

# An Automatic Insect Recognition Algorithm in Complex Background Based on Convolutional Neural Network



Xianrong Zhang<sup>1,2\*</sup>, Gang Chen<sup>1</sup>

<sup>1</sup>Zhejiang University, Hangzhou 310058, China

<sup>2</sup>Zhijiang College of Zhejiang University of Technology, Shaoxing 312030, China

Corresponding Author Email: [cg@zju.edu.cn](mailto:cg@zju.edu.cn)

<https://doi.org/10.18280/ts.370511>

## ABSTRACT

**Received:** 10 June 2020

**Accepted:** 25 September 2020

### Keywords:

*convolutional neural network (CNN), edgeless active contour, insect image recognition, complex background, narrow-band fast method*

The existing insect recognition methods mostly segment the target region by traditional classification technology, failing to achieve a high accuracy in complex background. To solve the problem, this paper introduces the morphology-based edgeless active contour strategy to segment insects in complex background. The strategy integrates the morphological operation of gray image, and detects insect contours by narrow-band fast method. To enhance the background diversity of new samples, the authors improved the synthetic minority over-sampling technique (SMOTE) algorithm into a variable weight edge enhancement algorithm. Based on the SMOTE algorithm, the proposed algorithm increases the weight of the edge area as adjacent images are superimposed into a new image, making the background of the new image more complex. Finally, the proposed method was coupled with DenseNet-121 to recognize insects in images with complex background. The results show that the accuracy of the network was nearly 10% higher on the balanced set than on the unbalanced set, suggesting that our method is feasible and accurate.

## 1. INTRODUCTION

The traditional insect recognition methods mainly rely on human vision: the insect in the target image is compared with known insect species in texture, color, and shape to determine its category. This subjective strategy inevitably leads to human errors and misjudgments. The recognition effect solely depends on the expertise and experience of the researcher.

With the advent of artificial intelligence (AI), great progress has been made in the deep learning (DL) technology for speech recognition, natural language processing, and computer vision. Considering its popularity in image recognition and face detection [1], this paper aims to utilize the DL to improve the accuracy of insect recognition in complex background.

In fact, most of the current image recognition algorithms are based on the DL. Their recognition results largely depend on the features of the dataset. However, the image set of insects is often highly unbalanced, owing to the varied difficulties in shooting images on different insect species. Thus, the image set must be balanced before the insect recognition.

Image segmentation, a key step of insect recognition, attempts to segment the region of interest (ROI) from the whole image. The segmentation quality directly affects the subsequent operations like feature extraction and image classification. Hence, the accuracy and stability of the segmentation algorithm are the prerequisite to effective insect recognition.

The complex background in many images adds to the difficulty in image segmentation. In many cases, the natural environment is complex and susceptible to the influence of various interference factors (e.g. sunlight and climate). Before insect recognition, it is important to eliminate the complex

background and green leaves from the original image.

This paper mainly puts forward a novel edge-based image segmentation model. Firstly, a morphology-based edgeless active contour algorithm was designed to detect insect contours by narrow-band fast method. Then, a variable weight edge enhancement algorithm was extended from the synthetic minority over-sampling technique (SMOTE) algorithm to enhance the diversity of image background. Through contrastive experiments on eight convolutional neural networks (CNNs), it is proved that the proposed method can achieve the best recognition effect on the DenseNet-121.

## 2. LITERATURE REVIEW

Since the 1980s, computer vision has been applied in insect recognition. Due to the limited computing speed, it is impossible to realize real-time recognition. Therefore, this technology was merely used to recognize insect images in specific environment. For instance, Wen and Guyer [2] binarized the collected insect images into black and white images, and recognized the insects by the morphological features.

With the advancement of computer technology, emerging technologies like digital image technology, support vector machine (SVM), and neural network (NN) have been adopted for insect recognition. Manickavasagam et al. [3] detected parasites on crops with SVM classifier. Zhang et al. [4] proposed an insect detection and recognition system, which captures video data with a camera and detects insects on crops through video analysis. Malvade et al. [5] extracted texture features in hue-saturation-intensity (HSI) space by color co-occurrence method, and effectively identified objects in

complex background.

Thanks to AI development, the DL has gradually replaced traditional image processing in target recognition. Hinton et al. [6] presented the unsupervised method for weight initialization in artificial neural network (ANN), and finetuned the initial weights through supervised training, thereby solving the problem of vanishing gradients. Shen et al. [7] recognized 11 insect species with the DL network, and achieved a mean accuracy above 99%. Zhu et al. [8] compared the recognition effects of several convolutional neural networks (CNN) on insect images, and concluded that DenseNet-121 realizes the best effect in the shortest time.

The key to insect image segmentation is to separate the complex background. After all, the insects usually vary in tilt and deformation, owing to the complexity of natural environment and the interference of numerous factors (e.g. soil, and weeds). Chen et al. [9] combined the local entropy threshold and Otsu's method to segment insect images. Moughal [10] studied insect classification and image segmentation, using hyperspectral reflectance and SVM. Arif and Akram [11] designed an insect classification method based on fuzzy least squares (LS) SVM, which segments insects in complex background in the YCbCr color space. Wen et al. [12] extended ant colony algorithm into an expert diagnosis system suitable for the discrimination between different insects: the system automatically selects a group of best features from various extracted features by ant colony algorithm, and performs SVM segmentation of insect images with complex background.

To sum up, fruitful results have been achieved on image recognition of insects based on computer vision. However, there is ample room to improve the application range and classification efficiency of the existing insect recognition methods, especially the recognition accuracy in complex environment.

### 3. MORPHOLOGICAL EDGELESS ACTIVE CONTOUR ALGORITHM

Edge-based image segmentation draws the ideal edge curve of the original image to determine the ROI, before formally segmenting the image. This paper puts forward a novel edge-based image segmentation model, which can be expressed as a mathematical function of the energy functional:

$$E_s = \int [E_{in}(v(s)) + E_{ex}(v(s))]ds \quad (1)$$

Thus, the drawing of the ideal edge curve is transformed into the search for the minimum energy functional. Formula (1) shows the energy structure is divided into internal energy  $E_{in}$  and external energy  $E_{ex}$ . Due to the interaction between the two kinds of energies, the edge curve shrinks continuously until reaching the ROI edge.

The external energy  $E_{ex}$  is generally defined empirically. It is easily affected by image features. The gradient of  $E_{ex}$  changes proportionally with the gray value of the image. The gradient change curve of  $E_{ex}$  skews outward under the guidance of the energy:

$$E_{ex}(x, y) = -|G(x, y) * \nabla f(x, y)|^2 \quad (2)$$

where,  $\nabla f(x, y)$  is gradient change of the image;  $G(x, y)$  is the

two-dimensional (2D) Gaussian function of image;  $*$  is the convolution of external energy.

The internal energy  $E_{in}$  can be defined by:

$$E_{in}(v(s)) = \mu(s) \left| \frac{dv}{ds} \right|^2 + \vartheta(s) \left| \frac{d^2v}{ds^2} \right|^2 \quad (3)$$

where,  $\mu(s)$  is the weight coefficient of elastic energy;  $dv/ds$  is the first-order derivative of elastic energy, which is inversely proportional to the elongation of the curve;  $d^2v/ds^2$  is the second-order derivative of curve energy, which is inversely proportional to the change degree of the ideal edge curve.

Under the action of internal and external forces, the proposed model seeks the minimum energy of elastic deformation through a given initial contour model. The search is not restricted by the morphology of the ROI. Instead, the proposed model can quickly detect the edge of any shape, merge various information (e.g. edge, initial estimation, and target constraint) into a whole, and accurately pinpoint the edge of the target.

Next, this paper designs a morphology-based edgeless active contour algorithm. Traditionally, image segmentation models need to solve a partial differential equation on floating-point array, which is time- and compute-intensive. Based on the edge-based image segmentation model, morphological gray scale and narrow-band were introduced to realize efficient image segmentation. The designed algorithm is robust against noise interference, providing a desired tool for the segmentation of images with complex background.

In the designed algorithm, the curve weight coefficients were adjusted and combined in the optimal manner to segment insects in complex background. The energy functional state of the algorithm can be expressed as:

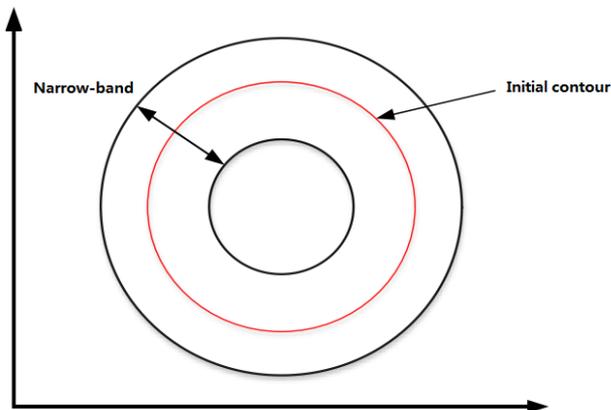
$$E(C, g_1, g_2) = \alpha L(C) + \beta A(\text{inside}(C)) + \rho_1 \int_{\text{inside}(C)} |g - g_1|^2 dx dy + \rho_2 \int_{\text{inside}(C)} |g - g_2|^2 dx dy \quad (4)$$

where,  $C$  is the boundary part enclosed by a contour;  $g_1$  and  $g_2$  are the mean gray values of the inner and outer parts of edgeless contour;  $\alpha$  and  $\beta$  are constant parameters limiting the length  $L$  and area  $A$  of the enclosed part, respectively;  $g$  is the pixel value of the image;  $\rho_1$  and  $\rho_2$  are the weights used to adjust contour changes.

The object of this research is the insect image after the removal of complex natural background. To realize automatic segmentation, the algorithm must be highly applicable and efficient at the same time. The proposed algorithm processes gray image with alternating sequence filter (ASF) [13] morphological operator: open and close operations are conducted as per the gray image cycle of insects, followed by the calculation of the area of smooth region. Once the preset termination condition is satisfied, the operations are terminated to obtain a smooth image of the insect area in the original image. The foreground, namely, insects, is highlighted through smoothing, reducing the gray redundant information of the complex background. This lays a good basis for subsequent operations like edge detection and contour segmentation.

In addition, the narrow-band fast method was employed to

accurately segment the foreground: when the contour moves to the narrow band range, a new narrow band is redefined based on the center of the current contour line; by limiting the time step, the calculation range is reduced to the narrow-band network of the contour lines; the expansion of contour line is stopped, once the contour line gets close to the edge of the image. Figure 1 illustrates the initial contour and narrow band.



**Figure 1.** The initial contour and narrow band

The narrow-band fast method can be implemented in the following steps:

Step 1: Define initial contour and narrow band. Summarize the current situation of distance by searching for the shortest distance from all points in the narrow band to each point of the contour line.

Step 2: Define the boundary points that currently appear in the narrow band, and set up observation points. The boundary points and observation points are close to the inner loop curve and within the outer loop curve, respectively.

Step 3: Define the termination condition. If the contour curve approaches the observation point in the preset narrow band, terminate the contour search, and redefine a new narrow band centering on the current contour. Since the width of narrow band is fixed, define the velocity of the nearest point in the contour line as the curvature velocity of each point in the new narrow band.

Step 4: Add a time step to each point of the redefined narrow-band, and calculate the new distance function. Since the value of the new function is zero, the points on the next contour curve can be identified, and the new initial contour lines can be connected in sequence.

Step 5: Repeat Steps 2 and 3 until meeting the preset number of iterations or the zero-level set curve does not change. Guide the movement of the closed contour curve in the ROI until meeting the termination condition.

By the morphological-based edgeless active contour algorithm, the insect segmentation can be implemented in six steps:

Step 1: Input an insect-containing gray image.

Step 2: Remove corner convexity and fill holes to obtain a smooth part through ASF morphological operation.

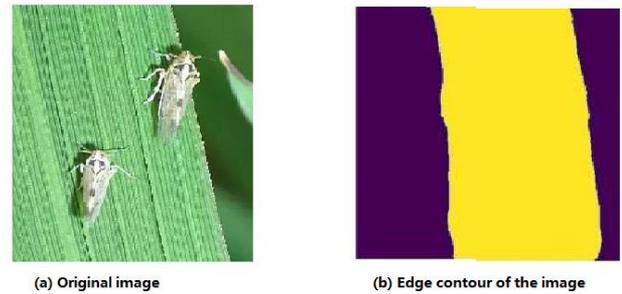
Step 3: Determine the location and size of the mask according to the insect area, and choose the rectangular mask based on the insect shape.

Step 4: Generate the initial contour and narrow band corresponding to the mask size.

Step 5: Generate a new narrow band based on the position of the next contour line, using the narrow-band fast method.

Step 6: When the contour curve shrinks to the image boundary or the number of iterations meets the preset value, the edge curve meets the definition of minimum energy functional.

Through the above steps, the authors obtained the edge contour of an insect image (Figure 2).



**Figure 2.** The result of insect segmentation by our algorithm

#### 4. VARIABLE WEIGHT EDGE ENHANCEMENT

Dataset imbalance can be solved by four popular methods: dataset expansion by data mining; dataset expansion through repeated sampling; supplementing small class samples with artificial data; modifying the classification algorithm. The first strategy continuously mines and increases the data of small classes to the mean size of the other classes. The second strategy expands the data in small classes by duplicating these small class samples. The third strategy obtains varied sample attributes from the diverse attribute space of small class samples, and uses them to produce new samples. The fourth strategy improves the classification algorithm based on balanced data, without changing the dataset.

The third strategy is the simplest among the four. But the small class samples in artificial data tend to be distorted, flipped, or compressed. Without considering the relationship between attributes, this strategy might destroy the linear relationship of each attribute. For this reason, the SMOTE algorithm [14] came into being, which does not undermine the linear relationship in unbalanced datasets. Nevertheless, the SMOTE algorithm has two obvious disadvantages:

(1) The number  $K$  of the nearest neighbors needs to be determined manually, and confirmed by the user based on his/her experience. The nearest neighbors are selected randomly, without a fixed direction.

(2) During the continuous selection of the nearest neighbors, the new samples tend to concentrate on edges, because the data samples in small classes gather on the edges. Thus, it is difficult for the classifier to distinguish between different classes.

In addition, the SMOTE algorithm uses the same weight for the original data, during the generation of new small class samples, while insects are mainly located at the center of the images. To enhance the background diversity of new samples, this paper improves the SMOTE algorithm into a variable weight edge enhancement algorithm. Based on the SMOTE algorithm, the proposed algorithm increases the weight of the edge area as adjacent images are superimposed into a new image, making the background of the new image more complex. The workflow of the proposed algorithm is explained as follows:

Step 1: Unify the size of original images to  $M \times M$  through

linear interpolation.

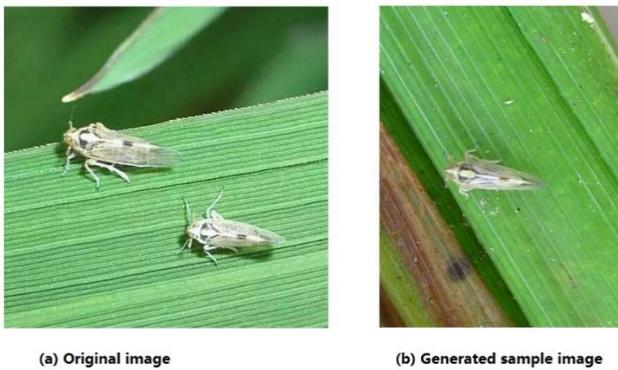
Step 2: Expand the red-green-blue (RGB) channels of each image into a one-dimensional (1D) array with a length of  $3 \times M \times M$ .

Step 3: Iteratively process each array  $X$  in small class, compute the Euclidean distance, and identify the  $K$  nearest neighbors in the array.

Step 4: Let  $X_n$  be the  $n$ -th nearest neighbor, and  $(x_1, x_2)$  be the position of each element  $x$  in the original image. If position satisfies  $20 < x_1$ , and  $x_2 < 204$ , set the weights of the center area and the edge as  $w_c = rand(0,1)$ , and  $w_b = 2 \times w_c$ , respectively. Then, generate a new sample by:

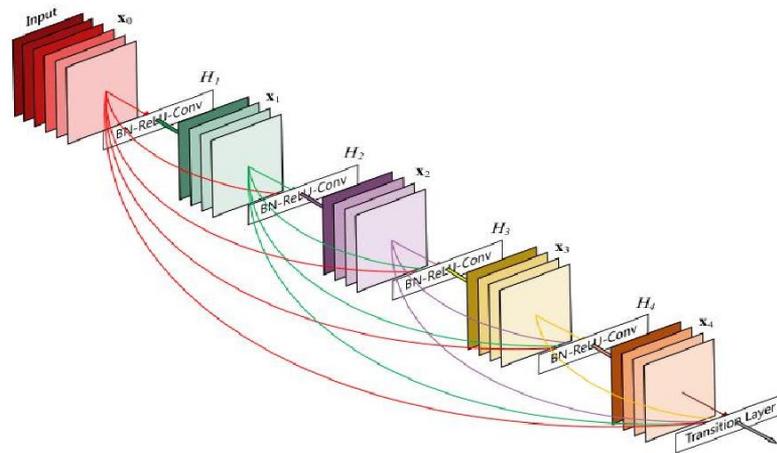
$$X_{\text{new}} = \begin{cases} X + w_c \times (X_n - X), & 20 < x_1, x_2 < 204 \\ X + w_b \times (X_n - X), & \text{otherwise} \end{cases} \quad (5)$$

Step 5: Divide new array into the three channels and restore the original image.

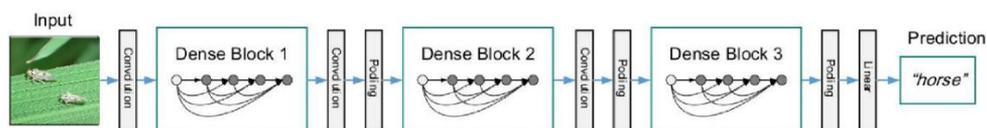


**Figure 3.** The result of the variable weight edge enhancement algorithm

As shown in Figure 3, the new image generated by our



**Figure 4.** The structure of the DenseNet



**Figure 5.** The dense blocks of the DenseNet

algorithm have more complex background than, yet similar object as the original image.

## 5. CNN-BASED INSECT RECOGNITION IN COMPLEX BACKGROUND

To verify its effectiveness, the proposed morphology-based edgeless active contour algorithm and variable weight edge enhancement algorithm were tested on DenseNet-121 trained by the insect image set processed by them. In the DenseNet [15], the layers are connected in the feedforward way (Figure 4).

General CNNs involve multiple operations, namely, convolution, pooling, and transition. To streamline the information flow in the network, a direct connection between any layer to all subsequent layers is designed as follows: Suppose layer 1 of the network receives feature maps  $x_0, \dots, x_{l-1}$  from all previous layers:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (6)$$

where,  $H_l()$  is the composite function of three continuous operations;  $[x_0, x_1, \dots, x_{l-1}]$  is the feature map set output by 0, ...,  $l-1$ -th layer. Because of this dense connection, this network is called DenseNet.

The cascade operation in formula (6) is not feasible if any change takes place to the size of the feature map. Thus, it is necessary to modify the size of the down sampling layer. To facilitate down sampling, the network needs to be divided into multiple interconnected dense blocks (Figure 5).

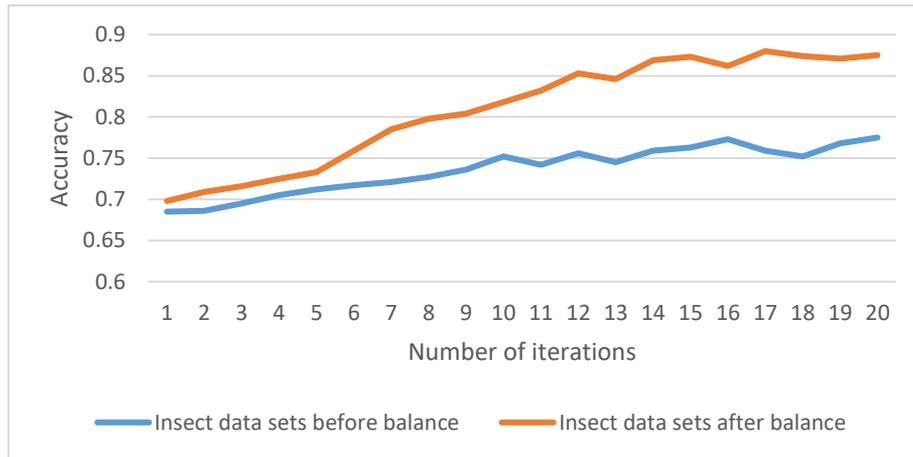
In a DenseNet with three dense blocks, the layer between two adjacent blocks is called a transition layer. In this paper, the transition layers include a batch normalization layer, a  $1 \times 1$  convolution layer, and a  $2 \times 2$  average pooling layer. The last layer of the network is a softmax classifier. Figure 6 shows the effect of automatic insect recognition by DenseNet-121.



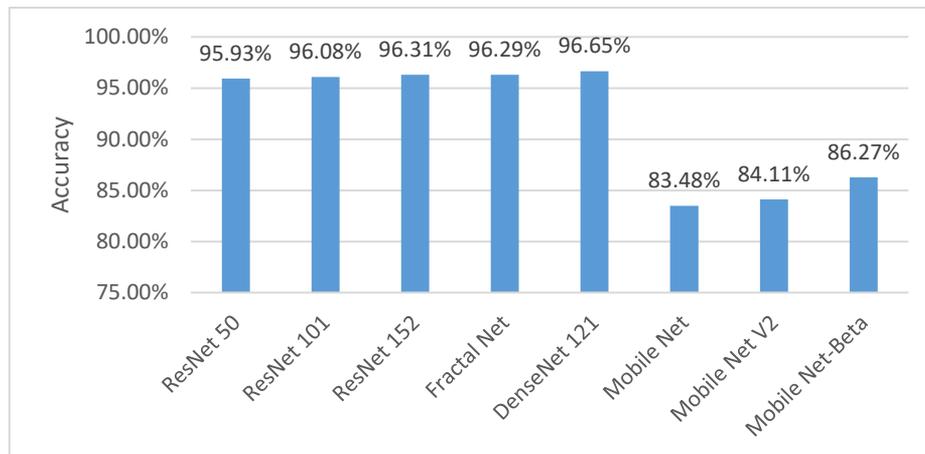
**Figure 6.** The effect of automatic insect recognition

**Table 1.** The recognition accuracies and training times of eight CNNs

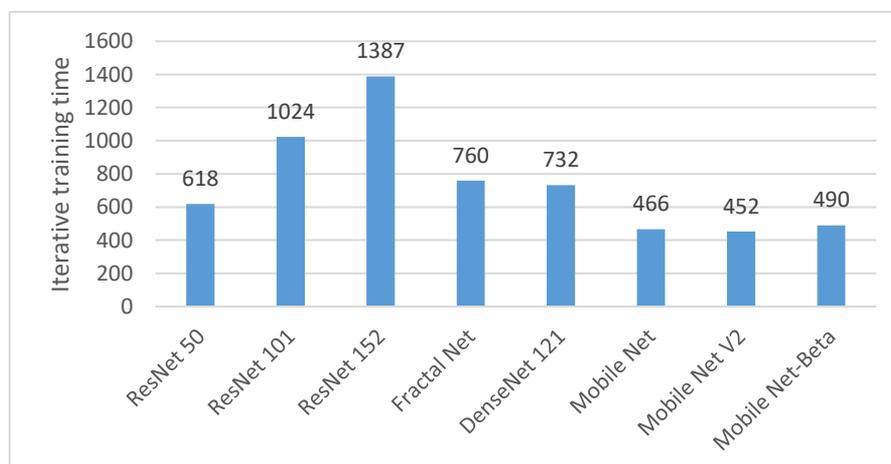
CNN	Accuracy	Training time (s)
<b>ResNet 50</b>	95.93%	618
<b>ResNet 101</b>	96.08%	1,024
<b>ResNet 152</b>	96.31%	1,387
<b>Fractal Net</b>	96.29%	760
<b>DenseNet 121</b>	96.65%	732
<b>Mobile Net</b>	83.48%	466
<b>Mobile Net V2</b>	84.11%	452
<b>Mobile Net-Beta</b>	86.27%	490



**Figure 7.** The results of DenseNet-121 on unbalanced and balanced insect image sets



**Figure 8.** The recognition accuracies of eight CNNs



**Figure 9.** The training times of eight CNNs

Figure 7 compares the results of DenseNet-121 on unbalanced and balanced insect image sets. It can be seen that the accuracy of the network was nearly 10% higher on the balanced set than on the unbalanced set.

Next, eight CNNs, including ResNet-50, ResNet-101, ResNet-152, Fractal Net, DenseNet-121, Mobile Net, Mobile Net V2, and Mobile Net-Beta, were compared on the balanced insect image set. As shown in Table 1 and Figure 8, DenseNet-121 achieved the highest recognition accuracy (96.65%).

In terms of training speed (as shown in Figure 9), MobileNet V2 with linear bottleneck inverse residual structure was the fastest, taking up only 452s in each iteration. Overall, DenseNet outperformed MobileNet, ResNet, and their improved networks.

## 6. CONCLUSIONS

This paper proposes a CNN-based automatic insect image recognition algorithm. Firstly, the morphology-based edgeless active contour was adopted to segment insect image with complex background. Next, the SMOTE algorithm was improved to enhance the background diversity of new samples. Through contrastive experiments, it is confirmed that the proposed method can achieve the best recognition effect on the DenseNet-121. The research results provide new insights to the automatic recognition of targets in images with complex background.

## REFERENCES

[1] Sun, G., Sheng, B., Dong, L. (2018). New trend of image recognition and feature extraction technology introduction. *Traitement du Signal*, 35(3-4): 205-208.

[2] Wen, C., Guyer, D. (2012). Image-based orchard insect automated identification and classification method. *Computers and Electronics in Agriculture*, 89: 110-115. <https://doi.org/10.1016/j.compag.2012.08.008>

[3] Manickavasagam, K., Sutha, S., Kamalanand, K. (2014). Development of systems for classification of different plasmodium species in thin blood smear microscopic images. *Journal of Advanced Microscopy Research*, 9(2): 86-92. <https://doi.org/10.1166/jamr.2014.1194>

[4] Zhang, H., Mao, H., Qiu, D. (2009). Feature extraction for the stored-grain insect detection system based on image recognition technology. *Transactions of the Chinese Society of Agricultural Engineering*, 25(2): 126-130.

[5] Malvade, N., Anami, B.S., Hanamaratti, N.G. (2017). Bulk paddy grain ageing period classification using RGB

and HSI color features. *International Journal of Computer Applications*, 176(5): 33-43. <https://doi.org/10.5120/ijca2017915577>

[6] Hinton, G.E., Osindero, S., Teh, Y.W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7): 1527-1554. <https://doi.org/10.1162/neco.2006.18.7.1527>

[7] Shen, Y., Zhou, H., Li, J., Jian, F., Jayas, D.S. (2018). Detection of stored-grain insects using deep learning. *Computers and Electronics in Agriculture*, 145: 319-325. <https://doi.org/10.1016/j.compag.2017.11.039>

[8] Zhu, L.Q., Ma, M.Y., Zhang, Z., Zhang, P.Y., Wu, W., Wang, D.D., Wang, H.Y. (2017). Hybrid deep learning for automated lepidopteran insect image classification. *Oriental Insects*, 51(2): 79-91. <https://doi.org/10.1080/00305316.2016.1252805>

[9] Chen, J., Guan, B., Wang, H., Zhang, X., Tang, Y., Hu, W. (2017). Image thresholding segmentation based on two dimensional histogram using gray level and local entropy information. *IEEE Access*, 6: 5269-5275. <https://doi.org/10.1109/ACCESS.2017.2757528>

[10] Moughal, T.A. (2013). Hyperspectral image classification using support vector machine. In *Journal of Physics: Conference Series*, 439(1): 012042. IOP Publishing. <https://doi.org/10.1088/1742-6596/439/1/012042>

[11] Arif, M., Akram, M.U. (2010). Pruned fuzzy K-nearest neighbor classifier for beat classification. *Journal of Biomedical Science and Engineering*, 3(4): 380-389. <https://doi.org/10.4236/jbise.2010.34053>

[12] Wen, C., Wu, D., Hu, H., Pan, W. (2015). Pose estimation-dependent identification method for field moth images using deep learning architecture. *Biosystems Engineering*, 136: 117-128. <https://doi.org/10.1016/j.biosystemseng.2015.06.002>

[13] Bin, S., Sun, G. (2020). Optimal energy resources allocation method of wireless sensor networks for intelligent railway systems. *Sensors*, 20(2): 482. <https://doi.org/10.3390/s20020482>

[14] Ramentol, E., Gondres, I., Lajes, S., Bello, R., Caballero, Y., Cornelis, C., Herrera, F. (2016). Fuzzy-rough imbalanced learning for the diagnosis of High Voltage Circuit Breaker maintenance: The SMOTE-FRST-2T algorithm. *Engineering Applications of Artificial Intelligence*, 48: 134-139. <https://doi.org/10.1016/j.engappai.2015.10.009>

[15] Zhou, F., Li, X., Li, Z. (2018). High-frequency details enhancing DenseNet for super-resolution. *Neurocomputing*, 290: 34-42. <https://doi.org/10.1016/j.neucom.2018.02.027>