Predictive Classification Model Based on Vespa Comprehensive Detection System

Ci Gong 1, *, Gefeng Deng 1, Haicheng Shan 1, Changxin Deng 2

1 School of Information Engineering, Wuhan University of Technology, Wuhan, China
2 Wuhan University of Science and Technology City College, Wuhan, China

*Corresponding author e-mail: gongci@whut.edu.cn

Abstract. Aiming at the problem that it is difficult to distinguish the authenticity of a large number of eyewitness reports when Asian giant hornet invaded, a multi-scale comprehensive detection model of Asian giant hornet was designed. The model uses cellular automata to process geographical information, YOLOv3 and improved Resnet101 to process image information, TF-IDF and Linear SVC to process textual information, and fuses the three kinds of information with a certain weight to obtain the final detection probability of Asian giant hornet. Compared with a single detection algorithm, this model makes use of multi-scale and multi-class information, and has higher classification accuracy and better robustness.

Keywords: Cellular automata method, YOLOv3, ResNet101, Linear SVC model.

1. Introduction
The Asian giant hornet (Vespa mandarinia) is the largest hornet in the world. In December 2019, an Asian bumblebee was spotted in Washington State. Asian giant hornet will attack other insects such as bees, and attack humans when threatened. If effective measures are not taken in time to eradicate the Asian giant hornet, its spread will have a negative impact on the ecological environment, socio-economic and public health of Washington State and the entire United States.

2. Cellular Automata Model
Cellular Automata (CA) is a method that simulates the overall behavior of the system by simulating local rules and local connections. We use the cellular automata model framework combined with the characteristics of the Asian Hornet to simulate and predict the spread of the Asian giant Hornet.

The strength of density-dependence experienced by nest i is determined by a Gaussian competition kernel measuring the strength of interaction with all other nests in the population [1]. To measure the interaction between two nests, we explore the competition between two nests for the same food source. We use twice the foraging distance as the standard deviation of the Gaussian distribution to obtain the strength of competition Si

$$S_i = \sum_{j=1}^{n} \exp \left( -\frac{\delta_{ij}^2}{2\theta^2} \right)$$
Where $\delta_{ij}$ is the distance between nest i and j, and $\theta_s$ is twice the foraging distance. According to the background information from Pennsylvania State University Extension, the average foraging distance of the Asian Hornet is 1-2 kilometers, and the farthest is no more than 8 kilometers.

The environmental suitability at a given site is the weighted average across all habitats that are likely to be visited. We take the Washington State Department of Agriculture as this given point. Regarding the influence of the location of the nest, we divide the type of habitat and assign a certain weight to determine its influence. Based on a Gaussian average, we get the equation for environmental suitability:

$$E_i = \epsilon \sum \exp \left( -\frac{\delta_y^2}{2\theta_f^2} \right) \cdot T_j$$

Where $\epsilon$ is a normalized parameter and $\theta_f$ is the average foraging distance.

We extract geographic information through the given table data, divide the map into grids, and calculate the sum of the number of reports submitted in the specific grid range. The larger the total number of report submissions in the grid, the closer the geographic category of the grid is to plain areas and residential areas, while the grids with a small number of report submissions are often distributed in waters and high mountains. Since one of the food sources of Asian bumblebees is bees raised by humans, the report data of other species of bumblebees also indicate that such areas are more suitable for the survival of bees. Therefore, we have reason to believe that the Asian Hornet is more likely to be distributed in areas closer to residential areas, that is, areas where more reports are submitted. In this way, we can determine the value of the geographic information weight $T_j$ as shown in the Figure 1.

![Figure 1. The value of geographic information weight $T_j$](image)

According to the assumption mentioned above, the number of queen bees conforms to the Poisson distribution. Combining the influence of nest latitude, we consider the reproductive ability of the queens, the environmental suitability of the nests and the competition between nests to obtain the relevant parameters of Poisson distribution. The number $N_i$ of new fertilized queens from nest i is formulated by:

$$N_i = \text{Poisson} \left( \frac{r \cdot E_i \cdot l_i}{1 + \gamma S_i} \right)$$

Where $r$ is the reproduction rate of the queens. Once the new fertilized queens are produced, the new nests they create will appear in different locations. The location of the nest is affected by the queen’s ability to spread and the type of habitat. The probability that a new fertilized queen from location i creates a nest in location j in the subsequent year is proportional to:
Where $\sigma$ is the average flight distance, and the value is 30 km [2]. Based on the above relationship, we set the proliferation rules of Asian Hornet nest cells contained in the cellular automata. For cells with nests $i$, after calculating the intensity of competition between nests and environmental fitness, we generate random numbers that satisfy the corresponding Poisson distribution according to the calculation formula of the number of newly fertilized queens in the nest. At the same time, we randomly assign a certain number of queen bees to neighbor cells according to the probability that the queen bee builds a nest in other positions in the neighbor, and complete the formulation of local rules between the cells.

### 3. Parameter Estimation

We use the distribution of Asian Hornet nests in 2019 as the initial conditions, and use the nest distribution represented by the 2020 Asian Hornet forward report and the location information of the Asian Hornet false report in 2019 to estimate the parameters with maximum likelihood. The estimated result is shown in Figure 2.

![Figure 2. The estimated result](image)

We determine that the estimated value of the parameter $r$ is 12, and calculate the confidence interval of bootstrap to get its 95% confidence interval of 4.28-22.6. Similarly, we determine that the estimated value of the parameter $y$ is 1.2, and its 95% confidence interval is 0.61-4.77.

### 4. Image RecognitionWasp Detection Based on YOLOv3

For image files, we process the files in different ways. We directly extract the pictures in the document, extract frames from the videos, and convert pictures in png format into jpg format. In this process, we do not consider the pictures whose lab status is unprocessed. Since the number of pictures with positive ID is small, we extract the pictures of the Asian giant hornet in the background information and get 19 pictures with positive ID. In addition, we randomly select 781 pictures containing common insects from the picture files.

For the image classification model, whether it is vespa or other insects, it belongs to small target recognition, that is, due to the small size of these insects, specific features are difficult to capture in the entire image. Therefore, in order to improve the accuracy of image classification so that it can recognize small vespa targets in unbalanced samples, first use the YOLOv3 model to identify and crop the entire insect, and then use resnet101 to perform the cropped insect target classification [3]. In this cropped
insect image, the specific characteristics of the insect can be extracted more easily, and the resnet will be less affected by other environmental features, and its classification accuracy will be higher than using the entire image for training [4].

5. Image Classification: ResNet101 Model
The pictures processed by Resnet are the insects pictures after YOLOv3 recognition and cropping, which contain positive samples. Since the positive samples in the given data are too scarce, there is a risk of overfitting when directly substituted into the training. Therefore, we perform data enhancement processing of random cropping, random horizontal flipping and random upside-down flipping on 19 cropped positive samples. After the treatment, the pictures of the positive samples become 70 cases.

After completing wasp detection based on YOLOv3, we perform image recognition to determine the identity of the target, that is, whether the target is an Asian giant hornet or other insects. ResNet101 uses a shortcut connection to form a building block structure. In order to reduce the number of parameters, the first 1x1 convolution reduces the 256-dimensional channel to 64-dimensional, and finally restores it through 1x1 convolution. The overall number of parameters used is 69632 [5].

The model first enters a $7 \times 7 \times 64$ convolution, and then passes through 33 building blocks, each with 3 layers, for a total of 99 layers. The last fc layer is used for classification to form a 101-layer network.

6. Model Update
In order to combat the imbalance of positive and negative samples caused by the scarcity of Asian giant hornet samples, we make some changes to the configuration of the ResNet101 network. The optimizer uses the Adam optimizer, and the cross-entropy loss function uses the Focal loss function that performs better in the case of unbalanced samples and a large number of simple samples.

7. Text Categorization
In order to predict the likelihood of a mistaken classification, we perform text classification on the notes in the data set file. Data preprocessing is a necessary process before training the classifier model. After we count the notes information of all data, we get the word cloud as shown in the Figure 3.

Figure 3. Word cloud

We quantify the text data. In order to evaluate the importance of a word to a document set or a document in a corpus, we choose to use TF-IDF (Term FrequencyInverse Document Frequency) to weight the data. The importance of a word increases in proportion to the number of times it appears in the document, but at the same time it decreases in inverse proportion to the frequency of its appearance in the corpus. We use sklearn’s TfidfVectorizer method to calculate the weight of the inverse word frequency of each word, and use the TF-IDF value to indicate the importance of each word in the text.

By calculating the weight of the inverse word frequency of each word, we generate a sparse matrix based on the weight, and get the weight value of the TF-IDF expressed as a floating point number. After processing the feature, we process the label data. In order to train the classifier more conveniently, we convert the label data into a bag of words (BOW) format. Then we use sklearn’s MultiLabelBinarizer method to binarize tags.
We train the classifier after extracting features and labels. The pipeline command (Pipeline) provided by sklearn can train the model in batch mode. Before training, we define an evaluation function to evaluate the accuracy of our prediction results.

8. Linear Support Vector Machine Model Based on TF-IDF

We establish the linear support vector machine model, which is commonly used in machine learning, and form the optimal combination with TF-IDF.

Linear support vector machines is originally formulated for binary classification. We enter linearly separable training set \( T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_N, y_N)\} \) \((x_i \in X = \mathbb{R}^n, y_i \in \{1, -1\}, i = 1, 2, \ldots, N\) ), and obtain the output separation hyperplane and classification decision function. The specific algorithm steps are as follows [6].

Step1. Select the penalty coefficient \( C > 0\), construct and solve the convex quadratic programming problem.

\[
\min_{\alpha} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) - \sum_{i=1}^{N} \alpha_i
\]

s.t. \( \sum_{i=1}^{N} \alpha_i y_i = 0 \)

\( C \geq \alpha_i \geq 0, i = 1, 2, \ldots, N \)

We find the optimal solution: \( \alpha^* = (\alpha^*_1, \alpha^*_2, \ldots, \alpha^*_N) \).

Step2. Calculate the following equation:

\[
\alpha^* = \sum_{i=1}^{N} \alpha_i^* y_i x_i
\]

Choose a component \( \alpha^* \) of \( \alpha^* \) to fit the condition \( 0 < \alpha^* < C \).

\[
b^* = y_j - \sum_{i=1}^{N} \alpha_i^* y_i (x_i \cdot x_j)
\]

Step3. The classification decision function is:

\[
f(x) = \text{sign}(w^* \cdot x + b^*)
\]

9. Image and Text Fusion: Comprehensive Asian Giant Hornet Detection Model

We calculate the probability that the report belongs to the positive state based on the geographical longitude and latitude information, image information and text information of the witness report. Therefore, we propose the Asian giant hornet comprehensive detection model, combine the probabilities of the three types of information extracted from the sighting report by the above three models, and give the comprehensive probability of whether the report is positive.

In all reports, image information is the most intuitive information. All messages with a status of positive have corresponding pictures, and most reports with a status of unverified have no relevant pictures. Therefore, the weight of image information should be the largest. Geographical information and textual information can also reflect the authenticity of the report at a certain level, so the weights of the two are equal. Finally, the probability calculation formula of Asian giant hornet is:

\[
P = r_1 \times p_1 + r_2 \times p_2 + r_3 \times p_3
\]

Where \( r_1, r_2, \) and \( r_3 \) are the weights of geographic, image, and text information. They are taken as 0.25, 0.5 and 0.25 in this article.

10. Cellular Automata Model results

After obtaining the estimated value of the parameter, we substitute it into the cellular automata model. We set the model evolution algebra to 3 and 6 generations respectively (according to the migration habits of the Asian Hornet, one generation per year), due to the existence of the evolution process In a random process, the result of each evolution is equivalent to a sampling of the whole.

Therefore, we use the Monte Carlo method to carry out multiple independent repeated experiments on the evolution of cellular automata to obtain the results in each region after the evolution of 3 and
generations. There is a probability distribution of Asian hornet invasion and nesting. The model solution result is shown in the Figure 4.

![Figure 4. The model solution result](image)

By calculating the confidence interval of bootstrapping, the 95% confidence interval of longitude is -122.68-122.42 for three generations, the 95% confidence interval for latitude is 48.733-49.248, and the 95% confidence interval for longitude is -122.83-122.1 when the evolution is 6 generations, and the latitude is 95%. The confidence interval is 48.334-49.262. The above figure shows the spread of the Asian giant hornet in Washington State. When the man-made removal after submission of the report is not considered, the Asian giant hornet will continue to spread at a faster rate under the estimated parameters. Due to the introduction of geographic factors, it can be seen that the propagation probability of Asian giant hornet in mountains and water terrain is relatively small.

11. YOLOv3 results

First, we use labelimg to frame the specific locations of the insects in the selected 800 pictures containing insects, and label the insects to ensure that YOLOv3 can identify these targets in any picture. Then we import a blank pre-weight file for YOLOv3, and train and test the 800 images we process through the model. The ratio of training set to test set is 0.8:0.2. Finally, we generate our own weight file after 1200 iterations.

We take a picture with a positive ID as an example. The recognition and cropping results of YOLOv3 are as follows.

1. Confidence

Confidence degree represents the probability of whether there is an object in the current bounding box, and can also represent the possible IOU value of the predicted box and the real box of the object when the current box has an object. The confidence of the jth bounding box of the ith grid cell is:

\[ C_j^i = \Pr(\text{Object}) \times IOU_{\text{truth}}_{\text{pred}} \]

We test the weights files after 1200 generations of training on the test set, and the confidence of the insect target recognition is above 0.9.

2. Accuracy and average loss

The changes in accuracy and average loss on the test set during the training of the YOLOv3 model are shown in the Figure 5. After 1200 iterations, the accuracy and average loss have stabilized. After training, the accuracy of the model on the test set reach 0.96, and the average loss is reduced to 0.12. It shows that our YOLOv3 model can well complete the recognition of insect targets.
12. Resnet101 results

We use 70 sample images of Asian giant hornet as positive samples for ResNet101 training and testing, and 730 cropped other insect samples as counter-example samples for ResNet101 training and testing to train ResNet101 classification tasks. The ratio of training set to test set is 0.8:0.2. We use the trained ResNet to implement the binary classification task of Asian giant hornet and other insects on the insects images cropped by YOLOv3.

The accuracy and average loss of the ResNet101 model on the test set during the 1000-generation training process are shown in the Figure. After 1000 iterations, the accuracy and average loss have stabilized. After training, the accuracy of the model on the test set reach 0.91, and the average loss is reduced to 0.25. The results show that the ResNet101 network can well complete the binary classification task for Asian giant hornet and other insects targets.

13. Text Categorization results

We implement a three classification task through text classification. Through the notes text of the sighting, we judge the probability that the sighting is positive, negative, and unverified.

We select 12 samples of sightings with positive status that contained notes text information, and randomly select 200 texts from sightings with negative and unverified status. Because the number of samples with positive status is too small, training directly through SVC may make it difficult for the
classifier to recognize positive information. Therefore, we use back translation and EDA methods to enhance the data of 12 positive texts to generate 100 texts. In this way, the number of our SVC training and testing samples is 500, among which 100 are positive information, 200 are negative information, and 200 are unverified information. The results of the model are as follows.

1. Confusion matrix

The main diagonal of the confusion matrix indicates the number of correct predictions, and the rest indicate the number of prediction errors. On the test set of a total of 80 samples, the predicted confusion matrix is shown in the Figure.7. All positive samples in the test set were judged correctly, indicating that the model has overcome the problem of sample imbalance. In addition, only a small number of samples of other insect types were judged incorrectly, indicating that the model can complete the text classification task well.

![Figure 7. The predicted confusion matrix](image)

2. Accuracy and F1-score-weighted

The accuracy of our Support Vector Machine Model on the test set is 0.82, and F1-score-weighted is 0.88. Its accuracy and F1-score scores are both high. Combining the above confusion matrix results, it shows that the model can more accurately complete text multi-classification tasks.

14. Example of Model Results

We take a picture as an example to calculate the overall probability of the Asian giant hornet.

According to its latitude and longitude information, we calculate the probability of being an Asian giant hornet by using model in the question 1 to be 0.86. According to its image information, we use the YOLOv3 model to identify and crop. The results are shown in the Figure 8.
**Figure 8.** According to its image information, we use the YOLOv3 model to identify and crop

| Table 1. The calculation results of the example |
|-----------------------------------------------|
| Geographic probability | Image probability | Text probability | Comprehensive probability |
|------------------------|-------------------|-----------------|--------------------------|
| 0.9459                 | 0.94              | 0.91            | 0.91                     |

**15. Conclusion**

On the basis of life habit analysis of Vespa mandarinia, the distribution probability of Vespa mandarinia in the United States in the coming years after its invasion is simulated and predicted by cellular automata, and the target detection system of Vespa mandarinia is constructed by YOLOv3 and resnet101 models. Text classification system of Vespa mandarinia is constructed by TF-IDF and Linear SVC algorithm, and comprehensive Vespa mandarinia detection system is constructed based on these three systems. Experiments show that the system has high accuracy and stability, and the designed system can accurately calculate the recognition probability of Vespa mandarinia according to different information, so as to meet the ideal detection requirements.

**References**

[1] Keeling, M.J., Franklin, D.N., Datta, S. et al. Predicting the spread of the Asian hornet (Vespa velutina) following its incursion into Great Britain. Sci Rep 7, 6240 (2017).

[2] Franklin D N, Brown M A, Datta S, et al. Invasion dynamics of Asian hornet, Vespa velutina (Hymenoptera: Vespidae): a case study of a commune in south-west France[J]. Applied entomology and zoology, 2017, 52(2): 221-229.

[3] Shi Yingying, Li Xiangrui, Sun Fan. Research on target detection method of tea buds in natural environment based on YOLOv3 [J]. Computer Knowledge and Technology, 2021, 17(03): 14-16.

[4] Chen Kai, Zu Li, Ou Yi. A photographic robot face recognition and tracking system based on YOLOv3 and ResNet101 [J]. Computer and Modernization, 2020(04): 30-36+41.

[5] D. Isa, L. H. Lee, V. P. Kallimani and R. RajKumar, "Text Document Preprocessing with the Bayes Formula for Classification Using the Support Vector Machine," in IEEE Transactions on Knowledge and Data Engineering, vol. 20, no. 9, pp. 1264-1272, Sept. 2008, doi: 10.1109/TKDE.2008.76.

[6] P.S. Bradley & O.L. Mangasarian (2000) Massive data discrimination via linear support vector machines, Optimization Methods and Software, 13:1, 1-10, DOI: 10.1080/105567800088057.