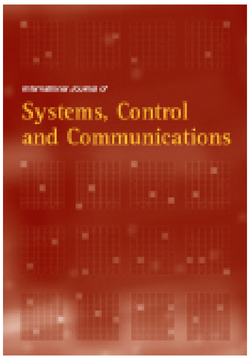


International Journal of Systems, Control and Communications

ISSN online: 1755-9359 - ISSN print: 1755-9340
https://www.inderscience.com/ijscc

Automatic insect identification system based on SE-ResNeXt

Yao Xiao, Aocheng Zhou, Lin Zhou, Yue Zhao

DOI: 10.1504/IJSCC.2022.10049960

Article History:

Received:	06 May 2022
Accepted:	07 July 2022
Published online:	06 December 2022

Automatic insect identification system based on SE-ResNeXt

Yao Xiao, Aocheng Zhou, Lin Zhou and Yue Zhao*

The School of Technology,
Beijing Forestry University,
Beijing 100083, China
Email: xy15944981269@outlook.com
Email: zhou_aocheng@bjfu.edu.cn
Email: Zhoulin1226@bjfu.edu.cn
Email: zhaoyue0609@126.com
*Corresponding author

Abstract: The Wudalianchi Scenic Area in Heilongjiang Province is the greatest place in the world to study species adaption and the evolution of biological communities. To solve the problems of heavy workload, poor timeliness, strong professionalism, and low accuracy in insect identification, an automatic insect identification system based on SE-ResNeXt is proposed. Firstly, to be suitable for the study of Wudalianchi insects, the dataset adopts the images of 105 species of eight orders insect in Wudalianchi. Then, through the comparison of three convolution neural networks, SE-ResNeXt has higher accuracy of insect identification than ResNet and Inception-V4, and its recall, precision, F1-score and accuracy all reach over 98%. Finally, based on Django framework, the website and app of system are built to realise the visualisation of identification results and the digital storage of insect data in Wudalianchi. The system has the characteristics of strong interactivity and convenient operation, and it was designed to provide technical assistance for insect protection, insect knowledge popularisation in agriculture and forestry, and a data foundation for the long-term evolution of insect variety in Wudalianchi, China.

Keywords: insect image; image identification; deep learning; convolutional neural network; CNN; Wudalianchi.

Reference to this paper should be made as follows: Xiao, Y., Zhou, A., Zhou, L. and Zhao, Y. (2023) 'Automatic insect identification system based on SE-ResNeXt', *Int. J. Systems, Control and Communications*, Vol. 14, No. 1, pp.81–98.

Biographical notes: Yao Xiao has been pursuing for her Bachelor's degree at the School of Technology, Beijing Forestry University, China, since 2019. Her major is Automation.

Aocheng Zhou has been pursuing for his Bachelor's degree at the School of Technology, Beijing Forestry University, China, since 2019. His major is Mechanical Engineering and Automation.

Lin Zhou has been pursuing for her Bachelor's degree at the School of Technology, Beijing Forestry University, China, since 2019. Her major is Electrical Engineering and Automation.

Yue Zhao received her BS in Measuring and Control Technology and Instrumentation from Harbin Engineering University in 2009 and PhD in Control Science and Engineering from Shanghai Jiao Tong University, Shanghai, China in 2014. She joined the Faculty of the Department of Automation, Beijing Forestry University in 2014, where she has been an Associate Professor, since 2018. Her research interests include deep learning, image processing, and data processing. In these areas, she has published dozens of technical articles. In addition, she is also a member of the IEEE Systems, Man, and Cybernetics (SMC) Society.

1 Introduction

Wudalianchi Scenic Area in Heilongjiang Province is a national protection zone with abundant animal and plant resources, and it is the best area in the world to study species adaptation and the evolution of biological communities. In recent years, due to the lack of protection for insects, the number of insects has been dropped sharply, and the insect ecosystem has been severely damaged, causing irreversible damage to the environment and ecological diversity. Therefore, automatic identification of insects in Wudalianchi Scenic Area can provide technical assistance for insect dataset construction, pest prevention and cure, environmental protection and border quarantine.

At present, the main method of insect identification is the visual method, which relies on the experience of workers, has problems such as heavy workload and low accuracy. Visual method requires the recogniser to have a high professional level of entomology, this kind of talent is in shortage. The existing insect dataset in Wudalianchi is mainly constructed in the form of specimens, which has poor timeliness, strong professionalism and difficult storage. These problems are not conducive to the statistics and preservation of various insect species in Wudalianchi. Hence, an efficient and accurate automatic insect recognition system is urgently needed.

With the development of image identification technology, automatic identification methods have been applied in the fields of environmental science (He et al., 2016) and agricultural engineering (Lowe, 2004), providing technical support for biological research (Weeks et al., 1999). At present, some traditional image identification methods, such as artificial neural network (ANN) (Russell et al., 2007), SVM algorithm (Arbuckle et al., 2001), DT algorithm (Wen and Guyer, 2012), and so on, have been successfully applied in the field of automatic insect identification. However, in the application of traditional methods, image features need to be extracted manually, which is not only very labour-intensive, but also poor in identification accuracy when there are too many insect categories and the feature similarity among insects is high.

Since AlexNet (Krizhevsky et al., 2012) put forward in 2012, the application of convolutional neural network (CNN) in image identification are increasing (Yang et al., 2019; Kamilaris and Prenafeta-Boldú, 2018; Ferentinos, 2018). CNNs, included VGG (Simonyan and Zisserman, 2014), GoogleNet (Szegedy et al., 2015), Inception-V4 (Szegedy et al., 2017), ResNet (He et al., 2016), SE-ResNeXt (Hu et al., 2018) and so on. Inception-V4 is the result of two modifications of GoogleNet, parallel structure and asymmetric convolution are used in its stem module, which can reduce the amount of calculation and improve the operation speed while ensuring the information loss is as low as possible. ResNet is proposed to solve the problem of network degradation in deep networks. On the basis of VGG19, ResNet adds residual unit through short-circuit mechanism, which can alleviate the problem of gradient disappearing during feedback and reduce the optimisation difficulty during network deepening. SE-ResNeXt adds squeeze-and-excitation (SE) module on the basis of ResNet network structure, which not only makes the network more nonlinear, but also can better fit the complex correlation between channels, and greatly reduces the amount of parameters and calculation. These CNNs have been successfully applied to face identification (Lu et al., 2018), character identification (Khan et al., 2021), medical image identification (Emara et al., 2019; Zhang et al., 2020; Al Husaini et al., 2022; Farooq and Hafeez, 2020; Linqi et al., 2022) and other fields, and have been proved that they have higher identification accuracy and faster identification speed than traditional methods (LeCun et al., 2015; Lauzon, 2012). Moreover, CNN takes the image itself as the input, which can automatically extract the features of the image, improve the defects of traditional methods and have higher efficiency of identification.

Hence, the study proposed the automatic insect identification system of Wudalianchi based on SE-ResNeXt. By training the networks, this system realises automatic and high-accuracy identification of 105 species of eight orders insects. The visualisation of automatic insect identification results and the digital storage of insect data in Wudalianchi are realised by building the website and app of system, which are based on Django framework. The system was built to provide technical support for insect protection and pest control on agricultural and forestry in Wudalianchi, and provide data basis for the sustainable development of insect diversity.

2 Related work

2.1 Inception-V4

Inception-V4 pre-processes the data before entering the Inception module by referring to the parallel structure and asymmetric convolution in the stem module, which can reduce the amount of network computation, improve the speed of network operation and keep the loss of information as small as possible. The three modules: inception-A, inception-B and inception-C are deeper and more complex convolution modules, which can improve the computing power of the network. The reduction-A and reduction-B modules in the structure are used to reduce the amount of calculation. At the same time, the dropout operation is introduced in front of the softmax layer to prevent the overfitting phenomenon in the training process.

2.2 ResNet

ResNet network is based on VGG19 network, the residual unit is introduced through the short-cut mechanism, so that it can realise residual learning and identity mapping, and then solve the problem of gradient disappearance caused by the increase of network training layer depth. There are two layers of residual blocks in ResNet, which can be expressed as:

$$F = W_2\sigma(W_1x) \tag{1}$$

$$y = F(x, \{W_i\}) + x \tag{2}$$

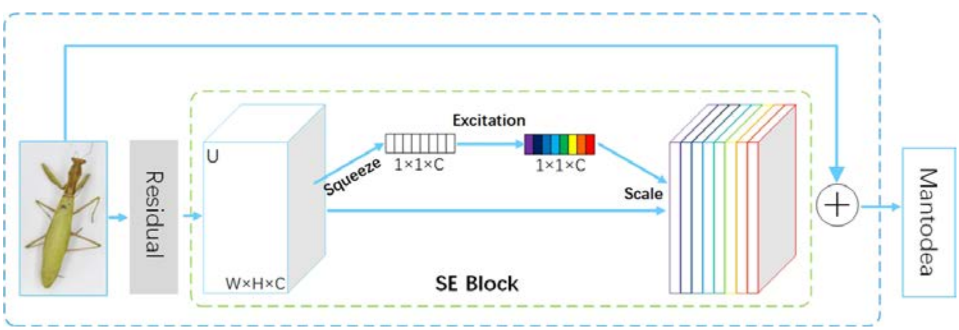
where F represents the residual function, σ represents the nonlinear function ReLU, x is the input of the residual unit, and W_1, W_2 are the convolution operation, $F(x, \{W_i\})$ is the residual mapping and y is the output of the residual unit.

ResNet network divides the output into two parts: $F(x) + x$ by short-cut mechanism: $y = x$ is the identity mapping; $F(x)$ is the residual, which is a complement to y. ResNet solves the problem of deep network degradation through residual unit, and can train deeper networks, which is a major breakthrough of deep learning networks.

2.3 SE-ResNeXt

Both SE-ResNeXt and ResNet have bottleneck characteristics, and shortcut is used to reduce the gradient dispersion when the network depth increases. The difference is that on the basis of ResNet network structure, SE-ResNeXt adds squeeze-and-excitation (SE) module, which can fit the complex correlation between channels better and reduce the amount of parameters and computation.

Figure 1 Chart of SE-ResNext (see online version for colours)



SE block is a computing unit, residual in Figure 1 is the traditional residual structure. After inputting insect images, they will be mapped and output as the feature maps $U(U \in R^{W\times H\times C})$, which are all existing in ResNet structure. The structure after the features U are the added part of SE block: first, squeeze the output $1 \times 1 \times C$ data; then perform excitation on the output $1 \times 1 \times C$ data; finally, limit it to the range of [0, 1] with Sigmoid, and its value is used as the scale, which is multiplied by channel C as the input data of the next level.

2.3.1 Squeeze

Each convolution kernel has a local receptive field, the output U cannot use the information outside the receptive field. In order to solve this problem, the squeeze structure adopts the simplest method of averaging, that is, global pooling [formula (3)].

$$z_c = F_{sq}(u_c) = \frac{1}{W \times H} \sum_{i=1}^{W} \sum_{j=1}^{H} u_c(i,j) \tag{3}$$

Through squeeze operation, the feature can be compressed along the spatial dimension, and each two-dimensional feature channel can be converted into a real number, so that the output U has a global receptive field to some extent. The output dimension matches the number of input feature channels and represents the global distribution of responses on the feature channels. The layer close to the input can also obtain the global receptive field.

2.3.2 Excitation

In order to using the data aggregated in squeeze, it will be performed to the excitation operation again. To ensure that each channel can be emphasised, the excitation structure selects a flexible activation gating mechanism with Sigmoid.

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)) \tag{4}$$

where δ is derived from the function ReLU. In order to simplify the model, excitation uses the bottleneck consisting of two full connections (FC) to reduce the dimension, then uses the function, and finally increases the dimension and returns to the dimension of the transformation output U after the squeeze link transformation. The final output of the excitation link is obtained by the following formula:

$$\tilde{X}_c = F_{scale}(u_c, s_c) \tag{5}$$

3 Performance evaluation

This paper adopts the evaluation indexes based on confusion matrix, recall, precision, F1-score and accuracy to evaluate the performance of the three CNNs.

3.1 Confusion matrix

Confusion matrix is a matrix in the form of n×n columns, which is a standard format to express accuracy evaluation. The columns represent the order of insects predicted by the model, and the total number of columns is the total number of images predicted for this purpose; the rows represents the real purpose of the insect image, and the total number of rows represents the total number of actual images for that purpose. Confusion matrix can clearly show the probability of correct classification and misclassification of insect images, which is convenient to analyse the identification results of the network.

3.2 Recall

Recall refers to the original sample, it indicates how many insect images of a certain purpose are correctly predicted for that purpose. The higher the recall, the better the network's performance.

$$R = \frac{TP}{TP + FN} \tag{6}$$

where TP (true positive) is the number of correctly judged samples that are judged as positive. FN (false negative) is the number of judgement errors in the sample judged as negative.

3.3 Precision

Precision refers to the prediction results. It indicates how many insect images predicted for a certain purpose really belong to that purpose. The higher the precision, the better the network's performance.

$$P = \frac{TP}{TP + FP} \tag{7}$$

where FP (false positive) is the number of judgement errors in the samples judged as positive.

3.4 F1-score

F1-score is the harmonic average of recall and accuracy. The recall and accuracy influence each other, and ideally both are high. But in general, when the accuracy is high, the recall rate is low, vice versa, which contradicts each other. Under the condition that both the P (precision) and the R (recall) have high requirements and the same weight, F1-score is used to comprehensively measure P and R.

$$F1 = \frac{2 \cdot P \cdot R}{P + R} \tag{8}$$

When multiple networks are compared, the higher the F1-score, the better the network's performance.

3.5 Accuracy

Accuracy represents the ratio of correctly predicted insect images (TP + TN) to all insect images (TP + FP + TN + FN). Generally, the higher the accuracy of identification, the better the network's performance.

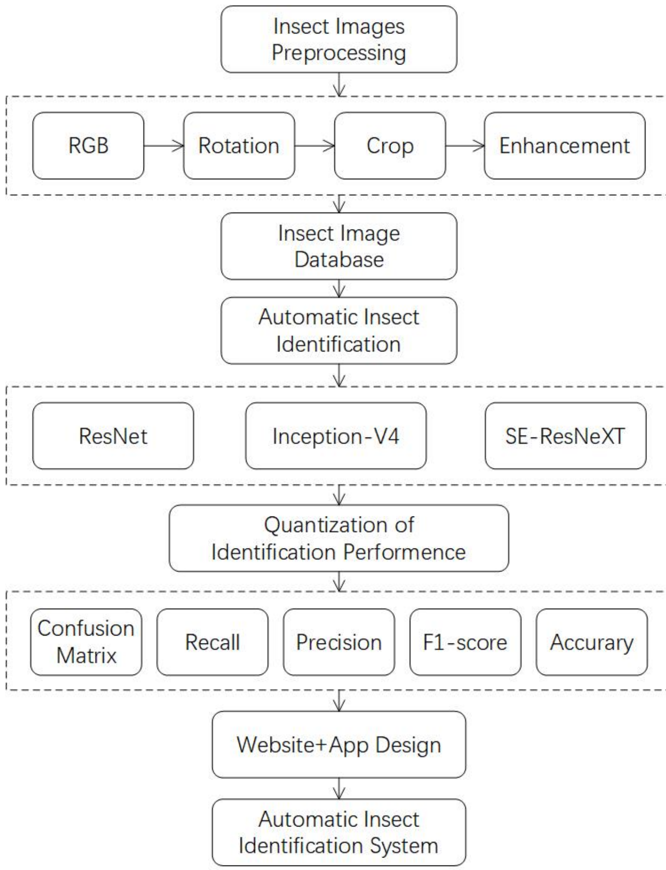
$$A = \frac{TP + TN}{TP + FP + TN + FN} \tag{9}$$

where TN (true negative) is the number of correctly judged samples that are judged negative.

4 Background

To solve the problems of heavy workload, poor timeliness and low accuracy of insect identification in existing methods, the study proposes an automatic insect identification system of Wudalianchi based on SE-ResNeXt (Figure 2).

Figure 2 System flowchart



The system consists of three parts:

1 Insect image acquisition and pre-processing: Firstly, obtain insect images with insect specimens, digital cameras, studios and other equipment. Then, construct the dataset of insects by cropping, rotating and enhancing the insect images.

2 Automatic insect image identification: SE-ResNeXt, ResNet and Inception-V4 are trained to obtain insect identification models. The best model for insect image identification is determined through quantitative analysis of their identification accuracy.

3 Automatic insect identification system construction: The visualisation and digitisation of insect identification is realised by building the website and app of

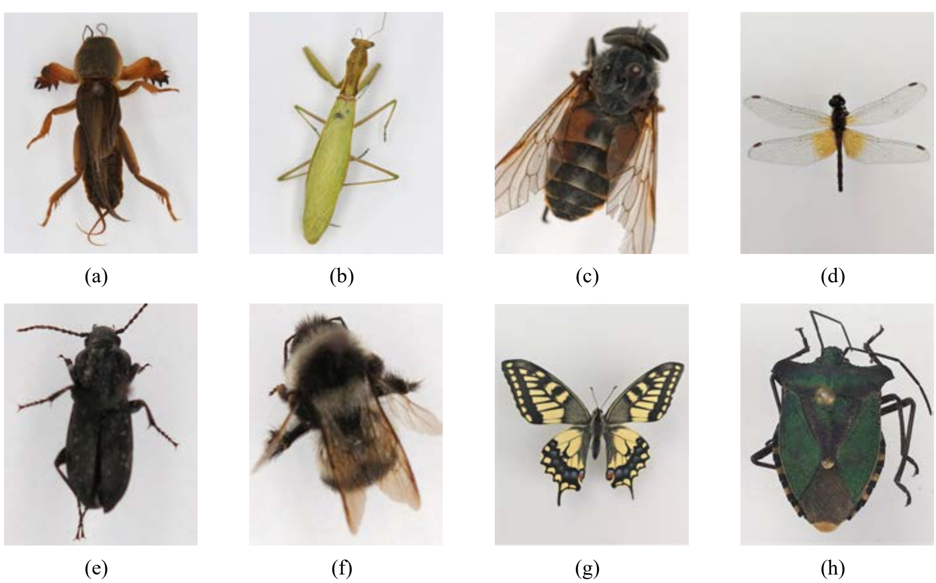
system, which are based on Django framework and combined with the best insect identification model.

4.1 Insect image collection

In this paper, we use 105 species of eight orders insects as experimental materials, which can be commonly found in Wudalianchi Scenic Area of Heilongjiang Province. Use professional equipment such as digital camera, background plate, reflector and spotlight to shoot insect specimens. Illumination was provided by a 150 W light source. During the experiment, 739 insect specimens were placed horizontally with their heads upward in the centre of the white background plate and shot from the top view.

There are 739 images of insects, including eight orders of insects: Orthoptera, Mantodea, Diptera, Odonata, Coleoptera, Hymenoptera, Lepidoptera and Hemiptera (Figure 3). The number of Orthoptera images is 82, Mantodea images is 9, Diptera images is 12, Odonata images is 233, Coleoptera images is 133, Hymenoptera images is 39, Lepidoptera images is 207 and Hemiptera images is 21.

Figure 3 Eight orders of insects, (a) Orthoptera (b) Mantodea (c) Diptera (d) Odonata (e) Coleoptera (f) Hymenoptera (g) Lepidoptera (h) Hemiptera (see online version for colours)



4.2 Insect image pre-processing

The original insect images have some defects, such as different target sizes, large background area, low contrast between target and background, and limited number of insect specimens, which affect the efficiency and accuracy of recognition. To improve the performance of the insect identification, we pre-process the images: firstly, rotate the image randomly; secondly, crop the image according to the largest circumscribed rectangle; finally, improve the contrast and brightness of the images.

4.2.1 Rotation

In order to enlarge the insect image dataset, the images are randomly rotated. The images before and after rotation are both saved in the dataset [Figure 4(b)]. By being trained from insect images with different rotation angles, the models can eliminate the random error in practical application and improve the identification accuracy.

4.2.2 Crop

The original insect images contain a large background area [Figure 4(a)]. In order to improve the efficiency of identification, we crop the insect images based on the maximum circumscribed rectangle method, retain the rectangular area of the insect target [Figure 4(c)], and store them in JPG.

4.2.3 Enhancement

The original image is dark in brightness, and the contrast between target and background is low. By improving the brightness and contrast of the image, enhance the target area in the insect images and highlight the colour and texture information of the insect [Figure 4(d)]. The calculation equation is:

$$g(i, j) = \alpha \cdot f(i, j) + \beta \tag{10}$$

where i and j represent the pixels in the I^{th} row and the J^{th} column, the original image and the output image. α and β are gain parameter and deviation parameter, which control contrast and brightness respectively. In this paper, α values 1.5 and β values 30.

Figure 4 Insect image pre-processing, (a) original (b) rotation (c) cropping (d) enhancement (see online version for colours)

(a) (b) (c) (d)

4.3 CNNs training

In this paper, we use the one-hot encoding to label eight orders of insects. For example, if we take Odonata as the first category, its corresponding one-hot encoding is [1, 0, 0, 0, 0, 0, 0, 0]; if we take Lepidoptera as the forth category, its corresponding one-hot encoding is [0, 0, 0, 1, 0, 0, 0, 0]. Four different layers of ResNet (18, 34, 50, 101) and two different layers of SE-ResNeXt (50, 101) are trained to evaluate the identification performance by using recall, precision, F1-value and accuracy.

As shown in Table 1, among the four different layers of ResNet, it can be seen that for the insect image database used in this paper, ResNet101 has the lowest four evaluation indexes and the largest standard deviation. ResNet18 has the highest four evaluation indexes and lower standard deviation, that is, ResNet18 has the best identification performance.

As shown in Table 2, among the two different layers of SE-ResNeXt, it can be seen that for the insect image database used in this paper, the four evaluation indexes of SE-ResNeXt50 are higher than those of SE-ResNeXt101 and have lower standard deviation, that is, SE-ResNeXt50 has the better identification performance.

As the lower the number of network layers, the faster the identification speed. In summary, this paper finally adopts three CNNs, namely, ResNet18, SE-ResNeXt50 and Inception-V4.

Table 1 Four evaluation indexes of ResNet with four different layers

Network	*Recall*	*Precision*	*F1-score*	*Accuracy*
ResNet101	87.34 ± 13.54%	87.34 ± 7.24%	87.85 ± 7.68%	87.34%
ResNet50	95.17 ± 3.87%	95.17 ± 5.62%	95.14 ± 4.42%	95.16%
ResNet34	93.04 ± 6.01%	93.04 ± 5.10%	93.13 ± 3.00%	93.07%
ResNet18	96.93 ± 4.38%	96.93 ± 4.38%	96.91 ± 2.25%	96.93%

Table 2 Four evaluation indexes of SE-ResNeXt with two different layers

Network	*Recall*	*Precision*	*F1-score*	*Accuracy*
ResNet	97.24 ± 2.55%	97.24 ± 5.10%	97.20 ± 2.71%	97.24%
SE-ResNeXt	98.39 ± 2.78%	98.39 ± 2.78%	98.38 ± 1.23%	98.39%

5 Results and analysis

Because the number of images of each insect target is different, for the target with a large number of images, the captured images are less; for the target with a small number of images, images will be taken repeatedly. Finally the average number of images for each order is between 70–120. The insect image dataset is divided into two groups, 80% is used for training, 20% for testing. Inception-V4, ResNet and SE-ResNeXt, are trained to compare and analyse their accuracy of insect identification. Inception-V4 adopts Rms optimiser; ResNet adopts 0.01 learning rate and 18 training layers; SE-ResNeXt adopts 0.01 learning rate and 50 training layers. Using Intel CORE i7 as the processor, Anaconda as the compilation environment, Pycharm as the development tool, Python as the development language.

5.1 System identification accuracy and result analysis

5.1.1 Confusion matrix

As shown in Figure 7(a), Inception-V4 has the highest accuracy in identifying Diptera, Hymenoptera and Mantodea, which are 100%, 100% and 99% respectively; the second is the identification accuracy of Hemiptera, Lepidoptera and Odonata, which are 95%, 91% and 89% respectively; for Coleoptera and Orthoptera, the accuracy are the lowest of 77% and 74%, respectively. In the confusion matrix, we can find an obvious defect: the probability of network misidentifying Coleoptera as Hemiptera is 20%. The reasons for this phenomenon are: as shown in Figure 5, the two are similar in shape, both oval or nearly cylindrical; their colour are close to black; the material of their body wall are similar, rough and horny.

Figure 5 Inception-V4 misidentification case, (a) Coleoptera (b) Hemiptera (see online version for colours)

(a) (b)

As shown in Figure 7(b), the identification accuracy of ResNet for Orthoptera, Mantodea, Diptera and Hemiptera are all reach 100%; the accuracy of Hymenoptera and Lepidoptera are the second, both are 97%; the accuracy of Odonata and Coleoptera are the lowest, which are 93% and 87% respectively. From the confusion matrix, it can be seen that ResNet has the probability to incorrectly identify Odonata, Coleoptera and Hymenoptera as Orthoptera. The reasons for this phenomenon are: As shown in Figure 6, the two pairs of wings of Odonata is transparent, and the network may ignore the wings and only extract the body part of dragonfly during feature extraction; some Coleoptera are slender; when the wings of Diptera contract, the network judges the wings as part of the insect's body. Odonata, Coleoptera and Hymenoptera are similar to Orthoptera: slender, wingless and dark, which leads to the wrong recognition of ResNet.

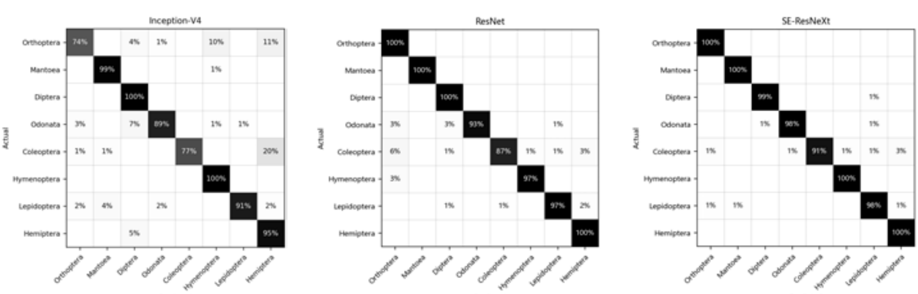
As shown in Figure 7(c), the identification accuracy of SE-ResNeXt for Orthoptera, Mantodea, Hymenoptera and Hemiptera reaches 100%; for Diptera, Odonata and Lepidoptera, the accuracy are 99%, 98% and 98% respectively; the accuracy of Coleoptera is the worst, which is 91%. The reason for the low identification accuracy of Coleoptera by SE-ResNeXt is that the insect image dataset used in this paper is small, which is not enough to cover all subordinate species of Coleoptera, so the network can't accurately learn the shape features of Coleoptera during training. Therefore, when the appearance of other insects are similar to that of Coleoptera, it is possible for SE-ResNeXt to misjudge other insects as Coleoptera.

Figure 6 ResNet misidentification case, (a) Odonata (b) Coleoptera (c) Hymenoptera (d) Orthoptera (see online version for colours)

(a) (b) (c) (d)

To sum up, the three networks all have the highest identification accuracy for Mantodea and Diptera (all above 99%), while all having the lowest identification accuracy for Coleoptera. The reason is that the appearances of different species of Mantodea [Figure 3(b)] are universal, and are more special than other orders: the shape is slender, the head is triangular, the front leg joints are sickle-shaped, and the colours are mostly green. Therefore, the probability that the model mistakenly identifies Mantodea as other orders is very low. There are two kinds of common Diptera insects, mosquitoes and flies. With obvious morphological features and little difference in appearance between different species, it is very unlikely that the three networks misidentifying Mantis and Diptera as other orders.

Figure 7 Confusion matrices of three networks, (a) Inception-V4 (b) ResNet (c) SE-ResNeXt



(a) (b) (c)

The three networks have the highest identification accuracy for Coleoptera. The reason is that Coleoptera is the most diverse and widely distributed group of the Insecta. Its various affiliated species have different shapes and colours, and there is no universal regularity. Moreover, some Coleoptera have similarities in appearance to other orders, which leads to the wrong identification. For the insect image dataset used in this paper, among the three networks, SE-ResNeXt has the lowest probability of false identification, which shows that SE-ResNeXt can accurately extract and learn the features of each insect even with limited dataset, and can also realise high accuracy identification.

5.1.2 Recall

As shown in Table 1, among the three networks, the recall of SE-ResNeXt network is the highest, which is 98.39%, which means that when there are 100 images of insects of one order, SE-ResNeXt can correctly identify 98 images of this order. The recall of SE-ResNeXt network is 1.48% higher than that of the second-best network ResNet, and the recall of Inception-V4 network is the worst, which is 91.16%. Furthermore, SE-ResNeXt network has the smallest standard deviation, followed by ResNet, and the fluctuation range of Inception-V4 is the largest.

5.1.3 Precision

Among the three networks, the precision of SE-ResNeXt network is the highest, which is 98.39%, which means that when 100 insect images are identified as an order by SE-ResNeXt model, 98 of them really belong to the order. The precision of SE-ResNeXt network is 1.48% higher than that of the second-best network ResNet, and the precision of Inception-V4 network is the worst, which is 91.68%. Overall, SE-ResNeXt has the smallest standard deviation, followed by ResNet, and the fluctuation range of Inception-V4 is the largest.

5.1.4 F1-score

The F1-score of SE-ResNeXt network is the highest among the three networks, which is 98.38%. The F1-score of SE-ResNeXt network is 1.47% higher than that of the second-best network ResNet, and the F1-score of Inception-V4 network is the worst, which is 91.07%. For the F1-score of the three networks, SE-ResNeXt has the smallest standard deviation, followed by ResNet, and the fluctuation range of Incepion-V4 is the largest.

5.1.5 Accuracy

Among the three models, SE-ResNeXt model has the highest accuracy of 98.39%, which means that 98 out of every 100 insect images are correctly identified by SE-ResNeXt network. The accuracy of SE-ResNeXt network is 1.46% higher than that of the second-best network ResNet, and the accuracy of Inception-V4 model is the worst, which is 91.15%. The accuracy is numerically equal to the average of recall, precision and F1-score, so there is no standard deviation of accuracy.

To sum up, results indicate the five indexes of SE-ResNeXt are the best, that is, its performance is the best among the three networks. This is mainly because the correct identification of insect images requires a deep network depth, and the residual unit can alleviate the gradient disappearance caused by increasing the depth in the CNN. SE-ResNeXt can improve the identification accuracy while deepening the network depth. And SE module reduces the amount of parameters and calculation, and also improves the identification effect to a certain extent. Therefore, SE-ResNeXt can obtain higher recognition accuracy when recognising insect images. The results show that SE-ResNeXt is more suitable for the field of insect image identification, which can improve the accuracy without increasing the complexity of parameters and reduce the number of super-parameters. It not only reduces the difficulty of network design and the complexity of time and space, but also has higher identification accuracy and shorter identification time.

Table 3 Four evaluation indexes of three networks

Network	*Recall*	*Precision*	*F1-score*	*Accuracy*
Inception-V4	91.16 ± 6.70%	91.68 ± 9.60%	91.07 ± 4.8%	91.15%
ResNet	96.93 ± 4.38%	96.93 ± 4.38%	96.91 ± 2.25%	96.93%
SE-ResNeXt	98.39 ± 2.78%	98.39 ± 2.78%	98.38 ± 1.23%	98.39%

5.2 Interface of automatic insect identification system

In this paper, based on SE-ResNeXt network and Django framework, with Pycharm as the development tool and python as the compilation language, the website and the app of the automatic insect identification system are designed. The combination of website and app can realise automatic identification of insect images in different scenes, popularisation of insect knowledge and digital storage of insect data in Wudalianchi. The system includes automatic insect identification module, insect knowledge popularisation module and feedback module. Results of multiple tests show that the network takes only 70ms to identify an insect image. Taking into account the performance of the equipment and the network delay, our insect automatic identification system developed in this paper can complete the identification and display the result within 1s on both PC and app.

5.2.1 Website

In this paper, the website of automatic insect identification system includes three functional modules and four display windows, which are as follows:

1 Automatic insect identification function: This function includes two display interfaces [Figures 8(a) and 8(b)]. First, the user can click the 'Choose File' button to arbitrarily select the insect image to be identified from the insect image datasets in different paths; secondly, click the 'Upload' button, the page will jump to the result page [Figure 8(b)]. In this process, the system uploads the picture to the static folder on the server. After successfully upload, the system will start two threads: obtaining image attributes and identifying insect.

2 The first thread: The system obtains information of the uploaded insect image such as name, size and upload time. To avoid repeat, convert the name of the image to MD5 value for storage. The second thread: input the insect image into the insect identification model to get the insect identification result, and display the insect representative picture and brief introduction in Figure 8(b). Finally, save the MD5 value, size, identification time, identification state and identification result of images into the database. In this paper, multi-thread technology is used in the system, which greatly improves the efficiency of automatic insect identification and enhances the user experience.

3 Insect science popularisation function: This function includes two display interfaces. Firstly, every time users refresh the homepage, the random function will randomly list the profiles of six insects in the encyclopaedia module [Figure 8(a)]. Secondly, when users click on the insect name or insect picture, the page will jump to the Baidu Encyclopedia page of this insect. This functional module can help users deepen their understanding of insect knowledge.

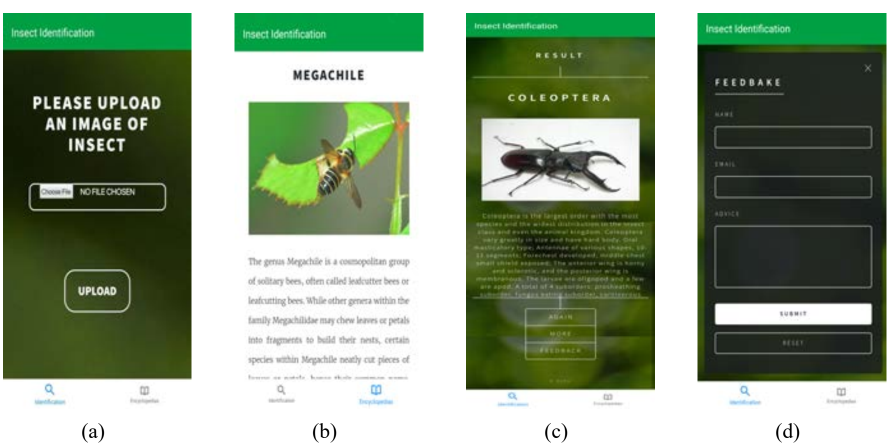
4 Feedback function: Due to the limitation of the identification accuracy of CNN, errors may occur during identification. Therefore, 'Feedback' is set to facilitate users to feedback relevant information [Figure 8(c)].

Figure 8 Pages of website, (a) homepage (b) result page (c) feedback (see online version for colours)

5.2.2 *App*

The app was transformed from the automatic insect identification website with Fushion app as the development tool, the app and the website of insect identification system have the same functions, including three functional modules and four display windows.

Figure 9 Pages of app, (a) identification (b) encyclopaedia (c) result (d) feedback (see online version for colours)



As shown in Figures 9(a) and 9(b), this app includes insect image selection and insect knowledge introduction. The automatic insect recognition function can be realised through key operation. Figure 9(c) shows the insect identification results retrieved from the Identification page, which can provide users with detailed insect knowledge. Users can also click the 'More' button to jump to the encyclopaedia interface to deepen their understanding of insects. Figure 9(d) is the feedback interface of the system, which can collect the feedback information from users in real time, is conducive to improving the experience of using the automatic insect identification app. The app is integrated in the smart phone, which has the feature of portable use, and can be applied to field investigation in forest areas and other places.

6 Summary

In this paper, 105 species of eight orders insects in Wudalianchi Scenic Area of Heilongjiang Province are taken as the research object. To solve the problems of heavy workload, poor timeliness, strong professionalism and low accuracy in the existing insect identification, this paper proposes an automatic insect identification system of Wudalianchi based on SE-ResNeXt, and draws the following conclusions:

1 According to the confusion matrix of SE-ResNeXt, ResNet and Inception-V4, it can be known that all three networks have the high identification accuracy (all above 99%) for Mantodea, while having the lowest identification accuracy for Coleoptera (77%, 87%, 91%, respectively).

2 Through the quantitative analysis of four indexes of SE-ResNeXt, ResNet and Inception-V4, it can be seen that the performance of identification of SE-ResNeXt is the best, ResNet is the second, Inception-V4 is the worst. The recall, precision, F1-score and accuracy of SE-ResNeXt network are all above 98%, which are 1.46%, 1.46%, 1.47% and 1.46% higher than that of the second-best ResNet respectively.

3 In this system, the website and the app of system can realise automatic identification of insect images in different scenes, popularisation of insect knowledge and digital storage of insect data. In addition, the website is suitable for entomological education and popularisation of science, the app is suitable for investigation in forest areas.

Acknowledgements

The project is supported by the Innovation and entrepreneurship training program for College Students, under Grant No. X202110022140.

References

Al Husaini, M.A.S., Habaebi, M.H., Gunawan, T.S., Islam, M.R., Elsheikh, E.A. and Suliman, F.M. (2022) 'Thermal-based early breast cancer detection using inception V3, inception V4 and modified inception MV4', *Neural Computing and Applications*, Vol. 34, No. 1, pp.333–348.

Arbuckle, T., Schröder, S., Steinhage, V. and Wittmann, D. (2002) 'Biodiversity informatics in action: identification and monitoring of bee species using ABIS', in *Proc. 15th Int. Symp. Informatics for Environmental Protection*, October, Vol. 1, pp.425–430, Google Scholar.

Emara, T., Afify, H.M., Ismail, F.H. and Hassanien, A.E. (2019) 'A modified inception-v4 for imbalanced skin cancer classification dataset', *2019 14th International Conference on Computer Engineering and Systems (ICCES)*, December, pp.28–33.

Farooq, M. and Hafeez, A. (2020) *COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs*, arXiv preprint arXiv:2003.14395.

Ferentinos, K.P. (2018) 'Deep learning models for plant disease detection and diagnosis', *Computers and Electronics in Agriculture*, Vol. 145, pp.311–318, ISSN: 0168-1699.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) 'Deep residual learning for image identification', *Proceedings of the IEEE Conference on Computer Vision and Pattern Identification*, pp.770–778.

Hu, J., Shen, L. and Sun, G. (2018) 'Squeeze-and-excitation networks', *Proceedings of the IEEE Conference on Computer Vision and Pattern Identification*, pp.7132–7141.

Kamilaris, A. and Prenafeta-Boldú, F.X. (2018) 'Deep learning in agriculture: a survey', *Computers and Electronics in Agriculture*, Vol. 147, pp.70–90, ISSN: 0168-1699.

Khan, M.M., Uddin, M.S., Parvez, M.Z. and Nahar, L. (2022) 'A squeeze and excitation ResNeXt-based deep learning model for Bangla handwritten compound character recognition', *Journal of King Saud University – Computer and Information Sciences*, Vol. 34, pp.3356–3364, ISSN: 1319-1578.

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) 'ImageNet classification with deep convolutional neural networks', *Commun. ACM*, Vol. 60, No. 6, pp.84–90, ISSN: 0001-0782.

Lauzon, F.Q. (2012) 'An introduction to deep learning', *2012 11th International Conference on Information Science, Signal Processing and their Applications (ISSPA)*, July, IEEE, pp.1438–1439.

LeCun, Y., Bengio, Y. and Hinton, G. (2015) 'Deep learning', *Nature*, Vol. 521, No. 7553, pp.436–444.

Linqi, J., Chunyu, N. and Jingyang, L. (2022) 'Glioma classification framework based on SE-ResNeXt network and its optimization', *LIET Image Processing*, Vol. 16, No. 2, pp.596–605.

Lowe, D.G. (2004) 'Distinctive image features from scale-invariant keypoints', *International Journal of Computer Vision*, Vol. 60, No. 2, pp.91–110.

Lu, Z., Jiang, X. and Kot, A. (2018) 'Deep coupled resnet for low-resolution face indentification', *IEEE Signal Processing Letters*, Vol. 25, No. 4, pp.526–530.

Russell, K.N., Do, M.T., Huff, J.C., Platnick, N.I. and MacLeod, N. (2007) 'Introducing SPIDA-web: wavelets, neural networks and Internet accessibility in an image-based automated indentification system', *Systematics Association Special Volume*, Vol. 74, pp.131–152.

Simonyan, K. and Zisserman, A. (2014) *Very Deep Convolutional Networks for Large-Scale Image Identification*, arXiv preprint arXiv:1409.1556.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. and Rabinovich, A. (2015) 'Going deeper with convolutions', *Proceedings of the IEEE Conference on Computer Vision and Pattern Identification*, pp.1–9.

Szegedy, C., Ioffe, S., Vanhoucke, V. and Alemi, A.A. (2017) ' Inception-v4, inception-resnet and the impact of residual connections on learning', *Thirty-First AAAI Conference on Artificial Intelligence*, February.

Weeks, P.J.D., O'Neill, M.A., Gaston, K.J. and Gauld, I.D. (1999) 'Automating insect indentification: exploring the limitations of a prototype system', *Journal of Applied Entomology*, Vol. 123, No. 1, pp.1–8.

Wen, C. and Guyer, D. (2012) 'Image-based orchard insect automated indentification and classification method', *Computers and Electronics in Agriculture*, Vol. 89, pp.110–115, ISSN: 0168-1699.

Yang, H., Gao, L., Tang, N. and Yang, P. (2019) 'Experimental analysis and evaluation of wide residual networks based agricultural disease indentification in smart agriculture system', *EURASIP Journal on Wireless Communications and Networking*, Vol. 2019, No. 1, pp.1–10.

Zhang, G., Yang, Z., Gong, L., Jiang, S., Wang, L. and Zhang, H. (2020) 'Classification of lung nodules based on CT images using squeeze-and-excitation network and aggregated residual transformations', *La Radiologia Medica*, Vol. 125, No. 4, pp.374–383.