Problem Chosen	2021 MCM/ICM Summary Sheet	Team Control Number
C		2105921

Summary

Asian giant hornets are the largest wasp species in the world, and predators of European honeybees, invading and destroying their nests. In this paper, we divide the problem into five aspects, and build a mathematical model to solve each problem.

For question 1, according to the hornets and other local bee predator - prey, and to consider the spatial distribution of population and hornets population has a propensity to migration phenomenon, in Loltka - on the model on the basis of the revised, joined the proliferation and migration, changed the equation for staggered reaction diffusion equations with hornets and distribution data as the initial condition of prey, and its numerical solution using the Euler's method, get the hornets after the invasion of reproduction in different area, diffusion, the relationship between migration change over time.

For question 2, we build a classification model based on Linear Regression and Support Vector Machine Model to process reports with unprocessed and unverified status. At the same time, this paper establishes the YOLO V5 bumblebee image detection model to classify the images and identify whether they are bumblebees or not.

For question 3, through the analysis of the model output, we can consider whether the report is positive and related to its location. Image detection can also be carried out on the uploaded pictures by sending them into the YOLO V5 model trained in the second question, and comprehensively assessing whether the pictures are Asian gaint hornets based on the geographical distribution of longitude and latitude.

For question 4, according to the problem of one equation described the characteristics of the nonlinear system, due to the initial value could not be completely accurate and in the process of calculation error accumulation, system will appear the phenomenon of chaos in the numerical simulation of a long period of time, for the problems need to be updated on the solution of the model, according to a new report to change the initial conditions for new simulation.

For question 5, because the hornets invasion of intervention is applied in the process of human, so the problem of a model on the basis of the equations to join elimination, and again, to simulate the activities of the swarm calculation shows the Washington area hornets population density is zero, the result can be used as the Washington area has been one of destroy the evidence of the hornets.

Keywords: Loltka-Volterra model, Reaction diffusion equation, Nonlinear system, Euler's method, YOLO v5

Content

1 Introduction

1.1 Background

Asian Giant hornets, which are also called Vespa mandarinia, are considered a potentially invasive wasp from eastern Asia, a colony was found in the fall of 2019 on Vancouver Island then destroyed. Based on previous research data, Asian Giant hornet is a predator of European honeybees. In the slaughter phase, a bunch of 20-30 Asian Giant hornets can kill 5000-25000 bees within a few hours, which caused hidden severe impact of danger on beekeeping industry in the area where the Asian Giant hornet is established. Moreover, Asian Giant hornets are large and deliver propotionally large volumn of venom per string. Every year 50 people are killed due to allergic reactions caused by stings. Although the damage caused by Asian Giant hornets to the beekeeping industry and the human body has not been clearly studied, if necessary, North America needs to take measures to deal with the invasion of Asian Giant hornets.

Helplines and website are built by the State of Washington for people to upload their observation materials of hornets. Some reports were determined to be Vespa mandarinia, while others turned out to be other species of insects. Due to the limited resources of government agencies, and the convenience to answer public questions while the effect of invasive Asian Giant hornets is not negligible, the team has to adopt certain strategies to prioritize investigation of the reports most likely to be positive sightings.

1.2 Problem Restatement

Most of the public's sighting reports were misidentifications. The images, statistics and government judgments in the reports provide reliable data. In order to meet the needs of local agriculture and ecological protection, the modeling team needs to use these data to figure out how to deal with Asian Giant hornets. The problem is that there may be some rules in the spread of Asian Giant hornets, or there may affect the livelihood of other species of bees.

(1) Discuss whether it is possible to predict the spread of this pest over time and with what accuracy.

(2) Use the data provided to create, analyse a model and discuss the likelihood of a mistaken classification.

(3) Predict which reports may be determined positive base on the created model.

(4) As time goes by, the forecast model needs to be updated at a certain frequency, discuss how often should it occur.

(5) Based on the model's predictions, discuss what signs that Washington State has eliminated the Asian Giant hornets.Write a two-page memorandum for summarizing results for Washington State Department of Agriculture.

2 Problem Analysis

2.1 Data Processing

The contest provides files in various formats such as numerical data, videos, and images, which would be used by the modeling team to make predictions on the recognition and spread of the Asian Giant hornets. For these data, different methods to deal with should be taken, and then combine them together for the judgment of public sighting reports and the prediction of the follow-up vespa mandarinia invasion and spread.

(1) Processing of numerical data

The numerical data contains two table files, import them into the database, and associate them with the same primary key. In fact, a record corresponds to the information of a public's sighting report. The locations of these reports are basically in the area of Washington State, and they also contain the judgment information of the local government.

We can infer that the Asian Giant hornets's invasion area is still in the border area of Washington State. In order to facilitate subsequent processing, the reports can be classified according to the judgment of the National Agricultural Department. They are: the report identified as the Asian Giant hornet, and the report identified as the Asian Giant hornet, reports that have not been judged, and reports that have not yet been processed. For unprocessed reports and unverified reports, data prediction can be used to evaluate the status of the report. According to preliminary statistics, the proportion of various reports in the data is as follows:

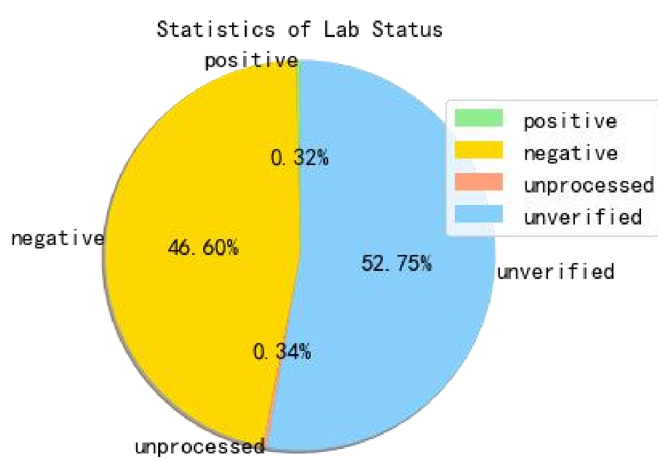


Figure 1 Statistics of Lab Status

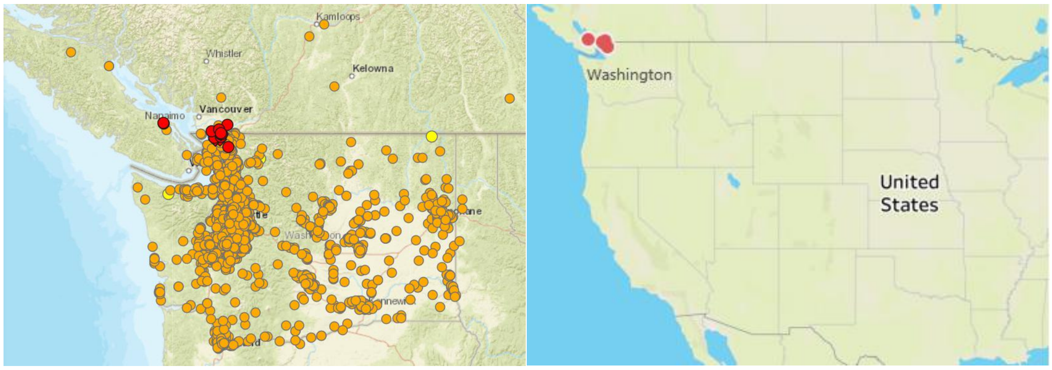


Figure 2 Distribution of Sighting Reports

The distribution of various samples is very uneven, and the number of positive reports is very small. It is likely that the Asian Giant hornets initially appeared in North America, which provides the possibility for the Washington state to eliminate Asian Hornets. This also shows that if we need to use the limited resources of the state to deal with people's sighting reports, positive and negative records will both play an important role.

(2) Processing of data in other formats

Through images, videos and other formats of data, the team can intuitively understand the morphological signs of Asian Giant Hornets. Most of the reasons for misclassification are due to the misunderstanding of the public. In terms of image processing, digital image processing methods can be used to process photographic photos of Asian Giant Hornets and other species of insects, and analyze the characteristics of Asian Giant Hornets that are different from other bees or insects in size and color. In addition, the use of deep learning image recognition can effectively extract the image features of the Asian Giant Hornets, establish a special recognition model, and make judgments on the image materials in the public's sighting report. Similarly, the proportion of these data is also very uneven, limited by the number of samples, it is necessary to additionally introduce other pictures confirmed as Asian Giant Hornets as complement of training data.

2.2 The Living Habit of Vespa Mandarinia

The living habits of the Asian giant hornet can provide a basis for the prediction of its spread. Like other bees, the Asian giant hornets are an annual species that build their new nests yearly. It appears outside the nest only in the spring before the appearance of workers or before hibernation. The most obvious physical characteristics are its eye-catching yellow head, black chest, yellow and black striped abdomens, and large in size. In the case of clear shooting or observation, there are still some similar species that are quite difficult to distinguish.

Figure 3 Asian Giant Hornet (Photograph by the Washington State Department of Agriculture via Flickr)

Asian giant hornets are natural enemies of other insects, and they also have predation relationships with other types of honeybees. This may be related to the decrease in the number of agricultural insects after the local Asian giant hornets invaded, causing the prey of Asian giant hornets to change. In the slaughter phase, multiple Asian giant hornets can kill almost a thousand times the number of bees, and only the creatures far away from their nests will not be disturbed. Due to the toxicity of the Asian giant hornets and the greater intensity of killing other organisms, we can think that the Asian giant hornets in the state of Washington where it invades has only natural factors or human culling that can cause death, and there are no other biological enemies.

2.3 Loltka-Volterra Model of Hornets Propagation Over Time

In order to model wasp dispersal over time, it is necessary to understand the factors that affect wasp migration and their interaction. First, hornets act as predators in the ecosystem, feeding primarily on other species of bees (including other insects). As an invasive species, hornets in the United States have few natural enemies other than humans. When hornets invade, they immediately form a predator-prey relationship with other local bees and insects. Therefore, the Loltka-Volterra model can be used to describe the relationship between the number of hornets and prey over time without considering the dispersal and migration of hornets and prey. But in fact, hornets and prey exists the phenomenon of diffusion, and because the hornets have outstanding learning and memory ability makes them to be able to place and food distribution has certain memories, that is to say, hornets have the ability to track its prey, it lead to it can according to the distribution of prey to a certain extent, the orientation of migration. Therefore, the Loltka-Volterra model can be modified by adding diffusion term and migration term, and the equation can be rewritten into a staggered reaction diffusion equation. The distribution data of hornets and predators can be used as the initial conditions, and the numerical solution can be obtained by using Euler method.

2.4 Asian Giant Hornets Image Detection Model Based on YOLO v5

The model uses YOLO v5 to generate target detection data sets on Asian bumblebee by identifying positive and negative data provided, as well as human recognition and automatic machine learning on unprocessed and unverified data. Bumblebee can be identified from images by learning its features through target detection.

3 Model Hypothesis

In this context, the predator refers to hornets , and the prey refers to other local bees and some insects.

1. If there are no predators, the intrinsic rate of prey population increase is proportional to the number of prey.

2. If there is no prey, the intrinsic rate of decrease in the number of predators is proportional to the number of predators.

3. If prey and predator coexist in the same environment, the rate of prey population decrease and the rate of predator population increase due to predation are proportional to the frequency of prey and predator encounter.

4. The frequency of encounters between prey and predator is proportional to the number of prey and predator.

5. Both predators and prey have the phenomenon of diffusion and migration. Predators tend to migrate to areas where the number of prey is more distributed.

4 Model Establishment and Solutions

4.1 Task1:Analyze the Spread of Asian Giant Hornets

4.1.1 Model Establishment

Hornets act as predators in ecosystems, feeding mainly on other species of bees (including other insects). As an invasive species, hornets in the United States have few natural enemies other than humans. When hornets invade, they immediately form a predator-prey relationship with other local bees and insects. Beretta pointed out that because of the widespread existence of predator-prey relationships, the dynamic relationship between predators and their prey has been and will continue to be an important research topic in ecology and mathematical ecology. The predator-prey model derived from the classic work of Lotka and Volterra is as follows:

$$\begin{cases} \frac{d}{dt}U = \Gamma U\left(R - \frac{R}{\kappa}U - \gamma_{12}V\right) \\ \frac{d}{dt}V = \Gamma V(-M + \gamma_{21}U) \end{cases} \quad (1)$$

Among them, the variables $U(\tau)$ said the prey population density at τ moments, $V(t)$ has said the predator population density at τ moments, $\Gamma, R, \kappa, \gamma_{12}, \gamma_{21}$ coefficient for life, and all is normal. After reduction, Model (1) becomes:

$$\begin{cases} \frac{d}{dt}u = \Gamma u(r - \gamma u - v) \\ \frac{d}{dt}v = \Gamma v(-1 + u) \end{cases} \quad (2)$$

Among them:

$$u(t) = \frac{\gamma_{21}}{M}U(t), v(t) = u(t) = \frac{\gamma_{12}}{M}V(t), t = M\tau, r = \frac{R}{M}, \gamma = \frac{R}{\kappa\gamma_{21}}$$

Diffusion is common in ecosystems due to the migration of species. There are general linear self-diffusion, nonlinear self-diffusion and staggered diffusion among species, and this law is also applicable to hornets and their prey. Model (2) does not take into account the heterogeneity of the distribution of the population within its environment, the migration in and out of the population, and the heterogeneity of the resources allocated to each individual. Therefore, the population in an ecosystem not only evolves with time, but also changes in spatial direction, which can be described by the reaction-diffusion equations. The semilinear reaction-diffusion model corresponding to Model (2) is:

$$\begin{cases} \partial_t u - d_1\nabla^2 u = \Gamma u(r - \gamma u - v),\ x \in \Omega, t > 0 \\ \partial_t v - d_2\nabla^2 v = \Gamma v(-1 + u),\ x \in \Omega, t > 0 \\ \frac{\partial u}{\partial n} = \frac{\partial v}{\partial n} = 0,\ x \in \partial\Omega, t > 0 \\ u(x, 0) = u_0(x), v(x,0) = v_0(x),\ x \in \Omega \end{cases} \quad (3)$$

Including Ω is bounded area boundary is smooth, n for boundary partial $\partial\Omega$ outside the unit normal vector of $d_1 > 0$ and $d_2 > 0$ is known as diffusion coefficient, respectively two species by high-density area to the low density area migration, $u_0(x)$ and $v_0(x)$ is a continuous function with zero negative is not constant.

Population growth rule, often depends on both time and space factors, also depends on the species and the role of another species or population interaction between internal: because the hornets has more excellent learning and memory ability makes them to be able to place and food distribution has certain memories, that is to say, hornets have the ability to track its prey, it lead to it can according to the distribution of prey to a certain extent, the orientation of migration. The phenomenon that predators tend to migrate to areas where the prey population is more distributed can be described by the staggered diffusion model. The staggered diffusion model corresponding to Model (3) is as follows:

$$\begin{cases} \dfrac{\partial u(x,y,t)}{\partial t} - d_1\nabla^2 u(x,y,t) = \Gamma u(x,y,t)[r - \gamma u(x,y,t) - v],\ x \in \Omega, t > 0 \\ \dfrac{\partial v(x,y,t)}{\partial t} - d_1\nabla^2 v(x,y,t) - d_{21}\nabla^2 u(x,y,t) = \Gamma v(x,y,t)[-1 - v(x,y,t)],\ x \in \Omega, t > 0 \\ \dfrac{\partial u(x,y,t)}{\partial n} = \dfrac{\partial v(x,y,t)}{\partial n} = 0,\ x \in \partial\Omega, t > 0 \\ u(x, 0) = u_0(x), v(x,0) = v_0(x),\ x \in \Omega \end{cases} \tag{4}$$

Where, d_{21} is the staggered diffusion coefficient, which is used to describe the tendency migration behavior of the population.

Model (4) can be used to describe the reproduction, diffusion, migration and the relationship between hornets and prey in a certain area after hornets invasion with given initial conditions. In order to determine the initial conditions, it is assumed that the density of hornet population and prey in a certain area is proportional to the reported number, and the distribution location is the same as the reported location, so the initial conditions can be obtained according to the topic data. In summary, the definite solution problem of the partial differential equations is formed.

4.1.2 Model Solving

Euler method is one of the most commonly used methods to solve the numerical solution of partial differential equations. The core idea of Euler's method is: using difference approximation to replace differential, the differential equation is discretized into difference equation, and then cycle iteration. The main steps are:

(1) Discretization: The system region is meshing first, and the continuous variation region of independent variables is replaced by a set of finite discrete points (grid points), and then the equations of continuous variables in the problem are replaced by the equations of discrete variables defined on the grid points.

(2) Solving the difference scheme: by replacing the partial differential with the difference of the partial differential equation on the grid point, the definite solution problem of the partial differential equation containing continuous variables is transformed into algebraic equations containing only a finite number of unknowns, namely the difference scheme.

(3) Cyclic iteration: the difference scheme is iterated several times in x, y and t dimensions to obtain the numerical solution of the original partial differential equation.

From the boundary value conditions, the partial differential equation solving area as the study area $\Omega : 0 \le x \le L_1, 0 \le y \le L_2$. Ω isometric partitioning into $N_1 \times N_2$ grid, get $N_1 \times N_2$

(x_i,y_i), among them, $i = 1,\ldots,N_1; j = 1,\ldots,N_2$. The step sizes of x_i and y_i are $h = \frac{L}{N}$, respectively. The time step is τ.

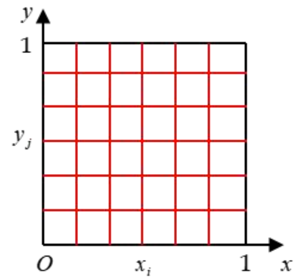


Figure 4 Schematic diagram of grid division

According to the idea of difference instead of differentiation in Euler's method, we write the first partial derivatives of the variable u with respect to t, u with respect to x and u with respect to y, and the second partial derivatives of x and u with respect to y in explicit difference form, where $O(\tau)$, $O(h)$ and $O(h^2)$stand for truncation error:

$$\frac{\partial u(x_i,y_j)}{\partial t} = \frac{u_{n+1}(x_i,y_j) - u_n(x_i,y_j)}{\tau} + O(\tau)$$

$$\frac{\partial v(x_i,y_j)}{\partial t} = \frac{v_{n+1}(x_i,y_j) - v_n(x_i,y_j)}{\tau} + O(\tau)$$

$$\frac{\partial^2 u(x_i,y_j)}{\partial x^2} = \frac{u(x_{i+1},y_j) - 2u(x_i,y_j) + u(x_{i-1},y_j)}{h^2} + O(h^2)$$

$$\frac{\partial^2 u(x_i,y_j)}{\partial y^2} = \frac{u(x_i,y_{j+1}) - 2u(x_i,y_j) + u(x_i,y_{j-1})}{h^2} + O(h^2)$$

$$\frac{\partial^2 v(x_i,y_j)}{\partial x^2} = \frac{v(x_{i+1},y_j) - 2v(x_i,y_j) + v(x_{i-1},y_j)}{h^2} + O(h^2)$$

$$\frac{\partial^2 v(x_i,y_j)}{\partial y^2} = \frac{v(x_i,y_{j+1}) - 2v(x_i,y_j) + v(x_i,y_{j-1})}{h^2} + O(h^2)$$

The above explicit difference scheme was put into Model (4) and the initial conditions were input. After cyclic iteration, the relationship of hornet reproduction, diffusion and migration in a certain area with time after invasion was obtained, as shown in the figure below:

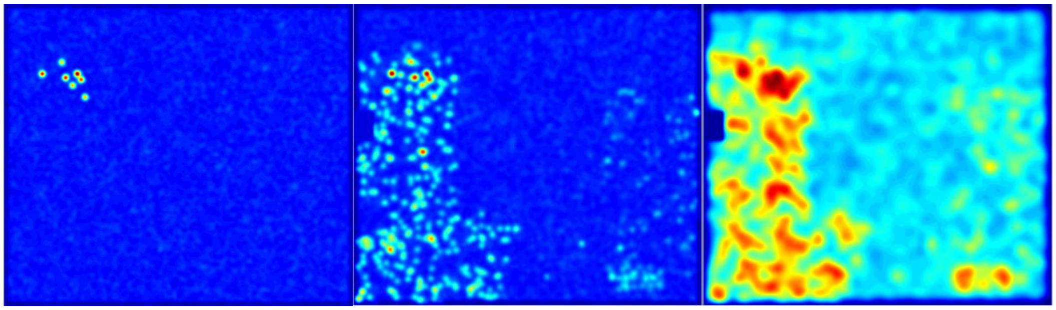


Figure 5 Changes in hornet population density

The color changes indicate hornet density:

According to the pattern, firstly, as a whole, it can be observed that the hornet population spread greatly with the increase of time, and mainly migrated to the south, but also had the trend of migrating to the east. Second, locally, the hornet population grows radially from the inside out, with the highest density in the center, which is the same as the actual situation -- the hornet's activity is always centered on the hive, and the queen is in the hive for the job of reproduction.

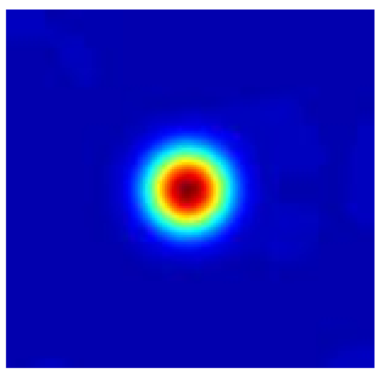


Figure 6 Radiant diffusion pattern of hornet population

In addition, using Euler's method to solve the partial differential equation, the truncation error is $O(h^2)$. With the increasing number of cycle iterations, the error will gradually increase. Therefore, this model is more suitable for the prediction of hornet population activity in a certain period of time, and for the prediction of a long time, the model needs to be updated, which will be mentioned in the fourth question.

4.2 Task2:Build a Classification Model

The dataset given in the question includes image data and numerical data, in order to make full use of them. The team will build models with numerical data and image data separately.

4.2.1 Build Predictive Models from Dataset Files

According to observations, the government-established website records the sighting reports submitted by the public in detail, including the time when the suspected Asian giant hornet was observed in the report, and the location where the photos, videos and other materials were filmed (located in the form of latitude and longitude) as shown in Figure 1. In terms of content, the number of positive reports is very small and concentrated on both sides of the Georgia Strait. There is no proliferation yet. We preliminarily infer that whether the status of the report is positive or negative is related to the place where the public submits the report. For example, the report formed on the border between Washington State and Canada is likely to be positive from the point of view of location and time. Numerical prediction methods can be used to test the correlation between location, time, and report status. Especially for classification tasks, two methods are common: linear prediction and SVM classification methods.

(1) Linear Regression Analysis

Linear regression is a statistical analysis method that uses regression analysis in mathematical statistics to determine the quantitative relationship between two or more variables. It is also the simplest and most basic model method in machine learning. Its expression form is $y = wx + b$. In the linear model prediction module of this article, all the columns of the dataset are set as independent variables x, normalized the data, and then modeled the data using the linear model of tensorflow.

The status of the report would be digitied before entering the model. Their equivalence is: positive ID=1, negative ID=2, unprocessed=3 and unverified=4. In this way, the report type can be judged by the model output value.

For the selection of the optimizer, Adam (Adaptive Moment Estimation) would be universal and reliable, Adam uses the first-order moment estimation and second-order moment estimation of the gradient to dynamically adjust the learning rate of each parameter. For the loss function, MSE (Mean Squared Error) would help to detect the deviation between the predicted value of the model and the true value.

From the training process in the figure below, after training about 100 times, loss tends to be stable, the final accuracy is around 92%.

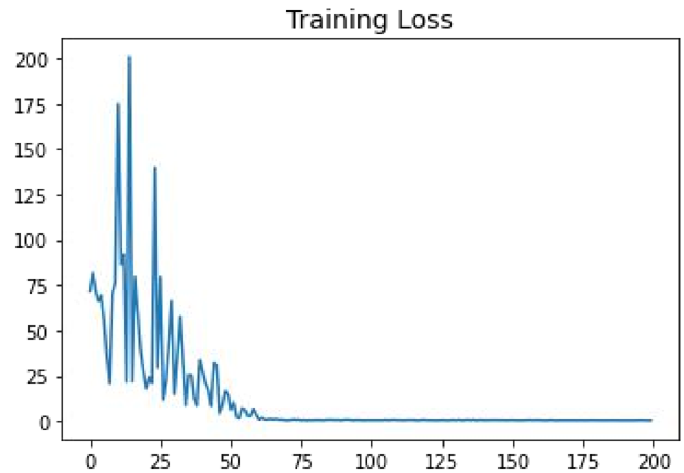


Figure 7 Change of Loss During Training

After the model training is completed, the report data with the status of unprocessed and unverified is input for judgment. Most reports are still judged as negative, and the number of positive reports is very small. Most reported sightings mistake other hornets or insects for the Asian giant hornets. The likelihood is about 98.739%.

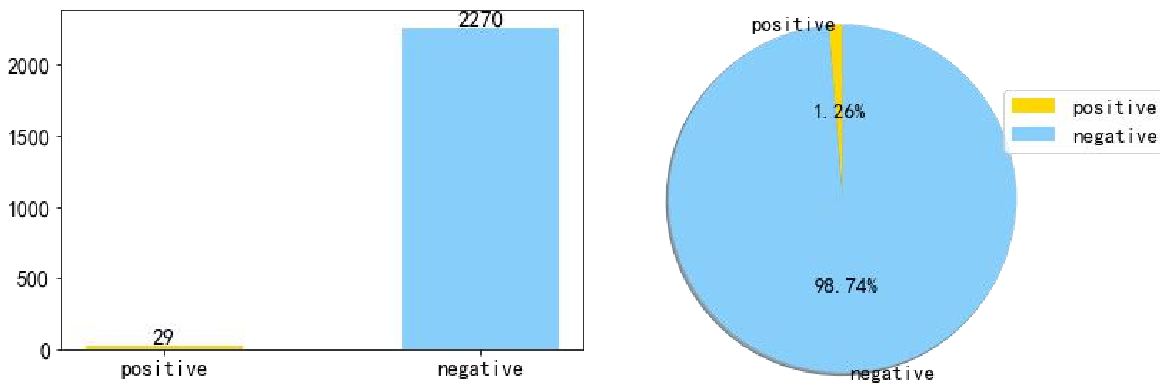


Figure 8 Statistics of Mistaken Classification

(2) SVM Classification Analysis

Support Vector Machine (Support Vector Machine) is a classifier that performs binary classification of data in a supervised learning method. Its decision boundary is the maximum margin hyperplane for solving the learning sample. In the report status classification of this article, we used existing positive or negative data for modeling.

First, perform One-Hot encoding on the multivariate data, and then normalize the data to divide the training set and test set. Finally, the Gaussian kernel function is used to model the data:

$$K(\vec{x}, \vec{l}^i) = e^{-\frac{\left\|\vec{x}-\vec{l}^i\right\|^2}{2\sigma^2}}$$

After a model of heart disease prediction was obtained. Below is the training result of the model:

Table 1 Trainning Result of SVM Model

	precision	recall	F1-score	support
positive	0.50	0.33	0.40	3
negative	1.00	1.00	1.00	515
accuracy			0.99	518
macro avg	0.75	0.67	0.70	518
weighted avg	0.99	0.99	0.99	518

Very similar to the results of linear regression prediction, almost all unprocessed and unverified reports in the test data are judged as negative, and the likelihood is quite high. In terms of dataset files, most reports are inherently negative. In the case of pure numerical data that is hard to enhance, it is necessary to additionally rely on image data submitted by the public.

4.2.2 Asian Giant Hornet Recognition Model

In the dataset, the only information that can be used as the linear prediction model and the SVM classification model is the latitude and longitude, the detection time of report and the status of the report. If the report can be classified by the image recognition method, because the Asian giant hornet has obvious morphological features, the image in the report can be used as the basis for judging whether the report status is positive with the help of the powerful feature capture ability of the image recognition model. We use the YOLO v5 model due to ies smaller size and superior performance.

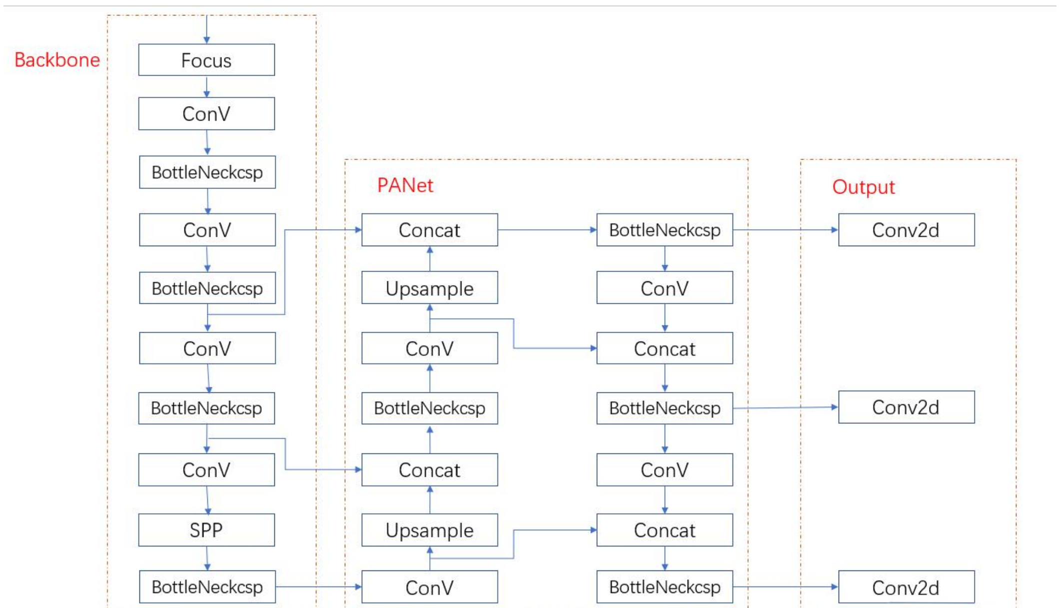
YOLO v5 flexibly configures models of different complexity by applying channel and layer control factors similar to EfficientNet, and adopts a cross-neighborhood grid matching strategy in the positive and negative sample definition stage, thereby obtaining more positive sample anchors and accelerating convergence.

This model uses YOLO v5, through the identification of the provided positive and nagetive data, as well as the human identification and automatic machine learning of unprocessed and unverified data, to produce an object detection data set on the Asian giant hornet.

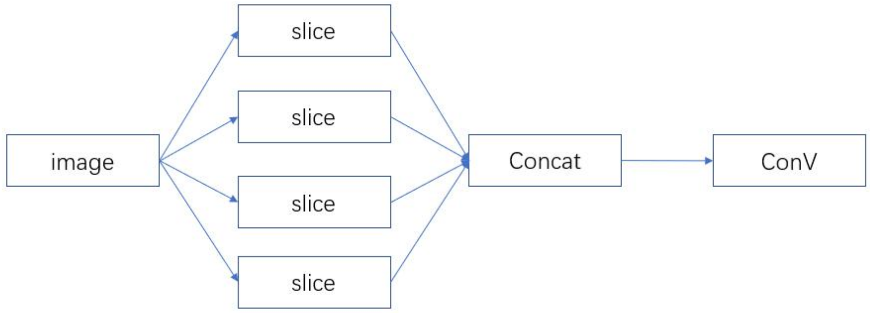
Provided images were labelled, and in order to reduce the unevenness of different sample amount, some images that have been confirmed as Asian giant hornet are added as supplements to ensure that the model can recognize Asian giant hornet in the public's sighting reports to a certain extent.

Figure 9 Overview of the Model's Training Dataset

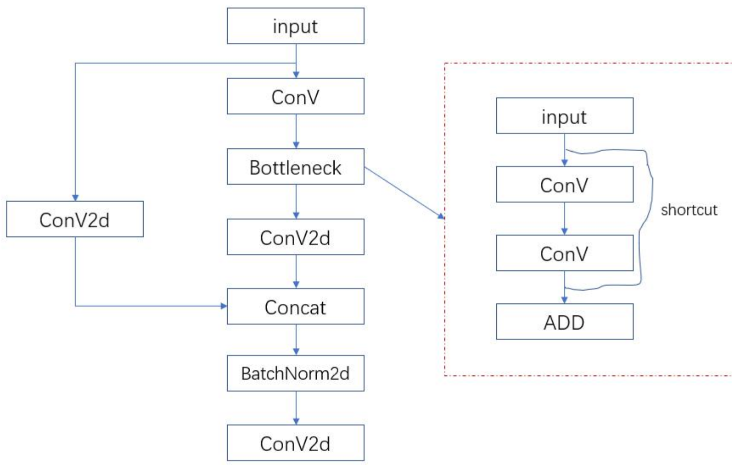
(1) The overall architecture of the YOLO v5 network model used is as follows：



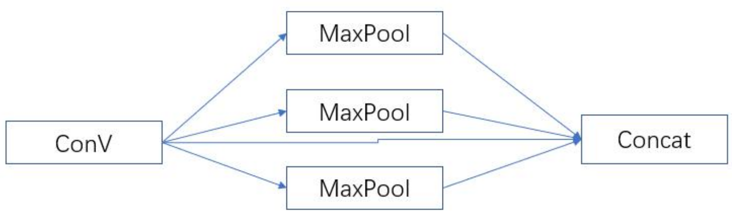
(2) Focus module：



(3) BottleneckCSP module：



(4) SSP module：



(5) Classification Model Configuration

For model tuning and training, the ratio of training set to validation set is 0.8:0.2, and 0.1 is extracted from it as the test set. The learning rate and network structure in the model are all the default YOLO v5 configuration：

① Adopt ’ warmup+cos lr’ learning rate strategy, no weight attenuation for bias.

② The gradient accumulation strategy commonly used in the yolo series is adopted, the batch size is increased, and the bias of the output head part is specially initialized.

③ Adopt class-balanced sampling strategy.

④ Multi-scale training, directly perform bilinear interpolation on batch images output by dataloader.

⑤ Use the ema strategy of model weight exponential moving average.

Table 2 Main Configuration of Training Process

parameter	Value	Parameter Meaning
batch-size	32	number of images in batch
epochs	100	number of training iterations
img-size	640, 640	Image size
workers	8	number of training cores
Specifications	CPU: 8 vCPUs 64 GIB GPU:1 x nvidia-v100-pcie-32gb 32GIB	CPU and GPU infomation
Compute Nodes	1	-

After training, the model predicts the images in the test set, and the overall accuracy rate is 88% through the test verification of 400 images. The following shows the prediction situation when using Internet search images. The model has certain discriminative ability. Using this model can reduce the processing pressure of the state government while analysing reports.

Figure 10 The Prediction Result of the Asian giant Hornet Image

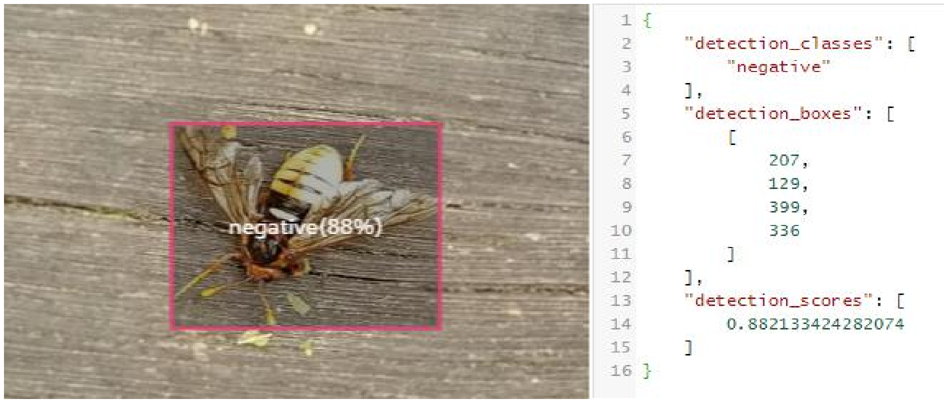


Figure 11 The Prediction Result of the Other Insect Image

It can be seen that through deep learning, this model can better assist in identifying the Asian gaint hornet and at the same time identify the area where the Asian giant hornet is located in the image. For a small number of misjudged and difficult-to-recognize iamges, the frequency of geographic latitude and longitude and manual recognition can be used to assist in the judgment.

4.3 Task3:Discuss What Leads to a Positive Report

4.3.1 Analysis From the Perspective of Classification Model

From the above classification model constructed using the dataset file, a model for judging the actual status of unprocessed and unverified reports can be obtained. After inputting these data into the model, the classification result produced. We plot these data separately on the map of Washington State, as shown below:

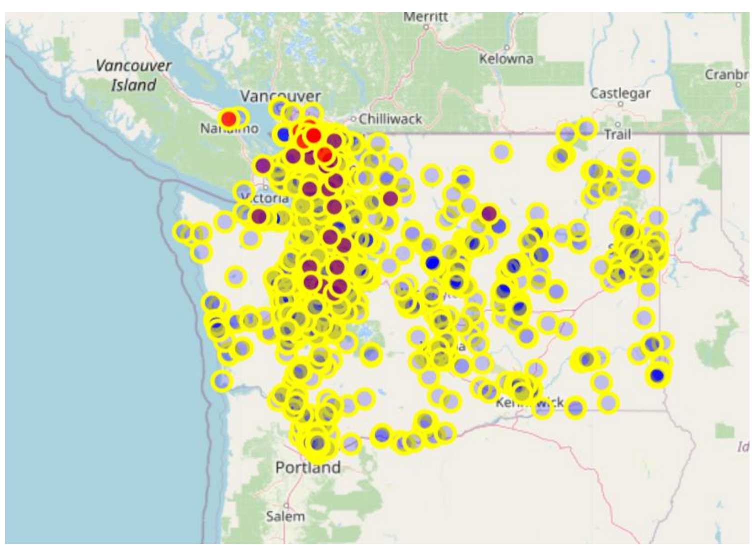


Figure 12 Map Visualization of Classification Results

It should be clearly noted that on the map above, the red dots represent the location of the report submitted with the original status as positive, the blue dots represent the location of the report submitted with the negative prediction result, and the dots shows purple mean that the report result is predicted to be positive. From the classification results alone, the purple dots are closer to the red dots in distance than most blue dots. We can infer that the report submission position is related to whether it is positive. The location information of these reports predicted positive is shown in Appendix 1.

4.3.2 Analysis From the Perspective of Detection Model

For image detection, we can carry out image detection on the uploaded images fed into the YOLO V5 model trained in Question 2.

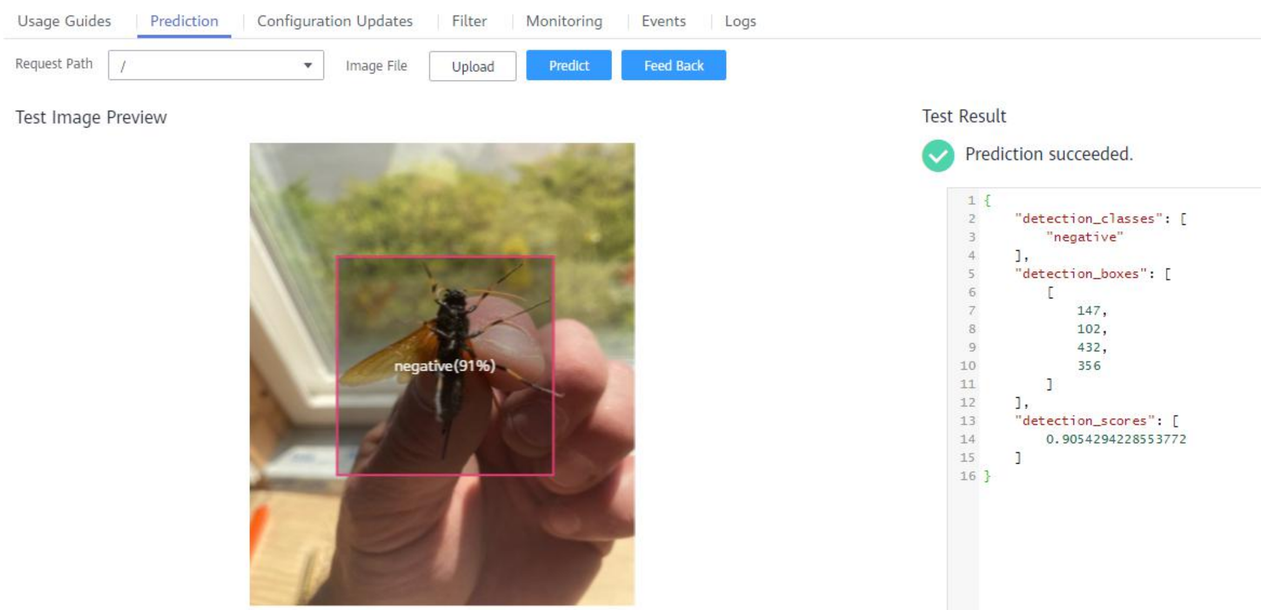


Figure 13 Model Online Test Effect

Through the preliminary judgment of the model, we can roughly classify the original data and comprehensively evaluate whether the picture is Asian bumblebee based on the geographical distribution of longitude and latitude.

4.4 Task4:Infer How Often the Model is Updated

Using Euler's method to solve the partial differential equation, the truncation error is $O(h^2)$. With the increasing number of cycle iterations, the error will gradually increase. In addition, because the measurement of initial data is impossible to be completely accurate, the system will appear chaos and other phenomena in a long time simulation:

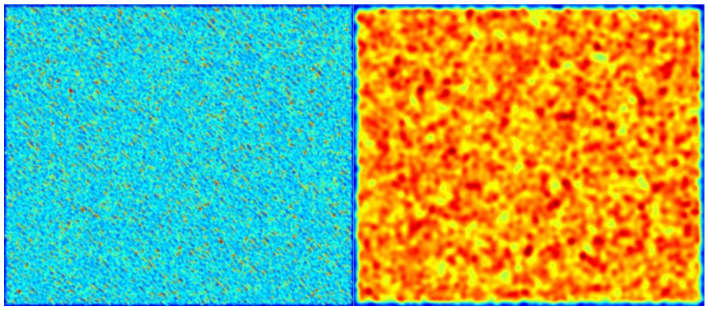


Figure 14 The System Produces Chaos

This is because the system is a nonlinear system, very sensitive to initial values, parameters and errors, the system is prone to large deviations in a long time simulation, so the model is more suitable for prediction in a short period of time. To predict hornet population density over a longer period of time, the model must be updated -- by changing the initial conditions of the system, which, according to the assumptions in the preceding paragraph, are determined by the given new reports. Therefore, when there is a large deviation between the calculation result and the actual report, the model should be updated, and the initial conditions should be updated according to the latest report before short-term simulation calculation.

For the YOLO v5 model, if there are other new reports over time, they should be added to the training set to re-train the model we built. This can make our image detection model more judgmental.

4.5 Task5:Vidence of the Eradication of Asian Giant Hornets

Rare in North America as the invasive species of hornets predators, in order to rein in wasp species breeding and spread, must be human intervention, therefore shall join in the evolution of the wasp model human intervention, namely elimination, assume that wasp species found in somewhere to eliminate the hornets population immediately after processing, to eliminate the speed is proportional to the found that the speed of the wasp species, and found that the speed of the hornets is proportional to the local wasps, population density, so the equation should join -elimination of dv, including constant d is greater than zero.

$$\begin{cases} \frac{\partial u(x,y,t)}{\partial t} - d_1\nabla^2 u(x,y,t) = \Gamma u(x,y,t)[r-\gamma u(x,y,t) - v],\ x\in\Omega, t>0 \\ \frac{\partial v(x,y,t)}{\partial t} - d_1\nabla^2 v(x,y,t) - d_{21}\nabla^2 u(x,y,t) = \Gamma v[-1-v] - dv,\ x\in\Omega, t>0 \\ \frac{\partial u(x,y,t)}{\partial n} = \frac{\partial v(x,y,t)}{\partial n} = 0,\ x\in\partial\Omega, t>0 \\ u(x,0) = u_0(x), v(x,0) = v_0(x),\ x\in\Omega \end{cases} \tag{4}$$

Using the same method as the first problem, the numerical solution of the equation can be obtained, that is, the relationship between the reproduction, diffusion and migration of hornets in a certain area with the intervention of human beings after the invasion with time.

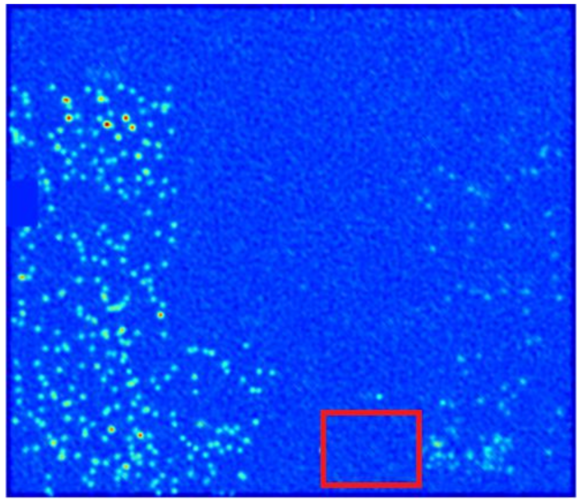


Figure 15 The Change of Asian Giant Hornet Population Under Human Intervention

The color changes indicate hornet density:

It can be seen that the spread of hornet populations was significantly limited after human intervention compared to the results in Question 1. The calculated hornet density in the Washington area (circled in the image above) is essentially zero, which can be taken as evidence that the area has been eradicated.

5 Evaluation and Extension of the Model

5.1 Advantages of the Model

(1) For the report recognition model in Task 2:

The model can detect a more direct relationship between variables. Since the indicators of the dataset file are not very complicated, it is more reliable to use a linear regression model.

(2) For the Asian Giant Hornet recognition model based on YOLO v5:

YOLO v5 has fast object recognition speed, small model parameter file and can maintain high accuracy.Easy to configure the environment, model training is also very fast, and batch inference produces real-time results. It’s possible to retrain the model with the update of public’s sighting reports.

5.2 Disadvantages of the Model

(1) For the report recognition model in Task 2:

Since we balance the scale gaps of different parameters when forecasting, we have adopted the method of data standardization, but this has a certain impact on the numerical relationship of latitude, longitude, and detection time, and cannot reflect the original size of the data. We have not studied how this change will affect the model accuracy.

(2) For the Asian Giant Hornet recognition model based on YOLO v5:

Due to the complex background in the images and mutiple species of insects, the classification model can't guarantee a very outstanding recognition accuracy rate.The detection accuracy is not higher than its earlier version while enhancing its speed.

5.3 Improvement of the Model

① For the propagation model and classification model, the data of the local species population can be introduced in the calculation, which will be more in line with the objective situation.
② For the Asian Hornet recognition model, use data enhancement methods to expand the dataset samples on the existing Asian giant hornets image. If time permits, increasing the number of training sessions will help to increase the accuracy rate.

6 Reference

[1] Zhang Siqing, Liu Yi, Liu Yiran, Shao Dongdong, Sun Limin, Zheng Shaoyan. Dynamic simulation of Spartina alterniflora population diffusion in the Yellow River Delta based on cellular automata[J/OL]. Journal of Beijing Normal University (Natural Science Edition): 1 - 8[2021-02-09].http://kns.cnki.net/kcms/detail/11.1991.N.20210120.1101.002.html.

[2]Wang Bingbing. Research on the spread and diffusion model of alien organisms[D]. Zhejiang University of Technology, 2018.

MEMO

To: Washington State Department of Agriculture

From: Modeling Team

Date: 2021/2/7

Theme: Brief Reasearch on Invasive Asian Giant Hornet

1. background:

Asian gaint hornets, also called Vespa mandarinia, was found in the fall of 2019 in Washington State and then a colony was destoryed. They are quite big in size and strikingly colored, and they are not good things for local species.

The government set up a helpline and a website for the public to upload their sighting reports. With data provided, we found that in the past year, most of the locations judged to be Asian giant bees are in northwestern Washington state. Due to the limited resource to follow up with additional investigation, build a model for prediction and anslysis.

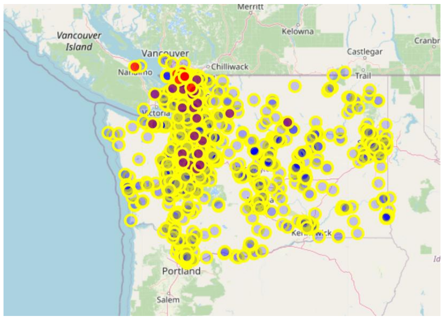
2. Our Research:

We have analyzed the living habits of the Asian giant hornets. The natural enemies of the Asian giant hornets in North America are not known in the current news reports and introduction information. And the species that Asian giant hornets prey were other insects at first, and then they turned to bees with no clear reason. In some cases, a few Asian giant hornets can attack hundreds of times as many bees. Because the bite of the Asian giant hornets delivers more venom, it is also harmful to the human body. This poses a great threat to the local beekeeping industry and the health of residents.

According to the report data, the distribution locations of some Asian giant hornets are clear. The Washington State government can detect nests in these areas and then use some means to eliminate them. As for the report that has not yet been processed, there may be other Asian giant bees that have not been discovered by the residents, or because of other climate and time factors, these giant bees did not move out of their nests.

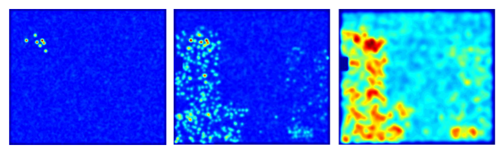
Our primary research is to detect whether there may be Asian giant hornets in sighting reports in other regions, so as to facilitate the government or relevant departments to detect Asian giant hornets nests in these areas. For the dataset file, we adopt the linear prediction model and the SVM classification model, and their test accuracy is higher than 90%. The linear prediction model is used to process the unprocessed and unverified reports, and it is still near the area where the positive report originally appeared. There are some reports that are predicted to be positive. They are shown on the map of Washington

State as shown below. We use purple dots to indicate possible locations of Asian giant hornets.

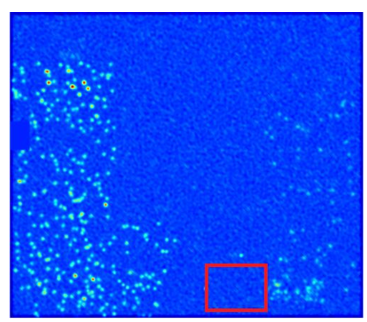


Considering that the website may increase the function such as automatic annotation in the future, and at the same time to reduce the workload of related staff, we have trained a model specifically for identifying Asian giant hornets. The YOLO v5 model used has the advantages of fast training and rapid image processing, but due to time constraints and data sample problems, the model does not have really perfect performance. But the model has a high accuracy rate, reaching 88%. The classification test effect for pictures on the Internet is also not bad.

In addition to detecting the presence of Asian giant hornets in the current report, it is necessary to predict the spread of Asian giant hornets. We regard other species of insects or bees appearing in reports as prey, and Asian giant hornets as predators to view their relationship. Since they all have proliferation, we use the Loltka-Volterra model which can describe this dynamic relationship. Taking the longitude, latitude, detection time and other information in the dataset file as the initial conditions of the input, using Euler's method to find the numerical solution of the partial differential equation, through loop iterations, we find the changes of the Asian giant hornets roughly multiplying, spreading and migrating in the Washington State area, the color represents the quantity :



We have added some factors to the above equation, which means that after the public submits the report, if it is determined that there are Asian giant hornets in the area, relevant departments will take measures to eliminate giant hornets and their nests, and the population of Asian giant bees will drop significantly.



7 Appendix

Appendix 1:The location information of these reports predicted positive

Latitude	Longitude	Latitude	Longitude	Latitude	Longitude
48.723779	-122.354431	47.896877	-122.323582	48.786349	-122.550631
48.986176	-122.69745	48.1496	-119.771316	47.57816	-122.21238
48.64683	-122.369771	47.372105	-122.460063	48.864727	-122.675513
47.8062	-122.1048	48.408405	-122.339303	47.583232	-122.653887
48.66002	-123.396765	48.936	-122.74729	48.948233	-122.748058
48.418776	-122.65316	48.50824	-122.23858	48.118052	-123.464837
48.2281	-122.266252	48.760013	-122.913326	47.375451	-122.168378
47.292402	-122.251238	48.664456	-122.395955	48.984144	-122.774438
48.752578	-122.622509	48.721109	-122.328668	48.924115	-122.255068
47.417355	-122.637451	48.316977	-121.355142		