Based on XGBoost Model and SAR Model to Identify New Species and Predict the Number Change and Distribution of New Species: Based on the Invasion and Spread of Asian Hornet in the United States

Huijin Han

North China Electric Power University (Baoding), Baoding, 071003, China

Abstract: New species may have a greater impact on the ecology and economy of the invaded area. Since November 2019, Asian hornet has begun to invade the United States, which has a greater impact on humans, invertebrates, and the economy. Based on the data related to the spread of Asian hornet in the United States, this paper proposes a plan to predict the extent of its spread and identify new species. The SAR model is used to predict the spread of the Asian Hornet in the United States. According to the spatial relationship of the model and the distribution data of the Asian Hornet in the United States from May 2020 to October 2020, it also considers environmental and climate factors such as temperature, humidity, and air pressure. Perform time series forecasts to obtain the extent of the spread of the Hornet in the United States. The average RSME is 91.27, which has achieved an ideal prediction result. At the same time, based on the data of 3000 reporters, this paper trains and uses the Bi-LSTM model and the CNN model to respectively identify the pictures of the Asian Hornet reporters' explanations and feedback, and through XGBoost, based on the above two recognition results, combined The suitable temperature, humidity, air pressure and other parameters of the Hornet finally identified the Asian Hornet. Compared with the 1,386 test data, the accuracy rate was as high as 0.956.

Keywords: SAR model; time series prediction; Bi-LSTM model; text sentiment analysis; CNN model; XGBoost

1. Introduction

The introduction of alien species may disrupt the ecological balance, change the local environment and affect biodiversity. (Wilcove et al., 1998; Dueñas et al., 2018). At the same time, it also has a greater impact on the economic environment. In the United States, the annual economic loss caused by invasive species to agriculture, forestry and public health is close to 120 billion U.S. dollars. Since November 2019, the dead specimens of bumblebees were found in Washington State, and Asian bumblebees were introduced into the United States. On May 5, 2020, the bumblebees became more raging in the United States. The New York Times published an article on stopping the Asian bumblebees [1,2].

Asian hornet, belonging to the Vespa family Hymenoptera (about 5000 species of Hufeng family worldwide), is the largest known wasp species in the world, ranging in length from 38-50 mm. Compared with other wasps, the head It is orange and bald (Lee 2010; Matsuura and Sakagami 1973). This type of wasp is native to India, Nepal, Sri Lanka, Vietnam, South Korea, Japan, Taiwan and China. Mainly distributed in dense woodlands and mountains, nests are also generally placed in existing caves, snake holes or rotting tree roots . The population consists of queen bees, worker bees, and drones, each of which performs its duties. There are usually 300 or more workers (up to 800-1000 workers) in each hive, and they usually move and forage in summer and autumn.

The Asian Hornet has a greater impact on humans, some invertebrates, and the economy. The venom of the bumblebee not only causes pain in humans, it may cause allergic reactions, and in severe cases, it may damage related human tissues (Schmidt 2019). In Japan, an average of dozens of people die every year from wasps and cause thousands of injuries (New York Times 2020). The venom of the bumblebee sprayed into human eyes can cause damage to the retinal function and so on [3,4]. This type of wasp is predatory and usually feeds on various terrestrial invertebrates, especially honeybees. Bees are the main way of pollen transmission. Due to the death of a large number of bees, it will have a serious impact on effective pollination, and then affect the development of agriculture. Alberto J Alaniz and others estimated the impact of Asian hornet invading the United States. According to estimates, 95216±5551 bee colonies may be threatened; the potential threat income related to bee pollination of farmland reaches 101.8 million US dollars per year [5,6].

Based on the above analysis, how to identify and predict the Asian Hornet is particularly important. Claudia Nuñez-Penichet et al., Alberto J Alaniz et al., avid A. Moo-Llanes et al. predict the spread of bumblebees based on niche models, which are based on species distribution Points and environmental parameters are used to estimate the ecological needs of species, and then to estimate the geographic distribution of species [7]. However, this model is conservative. It assumes that the niche of the species is unchanged during a period of history, the ecological needs and distribution are in a balanced state, and the species is in a saturated state. At the same time, it assumes that the migration ability of the species is unlimited, that is, it emphasizes the species. The migration ability of the species ignores the interaction between species, the migration ability of the species itself, and the barrier effect of geography. The above hypothesis is that the distribution of Asian bumblebees that have just invaded the United States is still not saturated, and they are new species in the United States. Under the new environment, the impact of the environment and species, etc., will undergo new changes according to local conditions. There may be some deviations in predicting the bumblebees in the United States based on the environmental factors of Asian bumblebees in other regions; At the same time, the niche model can only predict the presence of bumblebees in the area, and cannot measure the severity of the species's spread in the area.

This paper is based on the Spatial Autoregressive Model (SAR) to predict the spread of Asian Hornet in the United States. The SAR model was proposed by cliff and ord in 1973. The core of this model is the spatial weight matrix, which analyzes the influence of geographical adjacent areas on the local area, the so-called spatial spillover effect [8]. This article is based on the spatial relationship of this model and the distribution data of Asian Hornet in the United States from May 2020 to October 2020. At the same time, taking into account the environmental and climate factors considered in the niche model, the time series prediction is carried out. The RSME evaluation standard evaluates the forecast results from May 2020 to October 2020. The average RSME is 91.27, and a relatively ideal forecast result has been achieved [9].

Due to the variety of Hu Feng, as a new species in the United States, it is urgent to identify the Asian hornet. Based on the data of 3000 reporters, this paper trains and uses the Bi-LSTM model (bidirectional recurrent neural network) and the CNN model (convolutional neural network) to identify the images of the Asian Hornet reporters and feedback, and use the XGBoost model Based on the two recognition results, combined with the bumblebee's suitable temperature, humidity, air pressure and other parameters, the Asian bumblebee was finally identified. Through comparison with 1,386 test data, the accuracy rate was as high as 0.956.

2. Model

2.1 Spatial autoregressive model

2.1.1 Spatial autoregressive model

The spatial autoregressive model (Spatial Autoregressive Model, SAR) mainly explores whether each variable has a diffusion phenomenon (spillover effect) in a certain area. The model expression is:

$$y = \rho W y + X\beta + \epsilon \tag{1}$$

In this article, Y represents the extent of the Asian Hornet's spread, the parameter ρ reflects the influence of the spatial lag variable Wy on the Y value, and the parameter β reflects the influence of the Y value by the climate environment X. ρ is the coefficient of the spatial weight matrix, β is the coefficient of the extracted environmental variable, and ϵ is the error term.

2.1.2 Spatial weight matrix

The core of the spatial econometric model is the spatial weight matrix. Spatial autocorrelation can be expressed by the formula [10]:

$$\mathrm{Cov}(y_i, y_j) \neq 0 \text{ for } i \neq jc \tag{2}$$

In order to calculate the covariance i and j between the observations of the random variable at the location, a spatial weight matrix needs to be introduced. There are many kinds of matrices for spatial weights. This paper chooses to use the k-value nearest neighbor space weight matrix [11]

$$W_{ij}(d) = \begin{cases} 0 & \mathrm{dis(i,j)} \leq k \\ 1 & \mathrm{dis(i,j)} > k \end{cases} \tag{3}$$

And finally normalize the matrix:

$$W_{ij} = \frac{W_{ij}}{\sum_{j=1}^{n} W_{i,j}}, i,j = 1,2,\cdots,n \tag{4}$$

2.2 BI_LSTM model

Recurrent Neural Network (Recurrent Neural Network, RNN) is a type of recursive neural network that takes sequence data as input in table 1, recurses according to the evolution direction of the sequence, and all nodes (recurrent units) are connected in a chain. neural network) [12].

Research on recurrent neural networks began in the 1980s and 1990s, and developed into one of the deep learning algorithms in the early 21st century [13], among which Bidirectional RNN (Bi-RNN) And long short-term memory networks (Long Short-Term Memory networks, LSTM) are common recurrent neural networks [14].

$$\text{Where, } \begin{cases} \boldsymbol{I}_t = \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xi} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hi} + \boldsymbol{b}_i) \\ \boldsymbol{F}_t = \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xf} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hf} + \boldsymbol{b}_f) \\ \boldsymbol{O}_t = \sigma(\boldsymbol{X}_t \boldsymbol{W}_{xo} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{ho} + \boldsymbol{b}_o) \\ \widetilde{\boldsymbol{C}}_t = \tanh(\boldsymbol{X}_t \boldsymbol{W}_{xc} + \boldsymbol{H}_{t-1} \boldsymbol{W}_{hc} + \boldsymbol{b}_c) \\ \boldsymbol{C}_t = \boldsymbol{F}_t \odot \boldsymbol{C}_{t-1} + \boldsymbol{I}_t \odot \widetilde{\boldsymbol{C}}_t \\ \boldsymbol{H}_t = \boldsymbol{O}_t \odot \tanh(\boldsymbol{C}_t) \end{cases} \tag{5}$$

Table 1. Symbol.

Symbol	Definition
h	Number of hidden units

d	Enter of the number
n	Number of samples
$\boldsymbol{X}_t \in \mathbb{R}^{n\times d}$	Small batch input at a given time step t
$\boldsymbol{H}_{t-1} \in \mathbb{R}^{n\times h}$	Hidden status at previous time step
$\boldsymbol{I}_t \in \mathbb{R}^{n\times h}$	Input gate at time t
$\boldsymbol{F}_t \in \mathbb{R}^{n\times h}$	Forgotten gate at time t
$\boldsymbol{O}_t \in \mathbb{R}^{n\times h}$	Output gate at time t
$\boldsymbol{W}_{xi}, \boldsymbol{W}_{xf}, \boldsymbol{W}_{xo} \in \mathbb{R}^{d\times h}$	Weight parameter
$\boldsymbol{W}_{hi}, \boldsymbol{W}_{hf}, \boldsymbol{W}_{ho} \in \mathbb{R}^{h\times h}$	Weight parameter
$\tilde{\boldsymbol{C}}_t \in \mathbb{R}^{n\times h}$	Candidate memory cell at time step t

2.3 Convolutional Neural Networks

Convolutional Neural Networks (CNN) is a type of feedforward neural network (Feedforward Neural Networks) that includes convolution calculations and has a deep structure. It is one of the representative algorithms of deep learning [15,16].

2.4 XGBoost model

XGBoost is an improvement of the boosting algorithm based on the gradient descent tree (Gradient Boosting Decision Tree, GBDT), which consists of multiple decision tree iterations. The basic idea is to first build multiple CART (Classification and Regression Trees) models to predict the data set, and then integrate these trees into a new tree model. The model will continue to iteratively improve, and the new tree model generated in each iteration will fit the residual of the previous tree. As the number of trees increases, the complexity of the ensemble model will gradually increase until it approaches the complexity of the data itself, and the training achieves the best results.

The XGBoost algorithm model is as follows:

$$\hat{y}_u = \phi(x_i) = \sum_{t=1}^{T} f_t(x_i) \quad (6)$$

where $\{f_t(x_i) = w_q(x)\}$ is the space of CART, $w_q(x)$ is the scoring of sample x, and the predicted value $\hat{y}_i$ of the model is obtained by accumulation, and q denotes the structure of each tree, T is the number of trees. Each f_t corresponds to an independent tree structure q and leaf weight w_q.

3. Process and method

3.1. Research process

3.1.1 Predict the degree of spread

This paper first conducts a normal analysis of the influencing factors affecting the spread of the Asian Hornet. Based on the normal analysis, through the analysis of variance, the eigenvalues affecting the spread of the Asian Hornet are screened out, and then the spatial weight matrix is established. Through the SAR model, according to the exact bumblebee distribution data training parameters from May 2020 to October 2020 are used to obtain the prediction model of bumblebee propagation. Finally, the prediction results are evaluated using the RSME evaluation standard. The specific research process is as follows Fig.1:

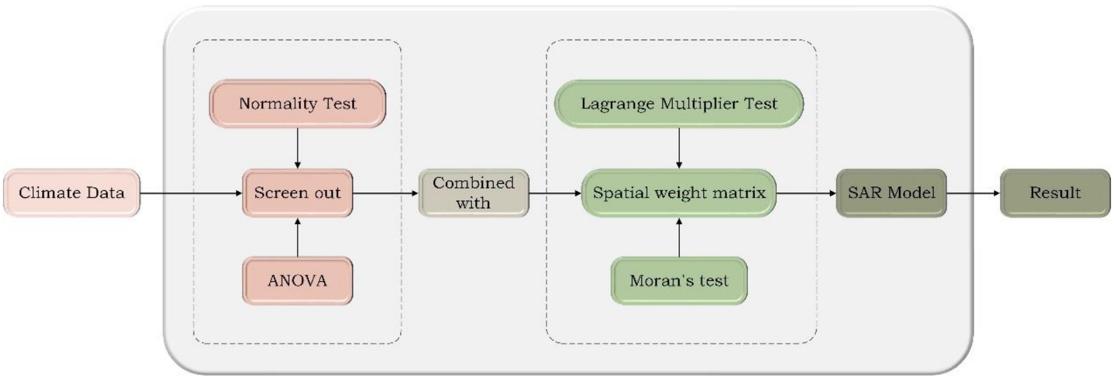

Figure 1. The specific research process

3.1.2 Identify the bumblebee

This paper uses the Bi-LSTM model (long short-term memory network) and the CNN model (convolutional neural network) to respectively identify the pictures of the Asian Hornet reporter's explanation and feedback, and through XGBoost, based on the two recognition results, combined The suitable temperature, humidity, air pressure and other parameters of the hornet finally identified the Asian hornet. The specific research process is as follows Fig.2:

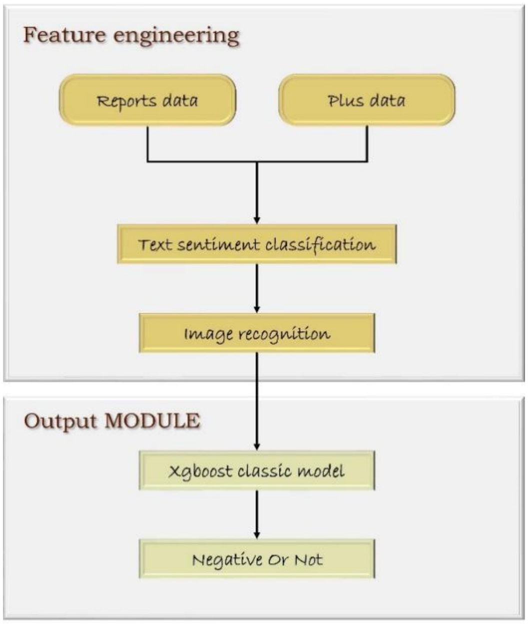

Figure 2. Classification Model

3.2. Research methods and results analysis

3.2.1 Predict the degree of spread

3.2.1.1Normal analysis and variance analysis

Since the one-way analysis of variance assumes that the sample obeys a certain normal distribution, this article needs to carry out a normal test on the collected climate and environmental data. The results are shown in the table.2 below:

Table 2. Shapiro-Wilk W Normal Test.

Symbol	Altitude	Temperature	Humidity	Precipitation	Wind speed	Air pressure	Cloud cover	Solar radiation
W	0.972	0.945	0.894	0.936	0.953	0.904	0.962	0.987
p	1.90E-08	1.10E-17	3.30E-42	6.40E-20	3.10E-15	2.40E-31	2.20E-12	2.00E-02

Through the above tests, this paper finds that climate and environmental data are similar to normal distributions. Set the significance level = 0.05. Although all p-values are less than 0.05, this article rejects H0. This article believes that the p-value is too small because the sample size is too large. Observing the value of W, it is found that the value of W is close to 1, so it can be considered that these data are approximately normally distributed.

On the basis of the above normal analysis, this article examines the influence of temperature, precipitation, humidity, wind speed and air pressure on the spread of Asian Hornet through analysis of variance. The test results are as follows Table.3:

Table 3. ANOVA Results.

Symbol	Altitude	Temperature	Humidity	Precipitation	Wind speed	Air pressure	Cloud cover	Solar radiation
W	1.526	144.667	3.464	83.441	7.646	12.093	0.628	0.649
p	0.217	8.30E-33	0.063	9.80E-20	0.006	5.10E-04	0.428	0.421

Through the above test results, this article found that temperature, precipitation, humidity, wind speed and air pressure have a significant impact on the spread of the vespa.

3.2.1.2 .SAR model analysis

Because the spread of the Asian Hornet is between 49.5 kilometers and 110 kilometers [17]. Therefore, this paper sets k to 50km, and the spatial weight matrix is:

$$W_{ij}(d) = \begin{cases} 0 & \text{dis(i,j)} \leq 50 \\ 1 & \text{dis(i,j)} > 50 \end{cases} \tag{7}$$

And normalize the weight matrix according to the following formula to get the final weight matrix:

$$W_{ij} = \frac{W_{ij}}{\sum_{j=1}^{n} W_{i,j}}, i,j = 1,2,\cdots,n \tag{8}$$

Based on the distribution data of the Asian Hornet in the United States from May 2020 to October 2020, this paper obtains data on characteristic factors such as temperature, precipitation, humidity, wind speed, and air pressure according to its latitude and longitude. Since the average number of worker bees is 300 [18], and considering the accuracy of the bumblebee report, the dependent variable is defined as:

$y =$ positive number $\times 300 +$ unverified number $\times 30$

Among them are the number of reports that accurately identify the hornet; the number of reports that fail to accurately identify the hornet.

If the predicted value of y in a certain area is greater than 300, it is considered that there is a nest of Asian hornet in this area. Due to the hibernation habit of the Asian Hornet, in order to effectively train the model parameters, this paper uses the data from May 2020 to October 2020 to train the model parameters ρ, β, and ϵ.

$$y^{i} = \rho W y^{i-1} + X^{i}\beta + \epsilon \quad (9)$$

Among them, y^{i} is the dependent variable of the i-th year, and X^{i} is the independent variable of the i-th year.

The training model is as follows:

$$y^{11} = 0.43Wy^{10} - 97.57x_1^{11} + 68.51x_2^{11} - 47.75x_3^{11} - 12.22x_4^{11} - 27.85x_5^{11} + 19.6042 \quad (10)$$

3.2.1.3 .Result analysis

This article analyzes the spread of the Asian Hornet from two aspects.

First of all, from the geographic location dimension, this article draws the forecast results for September, October, November, December 2020 and January and February 2021 into a heat map, as shown in the following Fig.3:

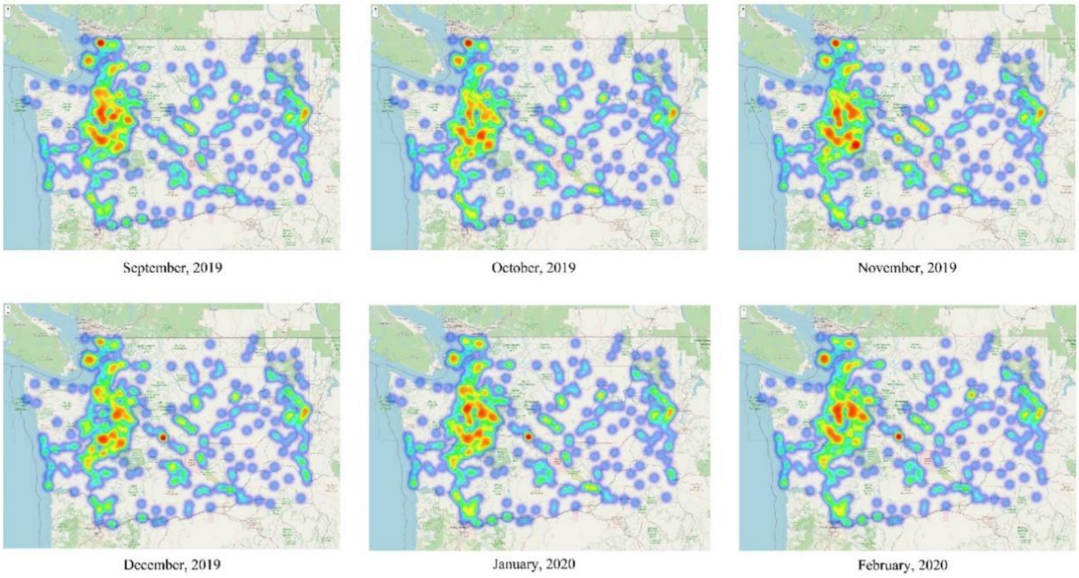

Figure 3. Heat map of predicted results.

According to the above heat map, it is found that the Asian Hornet is mainly distributed in northwestern Washington State, adjacent to Canada. At the same time, we can see that the center of the wasp gathering in Washington State is shifting from the northwestern United States like the central region. As of December 2020, it has been possible to discover that new gathering centers have emerged in parts of the central region.

Secondly, according to the value of the influencing factor Y, this paper selects five major cities, including Point Roberts, Blaine, Custer, Rockport, and Chelan Falls, and observes the changes in the number of bumblebees in a certain area according to the changes in the Y value.

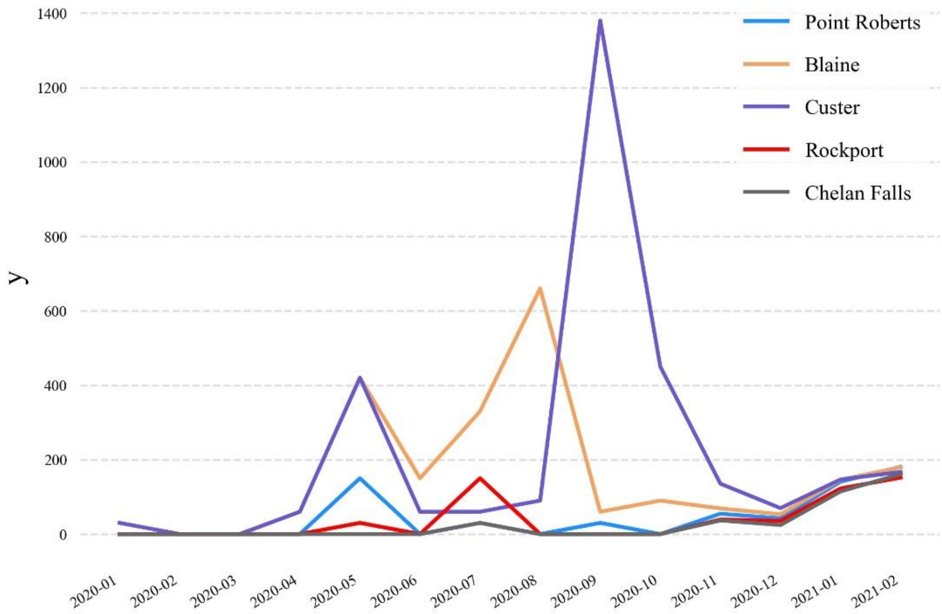

Figure 4. Changes in Y value of major cities.

According to the above Fig.4, the y value of important cities fluctuates. The number of bumblebees in Point Roberts and Chelan Falls has been declining since May 2020. On the contrary, Custer increased significantly in September and October, and it is likely to feed back in the summer and autumn of 2021. However, in February 2021, the number of them was less than 300, which confirmed the life habits of the bumblebee itself, that is, the main active seasons are summer and autumn, and winter is in hibernation, which also proves the accuracy of our model.

3.2.1.4 Inspection

This article uses the report data from June to October 2020 as a test set to evaluate our model in table 4. The evaluation standard is RMSE:

Table 4. The MSE of our model.

Month	5	6	7	8	9	10
RMSE	31.23	44.83	94.88	113.04	100.9	162.75

The average RMSE calculated in this paper is 91.27.

3.2.2 Identify the bumblebee

3.2.2.1 Text recognition model

The specific process of Notes for text emotion processing is as follows Fig.5:

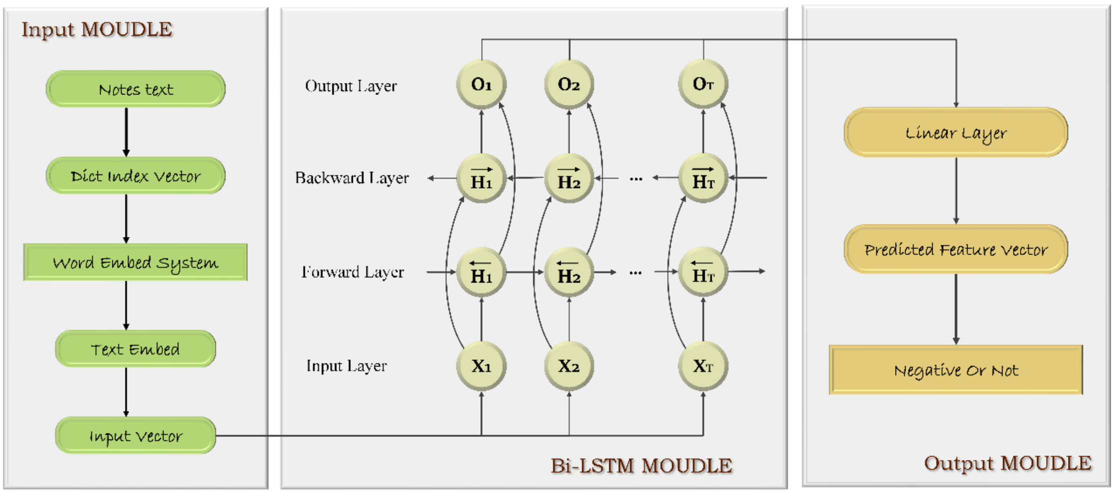

Figure 5. Text Sentiment Classification Model.

According to the above figure, text emotion processing includes three major parts: INPUT MOUDLE, LSTM MOUDLE, and Output MOUDLE.

In the LSTM model, the pre-trained word vectors are first loaded according to the current dictionary and index order. This article will use pre-trained word vectors and a bidirectional recurrent neural network with multiple hidden layers to determine whether a text sequence of variable length contains positive or negative emotions.

- *Input Moudle*

Data preprocessing, the process is as follows:

1. This article uses 2320 pieces of report information about the American Hornet as a training set and divides each piece of data with spaces, and keeps each word. Use the minimum number of occurrences of 5 as the standard to construct a vocabulary.

2. Create a dictionary and word index table, and convert the above data set text from string form to word subscript sequence form.

3. Put the above text in the embedding layer to reduce the dimension of the vector, and finally obtain the word vector of the LSTM model.

- *Bidirectional LSTM Moudle*

Because the LSTM model is unidirectional. Bidirectional LSTM is composed of two layers of LSTM. The input directions of the two layers are opposite. The newly added layer transfers the state to the previous element of the sequence step by step. See the BI-LSTM MOUDLE part in the figure above for details.

Two hidden layers, Backward Layer and Forward Layer, represent forward propagation and backward propagation, respectively. Finally, the output of the two layers is combined to get the final result.

- *Output Moudle*

Output Moudle means that a linear layer is placed on the basis of Bidirectional LSTM Moudle, and the predicted feature vector is obtained through the output of the linear layer, and then the subscript of a larger number is selected from it to obtain the output category, that is, the report is wrong.

- *Conclusion analysis*

This article uses 2091 reports about the American Hornet as a test set. First, the model results are tested through cross entropy. This article assumes that the probability distribution p is the expected output and the probability distribution q is the actual output. The test result can be as low as 0.0995. The smaller the value of cross entropy, the higher the accuracy of recognition.

Secondly, according to the ratio of the number of samples correctly reported to the total number of samples, this article calculates the accuracy rate of whether the forecast report is wrong, and the highest rate can reach 0.742.

3.2.2.2 CNN Model

The specific research process is as follows Fig.6:

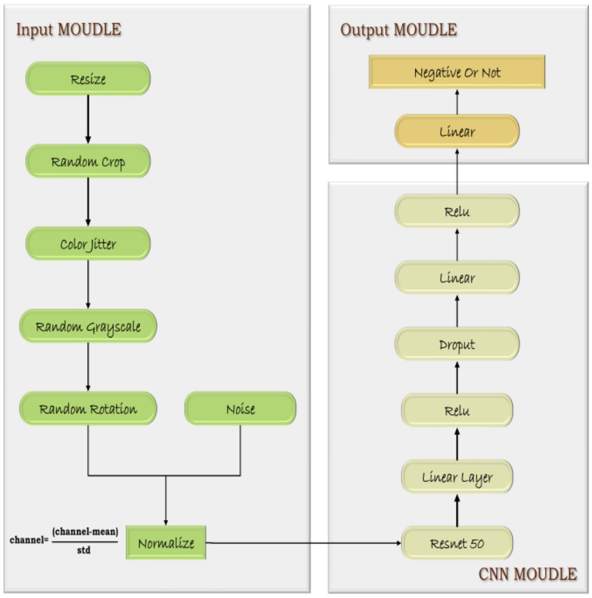

Figure 6: Image Recognition CNN Model

The model can be divided into three independent modules: Input Moudle, CNN module and Output Moudle.

- *Input Moudle*

Based on 4375 wasp images, this paper first preprocesses the image data, and divides the images into two categories: correct images and wrong images.

Put the above preprocessed data in the Input Module, and convert the image to the same capacity, and then enhance the image, including Random Crop, Color jitter, Raudom Grayscale, Rawdow Rotation, etc. On this basis, noise processing is added to the image. Image noise is an important enhancement step that enables the model to better separate the signal and noise in the image. Finally, for image standardization, the processing formula is as follows:

$$\text{channel} = \frac{\text{channel} - \text{mean}}{std} \tag{11}$$

The standardized data will be used as the input data of the CNN Module.

- *CNN Module*

The CNN model in this article uses the resnet50 model. The structure diagram is as follows Fig.7:

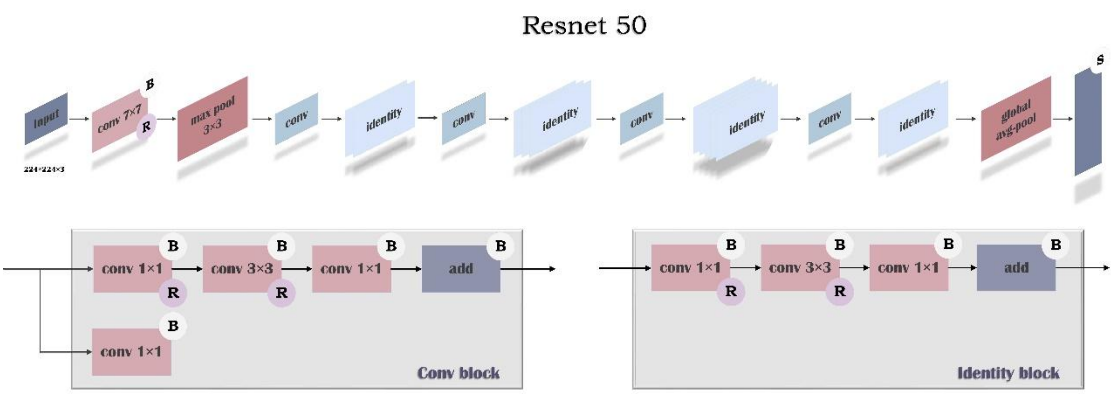

Figure 7. Resnet 50

The Resnet model proposes to use residual learning to solve the degradation problem. This model is a stacked layer structure (multi-layer stacking). When x is input, the acquired feature is recorded as H(x). This article hopes to obtain the residual F(x) = H(x)−x, so the original learning feature is actually F(x) + x. When the residual is 0, the accumulation layer only performs identity mapping at this time, the network performance will not decrease, and the actual residual error will not be 0, so that the accumulation layer can obtain new features after the input feature record, thus Have better performance. The structure of residual learning is shown in the figure below.

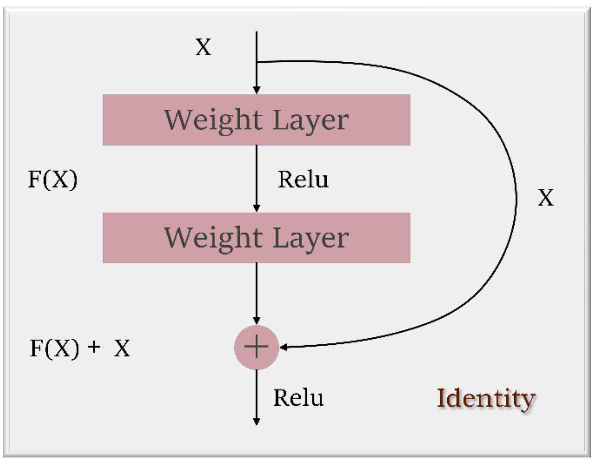

Figure 8. The structure of residual learning

The model used in this article has two linear layers behind resnet50. Between the linear layer and the linear layer, this paper uses the RELU activation function to transform the data in order to increase the nonlinearity of the model. The functional formula of RELU is as follows:

$$f(x) = max(0,x) \qquad (12))$$

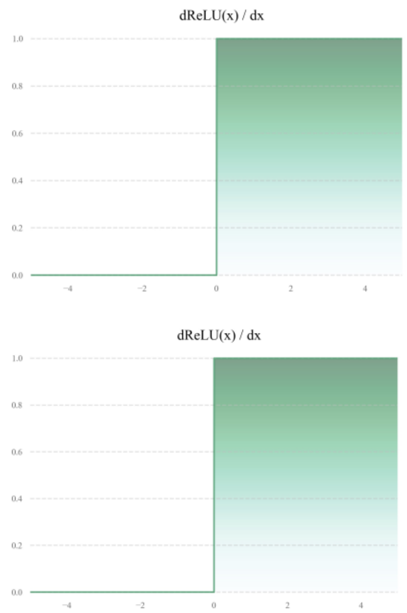

Figure 9: ReLU and dReLU(x)/dx

In the above training process, this article uses droupout. Dropout is to temporarily discard the neural network unit from the network according to a certain probability during the training process of the deep learning network. Note that for the time being, for stochastic gradient descent, because it is randomly discarded, each mini-batch is training a different network. The droupout can prevent CNN Moudle from overfitting.

- *Output Moudle*

In order to obtain the dimension of the output result, Output Moudle uses a linear layer to convert the output result of CNN Moudle into the data dimension, that is, whether the report data is wrong.

- *Result analysis*

The output module of this paper uses a linear layer to convert the output result of the CNN model into the data dimension, that is, whether the report data is wrong. This paper randomly selects 1000 photos as the test set, and the rest of the data as the training set. Using the model training set data to predict the test set, the accuracy rate can reach 0.943.

3.2.2.3 Classification model

First of all, this article queries its altitude, temperature, humidity, precipitation, wind speed, air pressure, cloud cover and average solar radiation based on the recorded data of the report.

Then, according to the text sentiment classification model, each report record data is predicted and classified, and the classification result is added to the above data as a new feature sentiment list.

In addition, this article uses the CNN model to predict and classify based on the bumblebee image data.

Finally, through the xgboost model, the above model and data are combined, and 3000 pieces of data are used as the training set to train this model, and finally compared with 1386 test set data, the accuracy can reach 0.956.

4. Conclusion

This article predicts the spread of Asian Hornet in the United States, and at the same time, proposes a model method to effectively identify Asian Hornet. However, there are still shortcomings. The life cycle of bees, seasonality of activities and other factors are not considered; due to limited effective data, the prediction model may not be able to reasonably fit the living environment of the Asian Hornet.

References

[1] Tony Stankus. Reviews of Science for Science Librarians: “Murder Hornets:” Vespa Mandarinia Japonica, Science & Technology Libraries, 2020. DOI: 10.1080/0194262X.2020.1796890

[2] Nuñez-Penichet C, Osorio-Olvera L, Gonzalez VH, Cobos ME, Jiménez L, DeRaad DA, Alkishe A, Contreras-Dáz RG, Nava-Bolaños A, Utsumi K, Ashraf U, Adeboje A, Peterson AT, Soberon J. 2021. Geographic potential of the world’s largest hornet, Vespa mandarinia Smith (Hymenoptera: Vespidae), worldwide and particularly in North America. PeerJ 9:e10690 DOI 10.7717/peerj.10690

[3] Erik D. Norderud 1, 2, Scott L. Powell 2, Robert K. D. Peterson 2. Risk assessment for the establishment of the Asian giant hornet (Vespa mandarinia) in the Pacific Northwest, DOI: https:/doi. org/10.1101/2021.02.01.429186.

[4] Koji Hiranoa Atsuhiro Tanikawab. Ocular Injury Caused by the Sprayed Venom of the Asian Giant Hornet (Vespa mandarinia). S. Karger AG, Basel. DOI: 10.1159/000508911O.

[5] Alberto Jalaniz, Mario A Carvajal. Giants are coming? Predicting the potential spread and impacts of the giant Asian hornet (Vespa mandarinia, Hymenoptera: Vespidae) in the United States. Pest Manag Sci, 2020. DOI 10.1002/ps.6063.

[6] United States Department of .Agriculture Asian Giant Hornet Control Program in Washington State.2020.

[7] David A.Moo-Llanes. Inferring Distributional Shifts of Asian Giant Hornet Vespa mandarinia Smith in Climate Change Scenarios. Neotropical Entomology.DOI: 10.1007/s13744-020-00840-4.

[8] Zhou Jian, Gao Jing, Zhou Yang Wenqian. Spatial econometric model setting theory and its new progress. Journal of economics, 2016, 3(2): 161-190.

[9]https://baike.baidu.com/item/%E8%83%A1%E8%9C%82/37732?fr=aladdin

[10]Anselin, Luc, and Anil K. Bera. Spatial Dependence in Linear Regression Models With an Introduction to Spatial Econometrics. New York: Marcel Dekker: 237289, 1998.

[11]Anselin, Luc. An Introduction to Spatial Regression

Analysis in R. New York: Marcel Dekker: 237289, 2003.
[12]Goodfellow, I., Bengio, Y., Courville, A.. Deep learning (Vol. 1): Cambridge: MIT Press, 2016: 367-415.
[13]Schmidhuber, J., 2015. Deep learning in neural networks: An overview. Neural networks, 61, pp.85-117.
[14]Ng, A., Kian, K. and Younes, B., Sequence Models, Deep learning. Coursera and deeplearning.ai. 2018
[15] Goodfellow, I., Bengio, Y., Courville, A. Deep learning (Vol. 1). Cambridge: MIT press, 2016: 326-366.
[16]Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Wang, L., Wang, G. and Cai, J., 2015. Recent advances in convolutional neural networks. arXiv preprint arXiv:1512.07108.
[17]Gengping Zhu, Javier Gutierrez Illan, Chris Looney, and David W. Crowder. Assessing the ecological niche and invasion potential of the Asian giant hornet. Proceedings of the National Academy of Sciences, 117(40):24646–24648, October 2020. Publisher: National Academy of Sciences Section: Biological Sciences.
[18]J.K. Moran Cliff, A.D. and Ord. Spatial processes: Models and applications. Pion Limited, London, 1981.