introduces noise, which interferes with the main information flow through the network. The ResNet network will appear to suffer neuronal apoptosis due to the use of the ReLU activation function during reverse transmission.

The convolution kernel in the convolutional layer captures local spatial relationships in the form of feature maps. Each feature map is considered independent and equal, which may cause unimportant features to propagate through the network, thereby affecting accuracy. Therefore, in order to model the relationship between feature maps, four SE modules are added to the four layers where the ResNet channel has changed. The features in the same channel are recalibrated and the global information is used to compare the information in the feature map. Richer features are weighted to improve their quality. Adding one, two or three SE modules cannot model all the feature information and is not sufficient for achieving the best recognition effect. Only after the four layers are connected

are SE modules used to model all the feature information obtained from the four layers. The quality of the network generated representation is improved through the interdependence between channels with a convolution property. Adding four SE modules can significantly improve accuracy at a small calculation cost. Compared with connecting SE modules after each ResBlock in SENet [27], it not only reduces the number of parameters but also reduces the information loss caused by weight calibration of different channels using global average pooling. In this paper, the network shortcut connection uses the maximum pooling layer, which helps to improve the translation invariance of the network and improves the recognition performance [28]. In addition, a 3×3 maximum pooling layer with a step size of 2 is used to make the convolution kernel of the maximum pooling layer. Its size is equal to that of the network convolution kernel (3×3 conv), ensuring that the sum of the elements calculated in the same spatial window is added. The output average value of the ELU activation function is close to 0, which makes the normal gradient closer to the natural gradient by reducing the influence of bias offset so as to accelerate the learning of the mean value to zero. The ELU activation function will be saturated to a negative value under a small input so as to reduce the forward propagation variation information [29].

The input image first passes through a convolutional layer with a size of 7×7 and uses a large convolutional kernel to retain the original features of the image at a large scale. Then, it uses a 3×3 maximum pooling layer with a step size of 2 to extract features from the feature map to compress the image. Then, we enter the four layers in turn; each layer includes a conv_block and a different number of identity_blocks. Because the facial features of sika deer are limited, the model presented in this paper mainly extracts the facial features through a shallow network. Therefore, the numbers of identity_blocks in Layer1 and Layer2 are set to 2 and 3, respectively. Layer3 and Layer4 only use the conv_block, and a Projection shortcut in the conv_block uses the largest pooling layer. As shown in Fig. 2b, to reduce information loss, the network uses the original identity_block; its structure is shown in Fig. 2c. The SE module is connected after each layer and all feature information obtained by each layer is compressed and excited to extract important features. Each convolutional layer is connected to the batch normalization layer. In the training phase, the sika deer facial recognition data set is small. Therefore, to prevent overfitting, a dropout layer is added before the full connection layer and the global average pool is used to optimize the network structure and increase the generalization of the model and its resistance to overfitting. Finally, a softmax classification layer is used for classification. We replace the ReLU [24] activation function in the whole network with the ELU activation function, which is more robust to the vanishing gradient problem. The above methods are used to improve the recognition accuracy of the network training. The improved network structure is shown in Fig. 2a.

3 Experiment

3.1 Data Collection and Processing

The data collection site was the Dongao Deer Industry Group Farm, Shuangyang District, Changchun City, Jilin Province, China. Ten sika deer were photographed from different angles at a resolution of 4288×2848 pixels in Joint Photographic Experts Group (JPG) format using a Nikon D90 digital camera in December 2020. This produced a sika deer data set, a sample of which is shown in Fig. 3. We aimed to prevent image saturation, avoid direct sunlight on the face in the images, and remove complex backgrounds [30] by extracting the face region of the image. A total of 2145 facial images were collected, which were randomly divided into training and verification sets at a ratio of 8:2 (1716 training images and 429 verification images). Usually, model training requires a sufficient sample size to avoid overfitting of the model; therefore, the training set was expanded. The traditional ways of expanding a dataset are image flipping, random cropping and colour dithering. In this study,

the image rotation method was chosen to expand the training set. The images were randomly flipped from three angles, which tripled the sample size of the training set. A 6864-image training set was obtained after data enhancement.

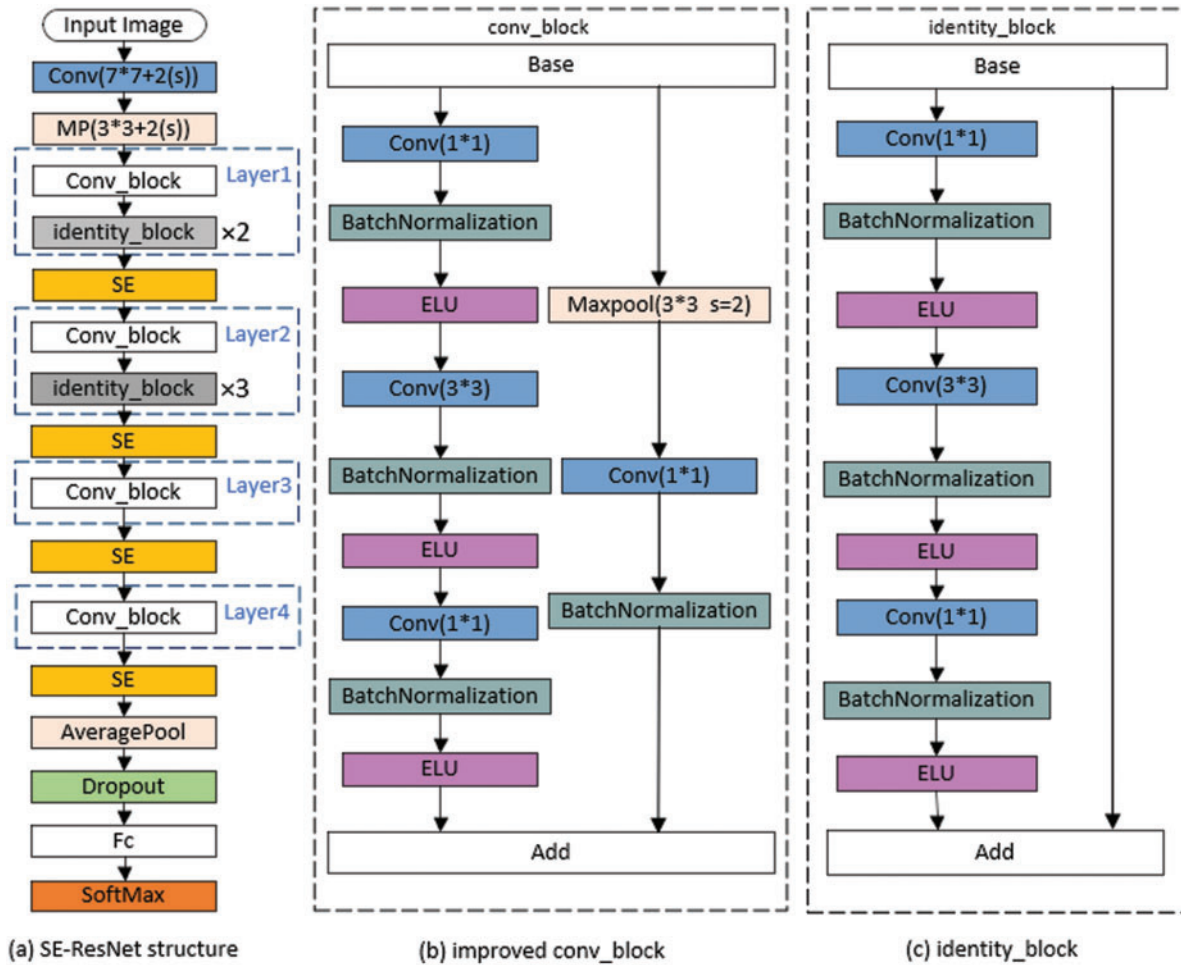


Figure 2: SE-ResNet model structure

3.2 Experimental Configuration

The computer used in this experiment had an Intel® Core™ i7-8700 CPU @3.20 GHz, 64 GB memory, and an NVIDIA GeForce GTX 1080Ti graphics card. A network training platform for deep learning 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.



Figure 3: Sample of sika deer facial data set constructed in this paper

In the experiment, to better evaluate the differences between real and predicted values, the batch training method was adopted. The other settings were: loss function = cross-entropy loss, weight initialization method = Xavier, initialization deviation=0, initial learning rate of the model=0.01, batch size = 16, 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.

3.3 Experimental Results and Analysis

Based on the basic ResNet model, this paper uses different numbers of SE initial modules to experiment on the constructed data set. The experimental results are shown in [Tab. 1](#). According to [Tab. 1](#), after adding one SE module to the model, the recognition accuracy of the model is improved. When adding two and three SE modules, the recognition effect is suppressed. When using the initial values of four SE modules, the model has better results and the model size is also within our acceptable range. This article also uses SENet to connect SE modules after all ResBlocks, but this operation adds a lot of parameters and the model recognition accuracy is lower than that of the model that connects four SE initial modules. Therefore, this paper uses four SE initial modules to build the model. It can be seen from [Tab. 1](#) that the accuracy of the model after adding four SE modules reaches 94.95%, which is higher than the initial value of the basic model.

Table 1: Influence of the number of different SE modules on the recognition results

Model	GFLOPs	Params size (MB)	Params	Accuracy (%)
1SE-ResNet	2.578	34.43	9,025,610	92.74
2SE-ResNet	2.578	34.55	9,058,378	90.85
3SE-ResNet	2.579	35.05	9,189,450	90.53
4SE-ResNet	2.579	37.05	9,713,738	94.95
All-ResNet	2.59	37.52	9,836,731	94.01

Based on the proposed SE-ResNet sika deer facial recognition model and recognition process, the SE module, shortcut connection and ELU activation function are explored for model construction. The influences of the different modules on the model are shown in Tab. 2. “Base” represents the basic model. As can be seen from Tab. 2, the SE module helps to improve the model’s recognition effect. After adding the SE module to the basic model, the recognition accuracy is improved by 4.41%. Adding maxpool to the shortcut connection part of the above model ensures that the growth rate of the model parameters is within our acceptable range while accuracy is improved. In the end, the model replaces ReLU with the ELU activation function. The model results confirm that this has the effect of improving recognition accuracy. Finally, the recognition accuracy of the model in this paper reaches 97.47%, which is 6.93% higher than that of the original model. The accuracy curve and loss value curve of the model are shown in Fig. 4.

Table 2: Effects of different model structures under a single background

Model	GFLOPs	Params size (MB)	Params	Accuracy (%)
Base	2.57	34.4	9,017, 418	90.37
Base_4SE	2.57	37.05	9,713, 738	94.95
Base_4SE_maxpool	2.58	37.08	9,719, 627	95.38
Base_4SE_maxpool_ELU	2.58	37.08	9,719, 627	97.48

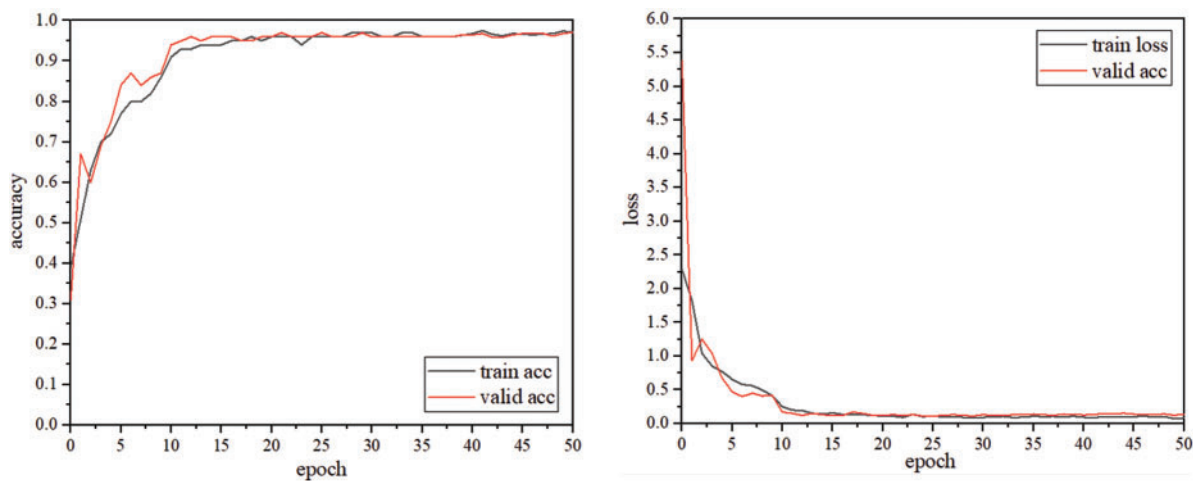


Figure 4: Recognition accuracy curve and loss curve of the proposed model

To verify the effectiveness of the proposed model, it was compared with the classical ResNet-50, SENet, DenseNet and GoogleNet models on the constructed sika deer facial dataset using the experimental environment and parameter settings described above. The results are shown in Tab. 3. The model size and the numbers of parameters and Floating point of operations (FLOPs) of DenseNet and GoogleNet are slightly lower than those of this paper; however, the models are unstable and their loss values are higher than those of the proposed models, which also have higher recognition accuracy and stability. The model size, number of parameters, FLOPs and loss values of the ResNet-50 and SENet models are higher than those of this paper, and the recognition accuracy is lower. It can be seen that the proposed SE-ResNet-based sika deer facial recognition model can greatly reduce the number

of model parameters while ensuring recognition accuracy, and can identify sika individuals faster. The accuracy and loss curves for the proposed and comparison models are shown in Fig. 5.

Table 3: Sika deer facial recognition results under simple backgrounds

Model	GFLOPs	Size (MB)	Params	Loss (%)	Accuracy (%)
DenseNet	2.86	26.60	6,964, 106	18.33	95.27
GoogleNet	1.50	21.40	5,611, 179	15.38	96.84
ResNet-50	4.10	89.75	23,528, 522	29.81	90.54
SENet	4.11	99.42	26,061, 563	26.01	92.11
SE-ResNet	2.58	37.08	9,719, 627	9.7	97.48

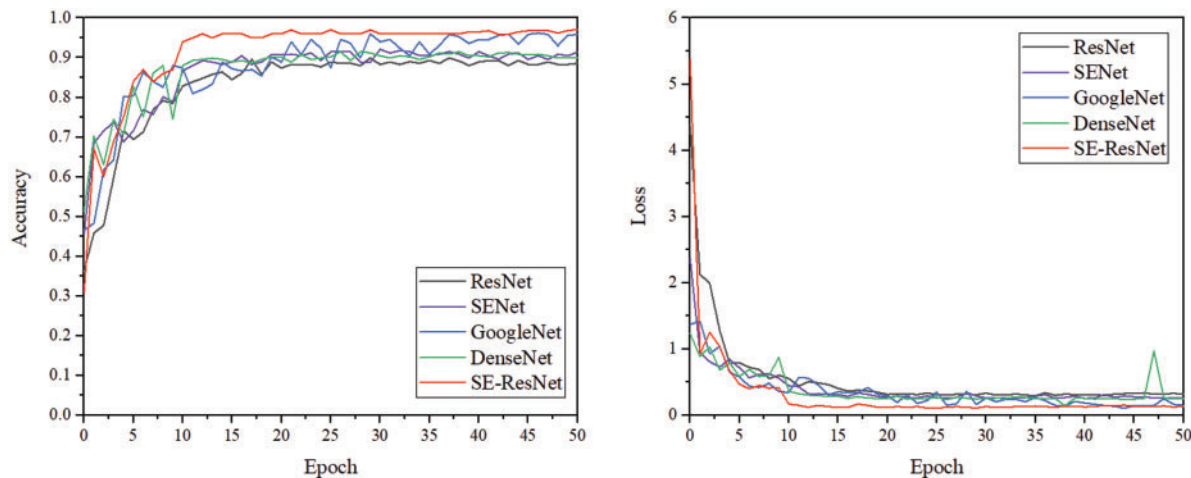


Figure 5: Recognition accuracy and loss curves of the comparison models

The confusion matrix shown in Fig. 6 shows almost 100% recognition rates for most sika deer, such as numbers 0, 1, 4, 7, 8 and 9, and relatively low recognition rates for others, such as numbers 2 and 3 (94% and 96% accuracy, respectively). As observed from the confusion matrix, in the process of model identification, 5.7% of No. 2 deer images were confused with No. 4 deer, 1.9% of No. 3 deer images were confused with No. 1 deer, and 1.9% of No. 3 deer images were confused with No. 6 deer. There are various potential causes for this phenomenon, such as the amount of training data, the angle of the face and changes in lighting, etc. If more data is used for training while taking these factors into account, it is expected that recognition accuracy will improve. In addition, the image recognition process of the model proposed in this study was visualised, and the heat map of the sika deer facial image for the recognition process of the model was output, as shown in the heat map of the sika deer face in part of Fig. 7, the dark colours are concentrated in the sika deer facial region, which proves that the model in this study is trained to recognize the sika deer facial features, and this result verifies the accuracy of the model proposed in this study for facial recognition.

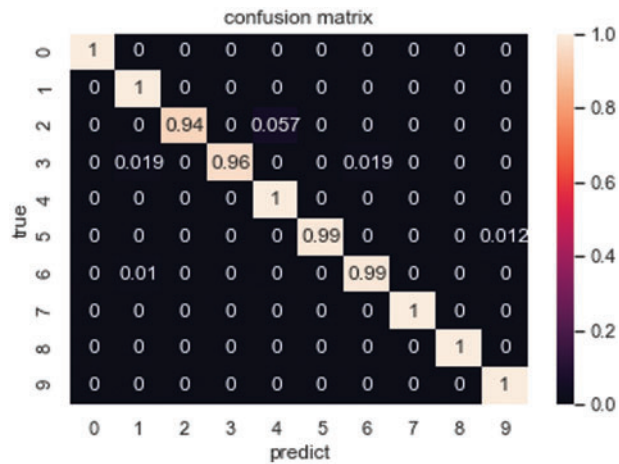


Figure 6: Standardized confusion matrix based on SE-ResNet

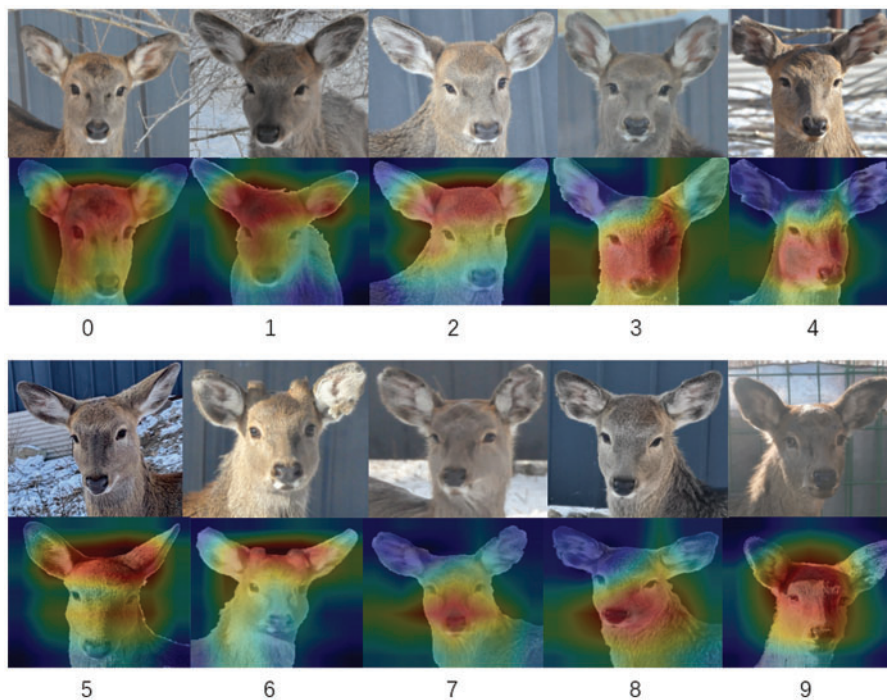


Figure 7: Sika deer facial heat maps

To verify the recognition effect of this model on sika deer faces photographed with a complex background, the model was trained with sika deer face images with complex backgrounds retained. The results of the experimental comparison are shown in [Tab. 4](#), which show that the recognition accuracy of the proposed model reached 96.28%, has a small loss value and is most suitable for use in a real sika deer breeding environment.

Table 4: Sika deer facial recognition results of various models using images with complex backgrounds

Parameter	GoogleNet	DenseNet	ResNet	SENet	SE-ResNet
Accuracy (%)	95.38	91.78	89.37	91.86	96.28
Loss (%)	15.12	24.37	31.37	24.12	13.24

3.4 Model Generalizability

The proposed model was applied to the facial recognition of other animals and its results 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 DenseNet network models were trained on the Long White Pig facial dataset and their results compared. The proposed model and classical models, such as ResNet and the sheep face breed recognition method proposed by Jwade et al., were compared on a dataset of sheep faces of four breeds [12]. The proposed model was compared with classical networks, such as ResNet and the pure-colour cattle face recognition method proposed by Chen et al., on a pure-colour cattle face dataset [11]. [Tabs. 5–7](#) show that the model in this paper outperforms all other models in recognition and has the lowest loss values.

Table 5: Comparison of Long White Pig facial recognition results

Parameter	GoogleNet	DenseNet	ResNet	SENet	SE-ResNet
Accuracy (%)	96.79	95.73	94.31	95.02	98.58
Loss (%)	13.47	25.12	25.76	18.75	12.91

Table 6: Comparison of sheep face recognition results

Parameter	GoogleNet	DenseNet	ResNet	SENet	Jwade et al.	SE-ResNet
Accuracy (%)	97.62	95.24	91.96	96.13	95	97.91
Loss (%)	10.15	15.84	34.49	11.37	16.42	9.18

Table 7: Comparison of pure-colour cattle facial recognition results

Parameter	GoogleNet	DenseNet	ResNet	SENet	Chen et al.	SE-ResNet
Accuracy (%)	98.68	98.42	97.36	97.89	95.53	99.47
Loss (%)	3.19	8.09	19.46	7.43	16.51	1.31

In order to verify the effectiveness of the model proposed in this study, the sika deer face dataset was trained on the pig face recognition model proposed by Yan et al. [10], the pure-colour cattle face recognition model proposed by Chen et al. [11] and the sheep face breed recognition sheep face recognition model proposed by Jwade et al. [12]. The experimental results are shown in [Tab. 8](#). The

recognition accuracy of the model proposed in this study is much higher than that of the other three 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 better recognition effect than the other models and is suitable for animal face recognition.

Table 8: Recognition results of sika deer facial dataset in different animal facial recognition models

Methods	Proposed AlexNet	PrimNet	Proposed VGG-16	SE-ResNet
Accuracy (%)	93.78	94.63	95.53	97.48
Loss (%)	19.51	15.42	16.51	9.7
Size (MB)	69.09	89.74	512.32	37.08

4 Conclusion

In recent years, with increases in the amount of sika deer breeding, the monitoring of individual information, health management and deer product quality control have become extremely important. Therefore, there is an urgent need for an intelligent, intensive and standardized method of managing individual sika deer on farms. Accordingly, this paper proposes a sika deer face recognition model based on SE-ResNet. A large number of sika deer images were taken from different angles (front, side, top and bottom) to create a face data set with the backgrounds removed. By obtaining faces with a single background, we can extract the contours, ears, and facial features such as the eyes and nose to recognize individual sika deer. By connecting the SE module behind the four layers of ResNet, the recognition accuracy is improved by making full use of the model channel information. The maxpool layer is added in the shortcut connection to reduce the loss of feature information. At the same time, the ELU activation function is used in the model to further improve accuracy without adding parameters. At the same time, to verify the effectiveness of the model, it was compared with four classic network models (DenseNet, ResNet-50, SENet and GoogleNet) on different data sets. The recognition accuracy of the proposed model was 97.48% with sika deer faces with a single background and 96.28% under complex backgrounds, which is better than other models. To verify the generalizability of the proposed model, a Long white pig dataset (with fewer facial features), a sheep face dataset and a pure-colour cattle face dataset were used; accuracies of 98.58%, 97.91% and 99.47% were obtained, respectively. Hence, the recognition accuracy of the proposed model is higher than those of the other models. The experimental results show that the proposed model can accurately identify individual sika deer, giving it great potential for application to sika deer breeding and the facial recognition of other animals with high facial similarity.

In the process of testing the facial recognition model of sika deer, it was found that the model has not achieved the best effect on facial recognition of sika deer in complex backgrounds. The method of improving the network structure can continue to be used to improve the recognition ability of the network in complex environments to adapt to large farms. complex and changing environment. At the same time, in the future, building a lightweight model and reducing the size of the model to improve the training speed and generalization of the model will be the focus of our work in the future without reducing the accuracy rate.

Acknowledgement: We are thankful to the Dongao Deer Industry Group Farm for their contribution.

Funding Statement: This research was supported by the Science and Technology Department of Jilin Province [20210202128NC <http://kjt.jl.gov.cn>], The People's Republic of China Ministry of Science and Technology [2018YFF0213606-03 <http://www.most.gov.cn>], the Jilin Province Development and Reform Commission [2019C021 <http://jldrc.jl.gov.cn>] and the Science and Technology Bureau of Changchun City [21ZGN27 <http://kjj.changchun.gov.cn>].

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] D. R. McCullough, Z. G. Jiang and C. W. Li, "Sika deer in mainland China," in *Sika Deer*, Springer, Tokyo, pp. 521–539, 2009.
- [2] W. Shen, H. Hu, B. Dai, X. Wei, J. Sun *et al.*, "Individual identification of dairy cows based on convolutional neural networks," *Multimedia Tools and Applications*, vol. 79, no. 21, pp. 14711–14724, 2020.
- [3] Y. Zhong, S. Qiu, X. Luo, Z. Meng and J. Liu, "Facial expression recognition based on optimized ResNet," in *2020 2nd World Symp. on Artificial Intelligence (WSAI)*, Guangzhou, China, IEEE, pp. 84–91, 2020.
- [4] M. Marsot, J. Mei, X. Shan, L. Ye, P. Feng *et al.*, "An adaptive pig face recognition approach using convolutional neural networks," *Computers and Electronics in Agriculture*, vol. 173, pp. 105386, 2020.
- [5] M. F. Hansen, M. L. Smith, L. N. Smith, M. G. Salter, E. M. Baxter *et al.*, "Towards on-farm pig face recognition using convolutional neural networks," *Computers in Industry*, vol. 98, pp. 145–152, 2018.
- [6] L. Yao, Z. Hu, C. Liu, H. Liu, Y. Kuang *et al.*, "Cow face detection and recognition based on automatic feature extraction algorithm," in *Proc. of the ACM Turing Celebration Conf.*, Chengdu, China, pp. 1–5, 2019.
- [7] Y. Otani and H. Ogawa, "Potency of individual identification of Japanese macaques (*Macaca fuscata*) using a face recognition system and a limited number of learning images," *Mammal Study*, vol. 46, no. 1, pp. 1–9, 2021.
- [8] J. Hou, Y. He, H. Yang, T. Connor, J. Gao *et al.*, "Identification of animal individuals using deep learning: A case study of giant panda," *Biological Conservation*, vol. 242, pp. 108414, 2020.
- [9] S. Guo, P. Xu, Q. Miao, G. Shao, C. A. Chapman *et al.*, "Automatic identification of individual primates with deep learning techniques," *Iscience*, vol. 23, no. 8, pp. 101412, 2020.
- [10] H. Yan, Q. Cui and Z. Liu, "Pig face identification based on improved AlexNet model," *Inmateh-Agricultural Engineering*, vol. 61, no. 2, pp. 97–104, 2020.
- [11] S. Chen, S. Wang, X. Zuo and R. Yang, "Angus cattle recognition using deep learning," in *2020 25th Int. Conf. on Pattern Recognition (ICPR)*, Milan, Italy, IEEE, pp. 4169–4175, 2021.
- [12] S. A. Jwade, A. Guzzomi and A. Mian, "On farm automatic sheep breed classification using deep learning," *Computers and Electronics in Agriculture*, vol. 167, pp. 105055, 2019.
- [13] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Boston, USA, pp. 770–778, 2016.
- [14] K. Deeba and B. Amutha, "ResNet-Deep neural network architecture for leaf disease classification," *Microprocessors and Microsystems*, vol. 10, pp. 103364, 2020.
- [15] V. Kumar, H. Arora and J. Sisodia, "ResNet-Based approach for detection and classification of plant leaf diseases," in *2020 Int. Conf. on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, IEEE, pp. 495–502, 2020.
- [16] M. Gour, S. Jain and T. Sunil Kumar, "Residual learning based CNN for breast cancer histopathological image classification," *International Journal of Imaging Systems and Technology*, vol. 30, no. 3, pp. 621–635, 2020.
- [17] D. Sarwinda, R. H. Paradisa, A. Bustamam and P. Anggia, "Deep learning in image classification using residual network (ResNet) variants for detection of colorectal cancer," *Procedia Computer Science*, vol. 179, pp. 423–431, 2021.

- [18] L. H. Shehab, O. M. Fahmy, S. M. Gasser and M. S. El-Mahallawy, "An efficient brain tumor image segmentation based on deep residual networks (ResNets)," *Journal of King Saud University-Engineering Sciences*, vol. 33, no. 6, pp. 404–412, 2021.
- [19] R. Rajakumari and L. Kalaivani, "Breast cancer detection and classification using deep cnn techniques," *Intelligent Automation & Soft Computing*, vol. 32, no. 2, pp. 1089–1107, 2022.
- [20] M. Wang, X. Zhang, X. Niu, F. Wang and X. Zhang, "Scene classification of high-resolution remotely sensed image based on ResNet," *Journal of Geovisualization and Spatial Analysis*, vol. 3, no. 2, pp. 1–9, 2019.
- [21] G. Li, L. Li, H. Zhu, X. Liu and L. Jiao, "Adaptive multiscale deep fusion residual network for remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 11, pp. 8506–8521, 2019.
- [22] Y. Cheng and H. Wang, "A modified contrastive loss method for face recognition," *Pattern Recognition Letters*, vol. 125, pp. 785–790, 2019.
- [23] Q. M. Ilyas and M. Ahmad, "An enhanced deep learning model for automatic face mask detection," *Intelligent Automation & Soft Computing*, vol. 31, no. 1, pp. 241–254, 2022.
- [24] A. F. Agarap, "Deep learning using rectified linear units (relu)," *arXiv Preprint arXiv*, vol. 1803, pp. 08375, 2018.
- [25] J. Liang, "Image classification based on RESNET," *Journal of Physics: Conference Series*, IOP Publishing, vol. 1634, no. 1, pp. 012110, 2020.
- [26] G. Abosamra and H. Oqaibi, "An optimized deep residual network with a depth concatenated block for handwritten characters classification," *Computers, Materials & Continua*, vol. 68, no. 1, pp. 1–28, 2021.
- [27] J. Hu, L. Shen and G. Sun, "Squeeze-and-excitation networks," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Hawaii, USA, pp. 7132–7141, 2018.
- [28] I. C. Duta, L. Liu, F. Zhu and L. Shao, "Improved residual networks for image and video recognition," in *2020 25th Int. Conf. on Pattern Recognition (ICPR)*, IEEE, Milan, Italy, pp. 9415–9422, 2021.
- [29] D. A. Clevert, T. Unterthiner and S. Hochreiter, "Fast and accurate deep network learning by exponential linear units (elus)," *arXiv Preprint arXiv*, vol. 1511, pp. 07289, 2015.
- [30] Y. Sun, Y. Mu, Q. Feng, T. Hu, H. Gong *et al.*, "Deer body adaptive threshold segmentation algorithm based on color space," *Computers, Materials & Continua*, vol. 64, no. 2, pp. 1317–1328, 2020.