



Article

Development of a Real-Time *Vespa velutina* Nest Detection and Notification System Using Artificial Intelligence in Drones

Yuseok Jeong ^{1,2,†}, Moon-Seok Jeon ^{1,2,†}, Jaesu Lee ¹, Seung-Hwa Yu ¹, Su-bae Kim ³, Dongwon Kim ³,
Kyoung-Chul Kim ¹, Siyoung Lee ¹, Chang-Woo Lee ² and Inchan Choi ^{1,*}

- ¹ Department of Agricultural Engineering, National Institute of Agricultural Sciences, Jeonju 54876, Republic of Korea; dbtjr2665@naver.com (Y.J.); sd118687@naver.com (M.-S.J.); butiman@korea.kr (J.L.); yush6210@korea.kr (S.-H.Y.); kkcmmole@korea.kr (K.-C.K.); leesy42@korea.kr (S.L.)
- ² School of Computer Information and Communication Engineering, Kunsan National University, Gunsan 54150, Republic of Korea; leecw@kunsan.ac.kr
- ³ Department of Agricultural Biology, National Institute of Agricultural Sciences, Jeonju 54874, Republic of Korea; subaekim@korea.kr (S.-b.K.); dongwonkim@korea.kr (D.K.)
- * Correspondence: inchchoi@korea.kr
- † These authors contributed equally to this work.

Abstract: *Vespa velutina* is an ecosystem disruptor that causes annual damage worth KRW 170 billion (USD 137 million) to the South Korean beekeeping industry. Due to its strong fertility and high-lying habitat, it is difficult to control. This study aimed to develop a system for the control of *V. velutina* nests using drones for detection and tracking the real-time location of the nests. *Vespa velutina* nest image data were acquired in Buan-gun and Wanju-gun (Jeollabuk-do), and artificial intelligence learning was conducted using YOLO-v5. Drone image resolutions of 640, 1280, 1920, and 3840 pixels were compared and analyzed. The 3840-pixel resolution model was selected, as it had no false detections for the verification image and showed the best detection performance, with a precision of 100%, recall of 92.5%, accuracy of 99.7%, and an F1 score of 96.1%. A computer (Jetson Xavier), real-time kinematics module, long-term evolution modem, and camera were installed on the drone to acquire real-time location data and images. *Vespa velutina* nest detection and location data were delivered to the user via artificial intelligence analysis. Utilizing a drone flight speed of 1 m/s and maintaining an altitude of 25 m, flight experiments were conducted near Gyeongcheon-myeon, Wanju-gun, Jeollabuk-do. A total of four *V. velutina* nests were successfully located. Further research is needed on the detection accuracy of artificial intelligence in relation to objects that require altitude-dependent variations in drone-assisted exploration. Moreover, the potential applicability of these research findings to diverse domains is of interest.

Keywords: artificial intelligence; *Vespa velutina* nest detection; unmanned aerial vehicle; real-time kinematics; long-term evolution; YOLOv5



Citation: Jeong, Y.; Jeon, M.-S.; Lee, J.; Yu, S.-H.; Kim, S.-b.; Kim, D.; Kim, K.-C.; Lee, S.; Lee, C.-W.; Choi, I. Development of a Real-Time *Vespa velutina* Nest Detection and Notification System Using Artificial Intelligence in Drones. *Drones* **2023**, *7*, 630. <https://doi.org/10.3390/drones7100630>

Academic Editor: Yangquan Chen

Received: 25 August 2023

Revised: 5 October 2023

Accepted: 9 October 2023

Published: 10 October 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Approximately 75% of the plant species grown for human consumption are pollinated by pollinators, such as bees and butterflies [1]. Honeybees are important for agriculture; therefore, population management is essential [2]. *Vespa velutina*, which was introduced to Busan via Shanghai, China in 2003, is a subtropical hornet native to tropical Asia that threatens honeybees and other pollinators in South Korea [3]. The annual damage caused by *V. velutina* is estimated to be approximately KRW 170 billion (USD 137 million) [1]. Unlike Korean hornets, *V. velutina* is highly successful in the rapid hunting of honeybees. This species accounts for 72% of the hornet populations collected nationwide and has strong reproductive power and adaptability to the urban environment [4]. A *V. velutina* nest colony includes 10,000–12,000 individuals, and during the breeding season, 300–500 *V. velutina* queens may be released from each nest, with efficient dispersibility [5]. Therefore,

V. velutina was classified as an ecosystem disturbance species in South Korea in 2019 [6]. From March to May, *V. velutina* queens that have overwintered begin to emerge. From April to May, when the weekly average maximum temperature is 27–28 °C, mating and larval development occur. Hornet numbers increase from the end of June, and intensive activities occur from August to September, which is the period when the number of *V. velutina* nests increase. These hornets are active until mid-October, when they overwinter and emerge in the following spring. *Vespa velutina* inhabit trees over 10 m high in the surrounding forests of nearby villages, and their life radius is approximately 1 km from the nests [7]. They are mostly undetectable with the human eye; hence, they are difficult to remove. Recent research has focused on automating the detection of *V. velutina* using optical sensors and machine learning to analyze the wingbeat frequencies of seven species, including bees, wasps, and hornets. The study was conducted within an entomological tent in a laboratory setting, where *V. velutina* individuals were differentiated from other species based on their wingbeat frequencies [8]. Previous research used a deep-learning-based portable monitoring system to mitigate the damage caused by *V. velutina* in apiaries [9]. Furthermore, research on the exploration of hornet nests has been undertaken. Moreover, dynamic hornet nest detection, tracking, and extermination systems based on unmanned aerial vehicles have been studied [10,11].

Artificial intelligence is a technology that allows machines to have reasoning and judgment capabilities that is applied in various fields, leading to technological innovation [12]. Artificial intelligence is divided into three main categories based on learning methods: supervised learning, unsupervised learning, and reinforcement learning [13]. Supervised learning involves the use of input and output data [14]. The artificial intelligence model analyzes input data and obtains previously provided predictive output data. Applications include classification or regression [15]. In unsupervised learning, data patterns and structures are identified based only on the input data [16], with applications for clustering or dimensionality reduction. Reinforcement learning maximizes the reward that a system can obtain when interacting with its environment [17]. The development of artificial intelligence technology is progressing rapidly due to advances in dealing with large amounts of data, computer hardware, and algorithms. Artificial intelligence can replace human labor in tasks ranging from simple to advanced [18]. Therefore, it is suitable for the detection of *V. velutina* nests, which requires accurate and fast analysis of large amounts of data.

Drones are used in various fields, such as industry, agriculture, the military, and medicine [19]. Aerial image acquisition using drones can collect visual information depending on the flight method and altitude [20]. However, as altitude increases, objects can be incorrectly detected as smaller than their actual size. Research on image analysis is also required for factors such as the flight speed, shooting angle, and light reflection of drones [21]. After the drone flight, the images stored in the drone are checked for additional analysis. High-definition images have large data sizes, making data transmission difficult due to the distance between the drone and the user. Therefore, delays occur in the stages of drone flight, image extraction, and analysis, resulting in poor real-time performance [22].

This study aimed to develop a notification system using artificial intelligence algorithms on drones for real-time detection of *V. velutina* nests, thereby enabling swift responses, collaboration, and precise actions. The research was conducted from October until overwintering (November). *Vespa velutina* nests are largest (80–90 cm diameter) before overwintering and become relatively easy to find as leaves fall. The number of overwintered queens emerging in the following year may be reduced by removing them when the colony is at its largest in order to reduce later damage.

2. Materials and Methods

2.1. System

The real-time *V. velutina* detection system using drones was designed to embed an artificial intelligence computer on the drone to alert the user during flight. When the *V. velutina* nest was detected by the analysis unit of the drone, the data were saved to the

database (DB) and could be accessed on the web service. Figure 1 depicts the overall configuration.

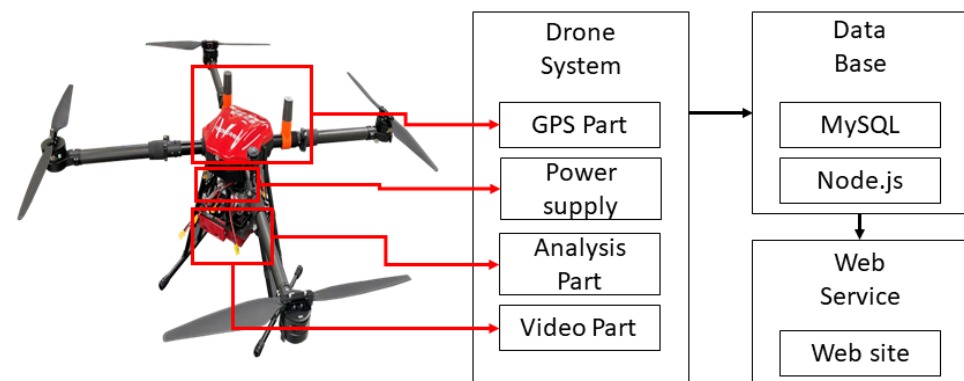


Figure 1. Diagram of a real-time *Vespa velutina* detection system using a drone.

A drone (YRX410 950 mm pre-M quadcopter, Shenzhen, China) was used to configure the system to fly with a take-off weight of 10 kg or more. The drone was equipped with modules for the real-time detection of *V. velutina* nests, including an information unit, a power supply unit, an analysis unit, and an imaging unit.

The power supply unit was composed of a 22.2 V, 22,000 mAh lithium-ion battery that powered the drone flight with a stable flight time of 20 min and a 12 V, 5 A battery that supplied the power for the analysis and imaging units. The information unit consisted of Pixhawk 4 M8N (HolyBro, Shenzhen, China) global positioning system (GPS) equipment mounted on the Pixhawk 4 (Holybro, China) flight controller and AsteRx-m3 (Septentrio, Leuven, Belgium) real-time kinematics (RTK)-GPS to communicate with the ground control system (ArduPilot Dev Team) module. RTK-GPS can obtain centimeter-level positional accuracy in real-time in the field using the correction value for the phase of the reference station that has precise positional information. The imaging unit was equipped with a camera for video confirmation during flight and a camera capable of acquiring images of up to 3840 × 2160 pixels in 30 frames on the gimbal. The analysis unit consisted of a computer (Jetson Xavier AGX, NVIDIA, Santa Clara, CA, USA) for artificial intelligence analysis. The DB stored information with MySQL (version 5.7) and provided web services with Node.js.

2.2. Data

Training Data Acquisition and Preprocessing

Vespa velutina nest data were collected from October 2021 to January 2022, the period when *V. velutina* colonies grew before overwintering. The size of *V. velutina* nests increased with increasing colony size, and in autumn, nest detection was easy with leaf loss exposure. More *V. velutina* queen candidates exist in the nest during this period. The data collection period was set for the ease of detection and elimination of *V. velutina* queen candidates. The *V. velutina* nest images were collected from 19 apiaries located in Buan-gun and Wanju-gun (Jeollabuk-do; Figure 2). The images were acquired at an altitude of 65–150 m above sea level using a red-green-blue (RGB) camera mounted on a drone. In the apiaries indicated by blue circles in Figure 2, *V. velutina* nests within 1 km of the surrounding area were detected, and *V. velutina* nests that existed continuously within a certain radius were additionally confirmed.

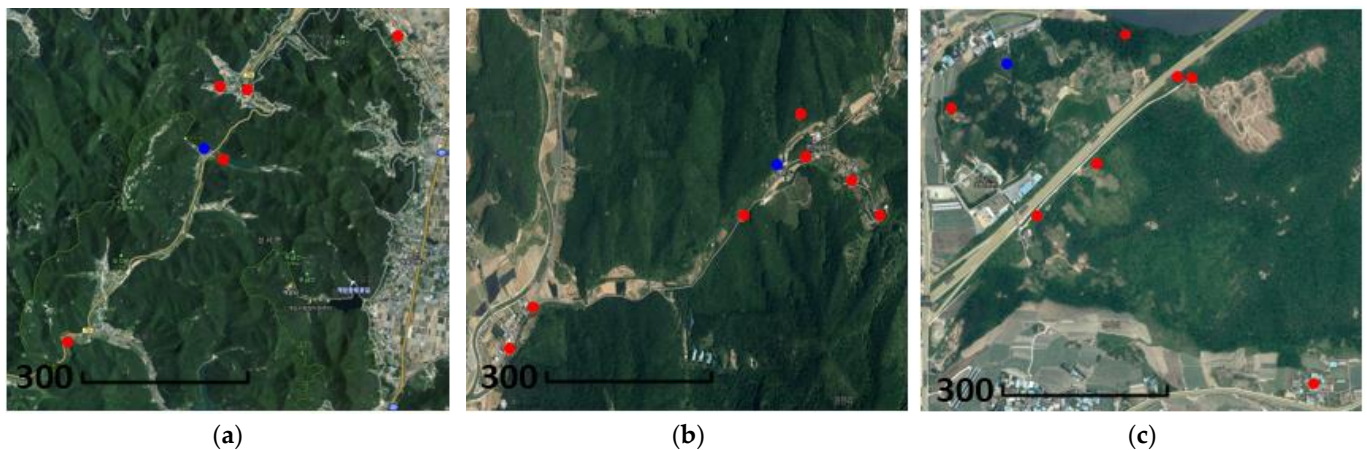


Figure 2. *Vespa velutina* nest location used for training. (a) Five nests detected in Sangseo-myeon (Buan-gun), (b) seven in Gyeongcheon-myeon (Wanju-gun), and (c) seven in Hwasan-myeon and Gosan-myeon (Wanju-gun) are shown. Red circles indicate *V. velutina* nests, and blue circles indicate apiaries.

A DJI Mavic2 Enterprise (DJI, Shenzhen, China) drone was used, and an RGB camera with a resolution of 3840×2160 pixels collected data at 30 frames/s. The video data were divided into frames, resulting in a total of 11,838 images. A total of 11,838 pieces of image data were used for learning, and 10,095 pieces of image data included the *V. velutina* nest coordinated information. The remaining 1173 were image data without *V. velutina* nests. Data preprocessing was performed using a Massachusetts Institute of Technology program (Labellmg v.1.8.6, MIT; Figure 3).



Figure 3. Data preprocessing using labellmg. Image depicts data at various angles and heights that were collected and preprocessed.

2.3. Artificial Intelligence Model Training

An Intel i7-11700k 3.6 GHz (CPU), a 128 GB memory card (RAM), and a NVIDIA GeForce RTX 3080TI 12 GB graphics card (GPU) were used for the artificial intelligence training. The artificial intelligence model was trained using YOLO-v5s. YOLO-v5s is largely composed of the backbone, neck, and head [23–26]. The backbone extracts the feature points of an object, the neck works to improve the performance through convergence and

data interpolation from the extracted feature points, and the head converts the feature points into a bounding box. The convolution layer (Conv) is a basic module used to extract feature points from an object, and C3 is a module that combines three Conv and bottleneck structures. Bottleneck is a module that is used to reduce the dimension of a channel when the channel value increases due to repeated Convs [27–29]. Spatial pyramid pooling–fast is a module that extracts fixed-size feature points for object images of different sizes [30,31]. Upsample is a module that increases the size of an image through data interpolation, and concat is a module that connects two or more data (Figure 4).

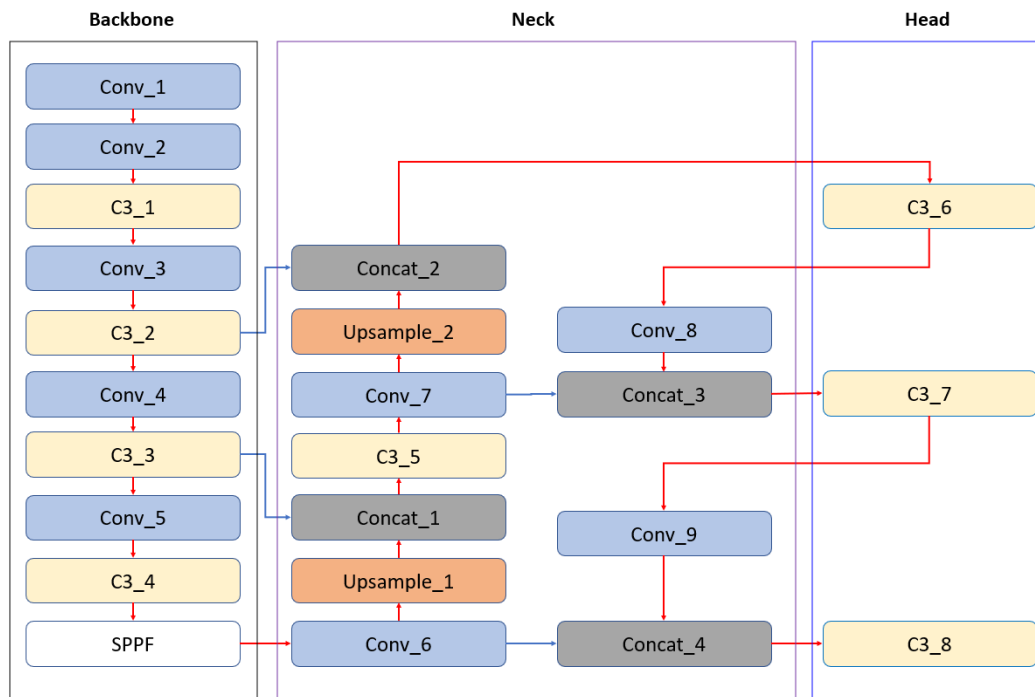


Figure 4. YOLO-v5s structure. The backbone extracts feature points, the neck improves the performance through convergence and interpolation of extracted feature points, and the head converts feature points into a bounding box. SPPF, Spatial pyramid pooling–fast.

The flight altitude is positively correlated with the range of data that can be obtained dependent on the angle of the installed camera equipment and the resolution. Resolution is a unit representing the number of pixels constituting the screen; generally, the higher the resolution, the clearer and more detailed the screen expression becomes. This means that the quantity and quality of information configured on the screen also increase; for data analysis, high-resolution images that contain more abundant and detailed data are advantageous for object detection using drones. As data continuously acquired in real-time must be processed without delay and accurate information must be detected, the learning image resolution was set and compared according to the performance of the analysis unit of the drone. In artificial intelligence learning, YOLO-v5s changes the size of the anchor box according to the input image size. Accordingly, the maximum and minimum sizes of detectable objects are different. Therefore, the model was trained by changing the size of the input image, changing the size of the detection box, randomly adjusting within the range of 50% of the input image (multi-scale) size, changing the minimum number of images used in learning (batch size), and setting parameters. To augment data on various image changes and locations according to the weather, image processing techniques, such as image mosaic and brightness contrast, were used to easily detect objects in drone flight images [32]. According to the processing speed in Table 1, four resolutions were set for learning.

Table 1. Real-time analysis unit processing speed performance table according to resolution.

Resolution (Pixels)	Analysis Part Processing Speed
640 × 480	60 frame/s
1280 × 960	30 frame/s
1920 × 1080	15 frame/s
3840 × 2160	3 frame/s

2.4. Model Learning Evaluation

2.4.1. Training

Artificial intelligence learning proceeded until there was no change in each training loss graph. Adaptive moment estimation (Adam) was used as the optimizer to find the minimum value of the training loss graph [33]. In Equation (1), θ was the loss function and g_t was the formula for obtaining the previous learning gradient.

$$g_t = \nabla_{\theta} f_t(\theta_{t-1}) \quad (1)$$

The formula for Adam provided an estimate of the primary momentum of the gradient m_t and an estimate of the secondary momentum v_t . The m_t was centered on the slope and v_t was calculated in inverse proportion to the square of the slope, with the advantage of being able to properly adjust the direction and movement distance (Equations (2) and (3)).

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (2)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (3)$$

where β was the decay rate for m_t and v_t , thereby playing the role that increased the proportion of the latest value and where the user could set the value to adjust the proportion. However, if the initial value of m_t , v_t was a 0 vector or β_1 , β_2 was close to 1, the bias became severe; therefore, the \hat{m}_t , \hat{v}_t equation with bias correction was used (Equations (4) and (5)).

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (4)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (5)$$

where η is the learning rate and ϵ is used to prevent division by 0 (Equation (6)).

$$\theta_{t+1} = \theta_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \hat{m}_t \quad (6)$$

2.4.2. Evaluating Confusion Matrix Performance

Artificial intelligence learning performance was judged by calculating the precision, recall, accuracy, and F1 score [34]. Precision refers to the proportion of actual correct answers among those classified as correct by the neural network. Recall refers to the proportion of actual correct answers that the neural network calls correct. Accuracy refers to the proportion of actual correct answers that the neural network considers correct and incorrect answers that the neural network considers incorrect. The F1 score is obtained using the harmonic mean of precision and recall. The confusion matrix consists of four values: true positive (TP), false positive (FP), true negative (TN), and false negative (FN). TP is when both the actual and predicted values are true. FP is when the predicted value is true but the actual value is false. TN is when both the actual and predicted values are false. FN is the case when the predicted value is false but the actual value is true. The precision,

recall, accuracy, and F1 score were determined using the confusion matrix value, and the formulae were as follows (Equations (7)–(10)).

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$F1 \text{ score} = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (10)$$

2.5. DataBase and Web Server

A DB was built using MySQL and a website was configured with Node.js. Image data and coordinate data were delivered to the analysis unit in real time, and when *V. velutina* nests were detected through artificial intelligence, coordinate information through RTK-GPS was saved in the DB and displayed on the website. The web service platform provided *V. velutina* notifications by accessing terminals such as personal computers, tablets, and mobile phones that users could communicate with at anytime and anywhere (Figure 5).

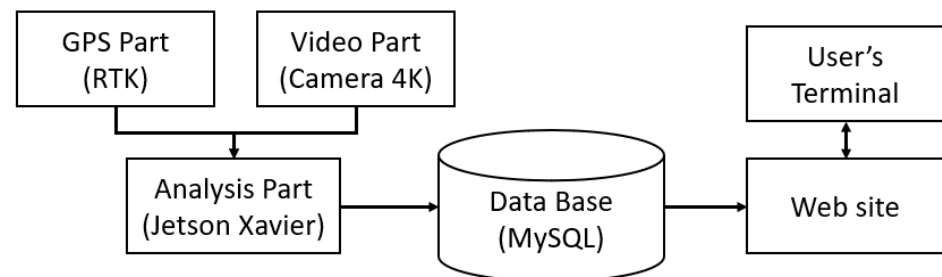


Figure 5. Web server configuration diagram. RTK, real-time kinematics.

3. Results

3.1. Artificial Intelligence Training Result

3.1.1. Training Graph

The model trained with 640-pixel resolution images was YOLO-v5s-default. The models trained with 1280-, 1920-, and 3840-pixel resolution images were YOLO-v5s-1280, YOLO-v5s-1920, and YOLO-v5s-3840, respectively (Figure 6).

The learning with 640-pixel resolution images had 50% random scaling, which was not applied to the input image size. This was due to image interpolation through resizing at low resolutions being small, which caused uncertain data increases (Table 2).

Table 2. Training settings for the four models.

Model	Epoch	Batch	Image Size	Multi-Scale	Image Processing
(a)	77	16	640	X	O
(b)	175	8	1280	O	O
(c)	179	4	1920	O	O
(d)	161	1	3840	O	O

(a) YOLO-v5s-default, (b) YOLO-v5s-1280, (c) YOLO-v5s-1920, (d) YOLO-v5s-3840.

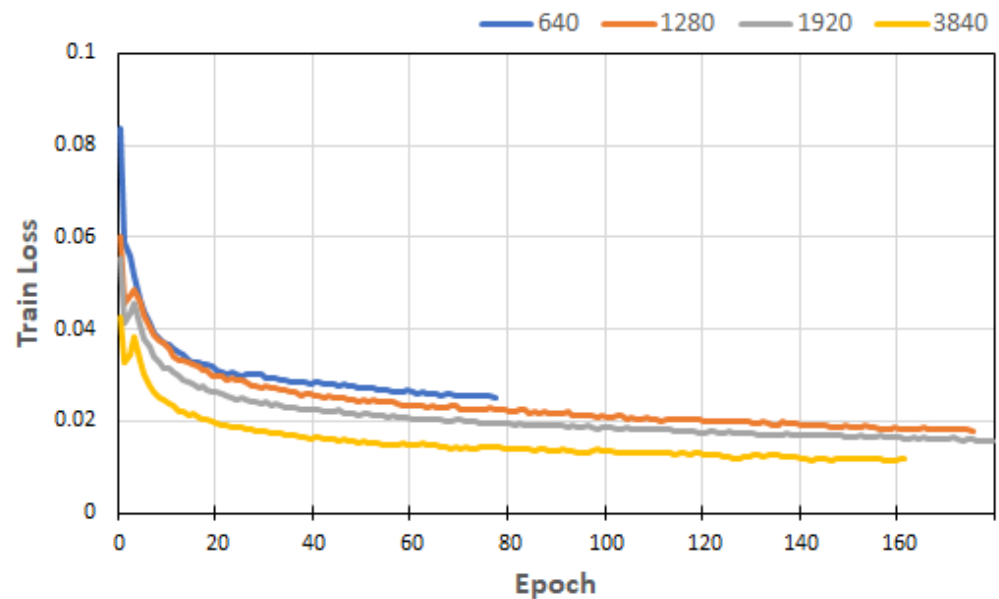


Figure 6. Training graph. Learning proceeded until there was little change in the graphs.

3.1.2. Validation

The TPs determined an actual *V. velutina* nest to be a *V. velutina* nest, the FPs determined a non-*V. velutina* nest to be a *V. velutina* nest, the FNs determined a *V. velutina* nest to be a non-*V. velutina* nest, and the TNs determined a non-*V. velutina* nest to be a non-*V. velutina* nest. A total of 7158 frames were used by dividing the 3840×2160 -pixel images taken for 3 min and 58 s at 30 frames/s in Wanju-gun (Jeollabuk-do) using the frame unit. The image included one *V. velutina* nest and was captured at an altitude of 40 m and a speed of 1 m/s.

At resolutions of 640, 1280, 1920, and 3840 pixels, the *V. velutina* nests were detected through each learning model. However, all 229 *V. velutina* nests in 7158 frames were not detected. The higher the resolution, the lower the number of false detections. YOLO-v5s-3840 trained with the highest resolution images obtained 98% better precision, 89.5% better recall, 7.5% better accuracy, and a 91.7% better F1 score than those of YOLO-v5s-default trained with the lowest resolution images (Table 3). False detections did not consistently decrease with resolution but occurred frequently at low resolutions, and the highest number of false detections was found with 1280-pixel resolution images.

Table 3. Confusion matrix.

Model	Confusion Matrix				Precision	Recall	Accuracy	F1 Score
	TP	FP	FN	TN				
(a)	7	332	222	6597	2.0%	3.0%	92.2%	2.4%
(b)	214	726	15	6203	22.7%	93.4%	89.6%	36.5%
(c)	213	220	16	6709	49.1%	93.0%	96.7%	61.2%
(d)	212	0	17	6929	100%	92.5%	99.7%	96.1%

(a) YOLO-v5s-default, (b) YOLO-v5s-1280, (c) YOLO-v5s-1920, (d) YOLO-v5s-3840, Confidence threshold ≥ 0.600 and IoU (intersection of union) ≥ 0.500 . TP, True Positive; FP, False Positive; FN, False Negative; TN, True Negative.

Figure 7 is a cropped image of the *V. velutina* nest detection area as a rectangle of the same size in each resolution image. The lower the resolution, the less data it contained in the wider range, and the higher the resolution, the more data it contained in the smaller range.

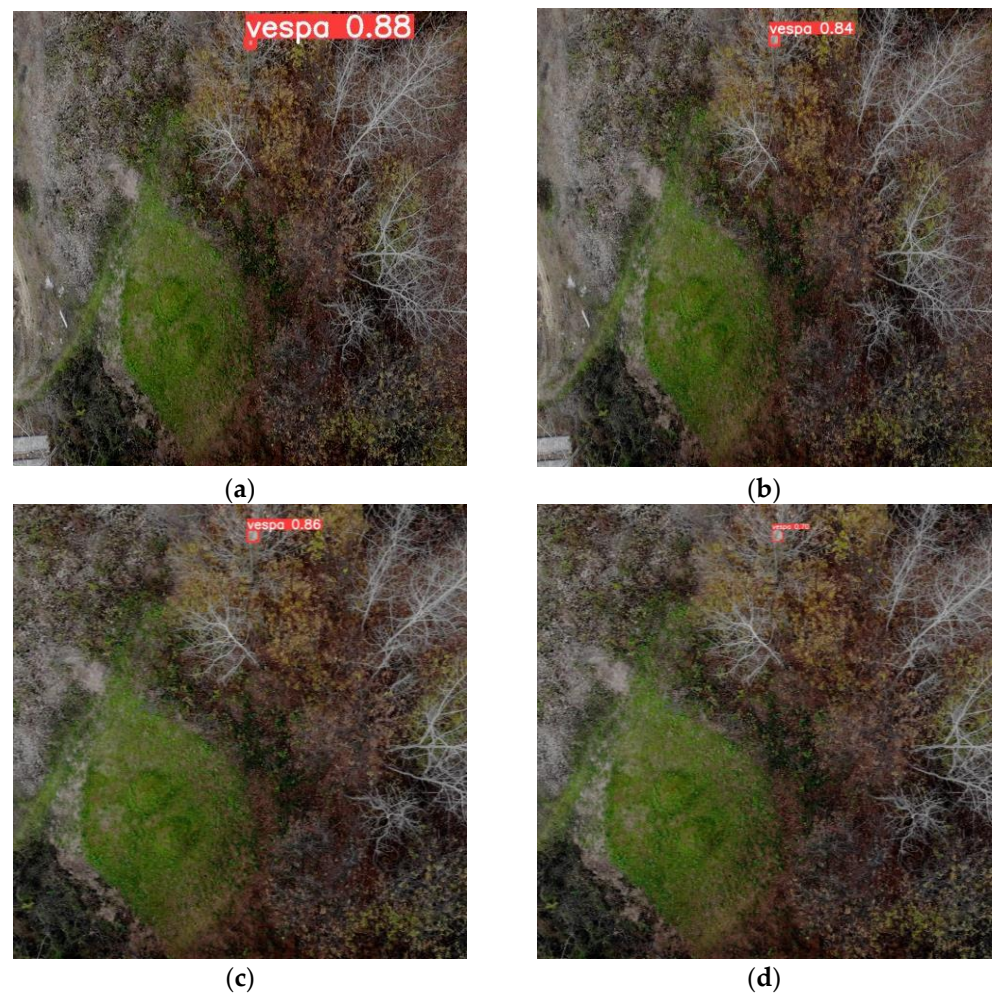


Figure 7. Detection of *Vespa velutina* nest in the image data. (a) YOLO-v5s-default (640-pixel resolution), (b) YOLO-v5s-1280 (1280-pixel resolution), (c) YOLO-v5s-1920 (1920-pixel resolution), and (d) YOLO-v5s-3840 (3840-pixel resolution).

For the 3840-pixel resolution model, there was no FP for the verification images. The 640-, 1280-, and 1920-pixel resolution models misdetected objects with similar colors and textures to *V. velutina* nests. Figure 8 is an image of false detection. Leaves, stones that looked white due to light reflection, and round stones with rough surfaces were mainly the objects falsely detected.

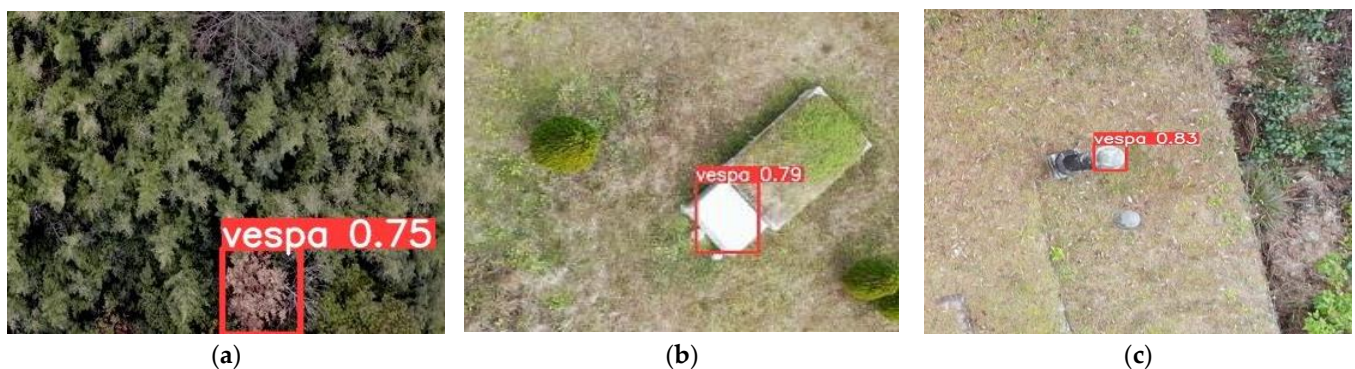


Figure 8. False detection image of three models. (a) YOLO-v5s-default (leaf), (b) YOLO-v5s-1280 (stone), and (c) YOLO-v5s-1920 (stone).

3.1.3. Selection of the Artificial Intelligence Model Used in the Drone System

When applying the trained artificial intelligence model to the real-time drone system, it was necessary to consider that object detection should be performed through artificial intelligence in the analysis unit of the drone for real-time captured images. The artificial intelligence object detection should proceed in the analysis unit until the next set of data is acquired, while obtaining the most accurate data without data loss based on the flight speed and altitude. Therefore, for accuracy, a detection model with a 3840-pixel resolution was selected; however, the processing speed was 3 frames/s. When 30 frames of real-time images are acquired per second, only 3 frames are used for artificial intelligence object detection. However, if a *V. velutina* nest was detected in even a single frame of 229 frames, coordinate information data that can be used to estimate the position of the nest could be obtained. YOLO-v5s-default detected the position of *V. velutina* nests at 7 frames, YOLO-v5s-1280 at 214 frames, YOLO-v5s-1920 at 213 frames, and YOLO-v5s-3840 at 212 frames. It was concluded that all four neural networks could estimate *V. velutina* nest locations. Moreover, images from drones flying at a constant speed at high altitudes had a large overlapping range of shooting areas and the size of objects changed every moment depending on the terrain and height of trees. Accurate high-resolution data were important, and even if a specific object appeared for only 1 s, it was repeatedly filmed at more than 30 frames. Even with the performance of 3 frames/s, it was possible to search for *V. velutina* nests, and the 3840-pixel resolution model was applied to the drone system to conduct demonstration search flights.

3.2. DataBase and Web Site

During drone flight, real-time RTK-GPS information was stored in the database through uniform resource locator (URL) transmission and added to the website map (Figure 9).

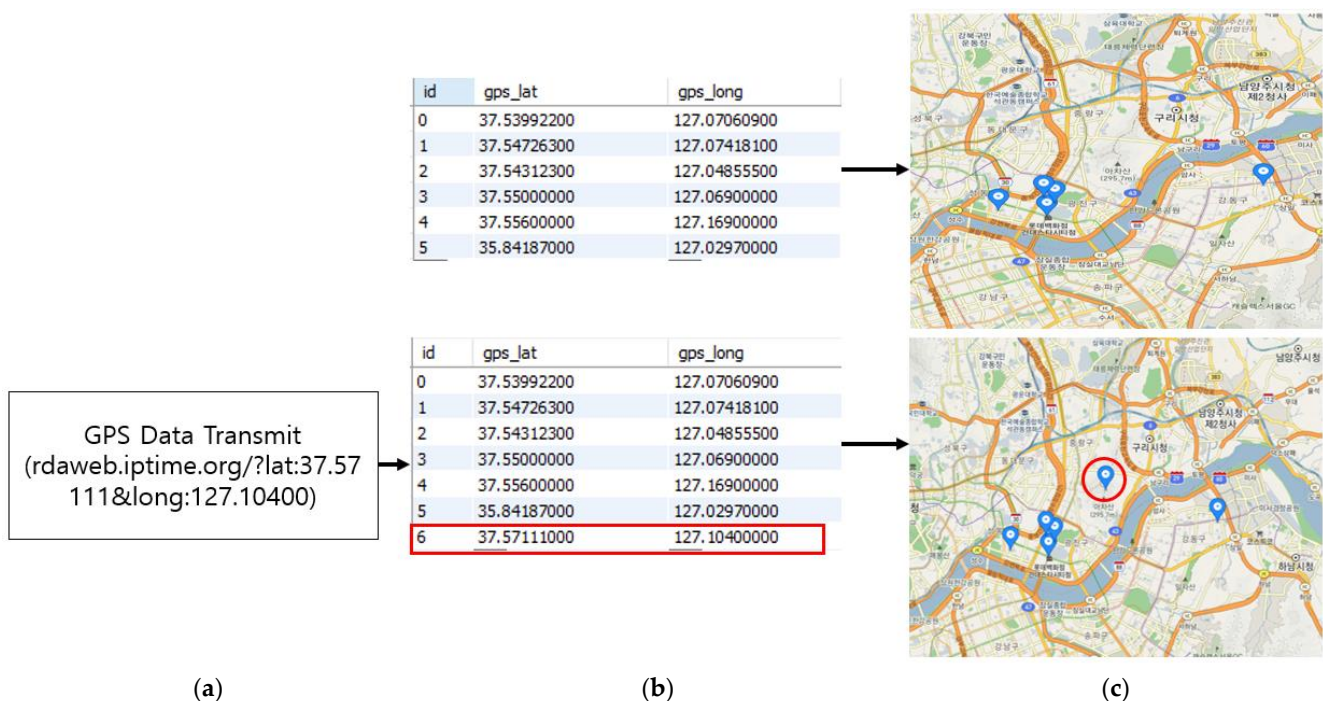


Figure 9. Implementation of the *Vespa velutina* nest location information and notification website. (a) The transmission of uniform resource locator (URL) coordinate information sends the command to save the coordinate information to the database (DB). (b) The DB where previously saved coordinate information and added information can be checked. (c) The website where a new marker created according to the changed coordinate information in the DB can be checked.

3.3. Demonstration of the *V. velutina* Nest Real-Time Detection System in Drones

3.3.1. Routed Detection Flight Settings

In November 2022, on-site verification was conducted in Gyeongcheon-myeon (Wanju-gun). To perform uniform and accurate detection over a certain range, the drone was flown with a routed automatic flight using an automatic flight setting program that specified the speed, altitude, route, camera angle, and direction of the drone (Figure 10). The analysis unit of the drone detects *V. velutina* nests using three frames at 30 frames/s with 3840×2160 -pixel resolution acquired from the imaging unit. The drone flew at 1 m/s at a set height of 25 m above the ground.



Figure 10. Routed automatic flight program for drones. (a) The mission planner main screen showing the status of the drone. (b) The process of specifying the drone flight path. (c) The process of setting drone flight details (speed, drone direction, etc.). (d) The command procurement button screen after setting.

3.3.2. Demonstration Detection Flight

Four *V. velutina* nests in Gyeongcheon-myeon (Wanju-gun) that had been spotted in advance were searched for using the routed automatic flight settings for the drone. *Vespa velutina* nests of various shapes were detected according to topography, *V. velutina* nest size, tree height, etc., and it was difficult to search due to the influence of light according to the weather and time of day. Image and coordinate information were saved only when the *V. velutina* nest was detected in the acquired image during the routed automatic flight. In addition, real-time coordinate information was created on the web server simultaneously with the *V. velutina* nest detection process (Figure 11).

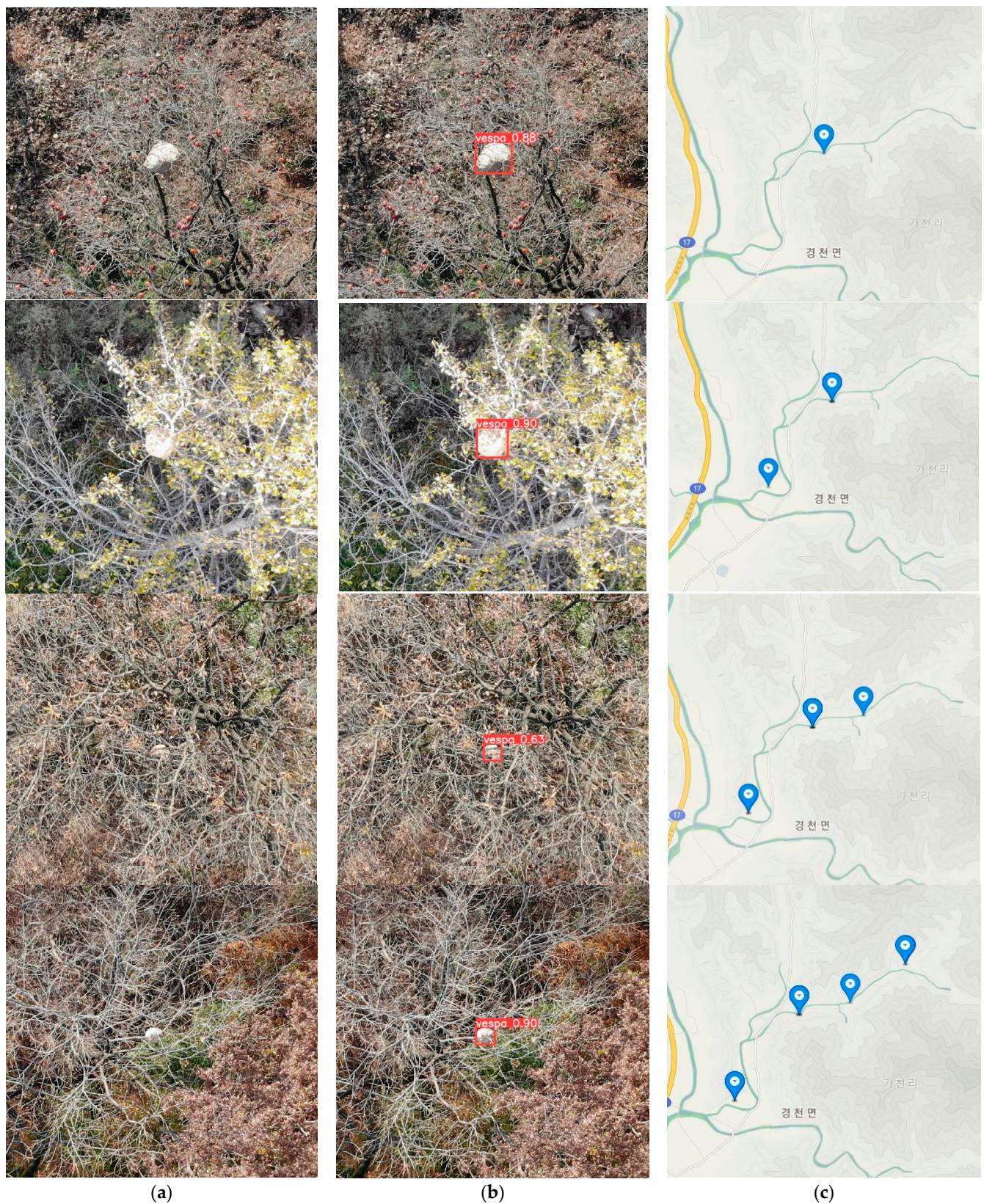


Figure 11. Images of detection and real-time coordinate generation during the demonstration flight. (a) Photo of a *V. velutina* nest taken by an actual drone in flight. (b) The result of *V. velutina* nest detection in the analysis unit of the drone during flight. (c) Image that the analysis unit transmits when saving the coordinate information on a DB and simultaneously informing the website (경천면 is Gyeongcheon-myeon).

4. Discussion

South Korea experiences very cold winters due to being a four-seasons country; during this time, insects hibernate. Except for the queen hornet, all the other hornets die during the winter, which can cause the extinction of the colony. Therefore, the queen hornet does not hibernate in the colony's hive but hibernates among the surrounding trees or fallen leaves. This leads the queen hornets becoming active in the spring. The queen hornet establishes a nest and produces worker hornets [5]. It is advisable to remove the queen hornet and the nest early on. Due to that reason, we are conducting research to detect and remove hornets in apiaries in the spring [9]. However, finding hornet nests is exceedingly difficult in the summer when the nest is small and the forest is dense. Therefore, our goal was to search and remove the nest during the fall when the colony is at its largest, reducing the number of hornet individuals that could possibly emerge the following year. As a result, we were determined and have conducted some research during the period when we could find the most nests in a short period of time. This study was conducted from October until winter. The nests are at their largest sizes (80–90 cm diameter) before winter and become relatively easy to find as leaves have fallen from the trees.

In this study, exploration of designated areas was possible using predefined drone paths and both exploration and notifications occurred simultaneously. Moreover, general drone image analysis was feasible; however, the drone's flight time and image analysis time are separate, thereby necessitating additional time. This results in a time delay for subsequent actions, such as a delay in emergency evacuation notifications. The real-time detection and notification system studied in this paper offers the advantage of enabling immediate follow-up. It can also explore multiple nests on a single flight by navigating local ranges rather than tracking a single object.

However, there were shortcomings in terms of the artificial intelligence algorithms. There were no mid-flight misidentifications, but the frequency of *V. velutina* nest detection was low. The detection numbers in the first, second, third, and fourth flights were 3, 5, 3, and 8, respectively (Figures 12–15).

It is expected that search performance can be enhanced by improving performance through the addition of *V. velutina* nest data acquired in more diverse regions or by utilizing artificial intelligence models with higher performance than YOLOv5 [35,36].

In this study, the field inspections for nests were conducted in locations entirely different from the ones used for training video data. The inspections took place in four different locations, each with varying shapes and sizes of the nests. To achieve robust results and confidence, extensive testing is necessary. However, since drone flights and video shootings are only possible in places permitted by Korean law, the study was only conducted in a limited number of places. Therefore, a future detection plan is needed.

Due to the current size of drones, exploration in areas with many obstacles is not feasible. This leads to the conclusion that research on miniaturized and lightweight drone designs is crucial to address the challenges of our research.



Figure 12. *Vespa velutina* nest detection images from the first demonstration flight.

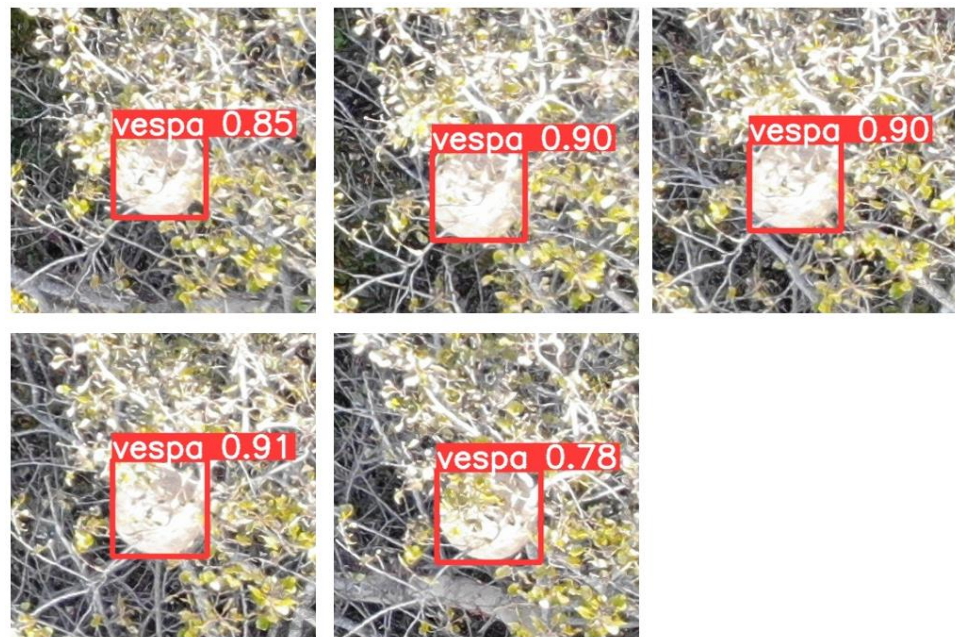


Figure 13. *Vespa velutina* nest detection images from the second demonstration flight.



Figure 14. *Vespa velutina* nest detection images from the third demonstration flight.

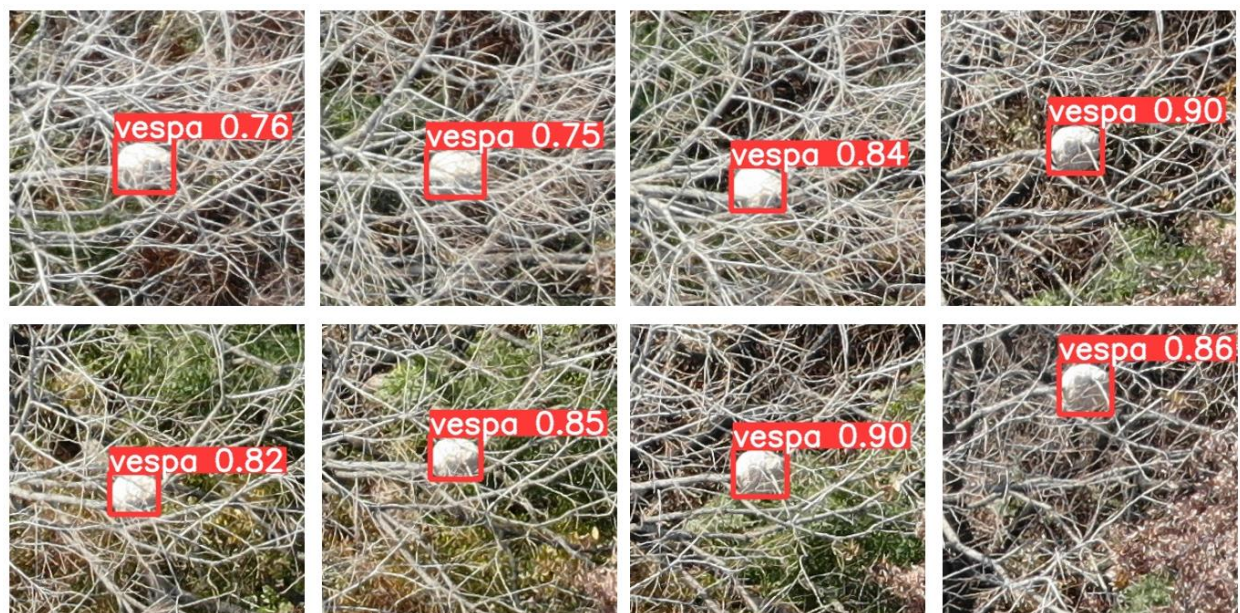


Figure 15. *Vespa velutina* nest detection images from the fourth demonstration flight.

5. Conclusions

This study aimed to reduce apiary damage in the following year by targeting *V. velutina* queen candidates before they overwintered using real-time location information notifications. To obtain an artificial intelligence model suitable for real-time video analysis during drone flights, we conducted comparative experiments based on different resolutions. The 3840-pixel resolution model had no false detection for the verification image and obtained the best detection values, with a precision of 100%, a recall of 92.5%, an accuracy of 99.7%, and an F1 score of 96.1. As a result of routed and automatic flight detection at a drone flight speed of 1 m/s and an altitude of 25 m, four *V. velutina* nests were detected. In addition, we were able to exploit the advantage of receiving real-time alerts during the flight through actions such as pest control and evacuation. Further research is needed to increase the detection accuracy of artificial intelligence for objects that vary according to altitude in detection using drones; this is expected to be able to be applied in other fields.

Author Contributions: Conceptualization, I.C. and K.-C.K.; data curation, Y.J. and M.-S.J.; formal analysis I.C., J.L., C.-W.L., Y.J. and M.-S.J.; investigation, I.C., S.-H.Y., S.-b.K., D.K., K.-C.K., Y.J. and M.-S.J.; Methodology, I.C., J.L., K.-C.K., S.L. and C.-W.L.; resources, S.-H.Y., S.-b.K., D.K. and S.L.; software, Y.J. and M.-S.J.; supervision, I.C.; validation, I.C., J.L., Y.J. and M.-S.J.; writing—original draft preparation, I.C., Y.J. and M.-S.J.; writing—review and editing, I.C., Y.J. and M.-S.J. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Research Program for Agricultural Sciences (PJ01497002), National Institute of Agricultural Sciences, Rural Development Administration, Republic of Korea.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Jung, C.E.; Kang, Y.R.; Oh, H.A.; Bak, S.; Hong, D.; Kwon, S. Bee Crisis and Protection Policy Proposal. *Greenpeace* **2023**, 59.
2. Kay, A.D.; Bruning, A.J.; van Alst, A.; Abrahamson, T.T.; Hughes, W.O.H.; Kaspari, M. A carbohydrate-rich diet increases social immunity in ants. *Proc. R. Soc. B Boil. Sci.* **2014**, *281*, 20132374. [[CrossRef](#)]
3. Jung, C.E.; Kang, M.S.; Kim, D.W.; Lee, H.S. Vespidae wasps (Hymenoptera) occurring around apiaries in Andong, Korea I. Taxonomy and life history. *Korean J. Apic.* **2007**, *22*, 53–62.
4. Jung, C.E.; Kang, M.S.; Kim, D.W.; Lee, H.S. Vespidae wasps (Hymenoptera) occurring around apiaries in Andong, Korea II. Taxonomy and life history. *Korean J. Apic.* **2007**, *22*, 63–70.
5. Jung, C.E. Initial stage risk assessment of an invasive hornet, *Vespa velutina nigrithorax* Buysson (Hymenoptera: Vespidae) in Korea. *Korean J. Apic.* **2012**, *27*, 95–104.
6. Sim, H.; Lee, M.L.; Choi, Y.S.; Kim, H.Y.; Hong, I.P.; Woo, S.O.; Byen, K.U.; Lee, M.Y. Pattern of Emergence of *Vespa velutina nigrithorax* Buysson (Hymenoptera: Vespidae) on Spring in South Part of Korea. *J. Apic.* **2014**, *29*, 353–358. [[CrossRef](#)]
7. Abrol, D.P. Ecology, behaviour and management of social wasp, *Vespa velutina* Smith (Hymenoptera: Vespidae), attacking honeybee colonies. *Korean J. Apic.* **1994**, *9*, 5–10.
8. Herrera, C.; Williams, M.; Encarnação, J.; Roura-Pascual, N.; Faulhaber, B.; Jurado-Rivera, J.A.; Leza, M. Automated detection of the yellow-legged hornet (*Vespa velutina*) using an optical sensor with machine learning. *Pest Manag. Sci.* **2023**, *79*, 1225–1233. [[CrossRef](#)]
9. Jeon, M.-S.; Jeong, Y.; Lee, J.; Yu, S.-H.; Kim, S.-B.; Kim, D.; Kim, K.-C.; Lee, S.; Lee, C.-W.; Choi, I. Deep Learning-Based Portable Image Analysis System for Real-Time Detection of *Vespa velutina*. *Appl. Sci.* **2023**, *13*, 7414. [[CrossRef](#)]
10. Gao, P.; Lee, K.; Kuswidiyanto, L.W.; Yu, S.-H.; Hu, K.; Liang, G.; Chen, Y.; Wang, W.; Liao, F.; Jeong, Y.S.; et al. Dynamic Beehive Detection and Tracking System Based on YOLO V5 and Unmanned Aerial Vehicle. *J. Biosyst. Eng.* **2022**, *47*, 510–520. [[CrossRef](#)]
11. Lee, C.G.; Yu, S.H. Exterminator for the Nests of *Vespa velutina nigrithorax* Using an Unmanned Aerial Vehicle. *Drones* **2023**, *7*, 281. [[CrossRef](#)]
12. Hamet, P.; Tremblay, J. Artificial intelligence in medicine. *Metabolism* **2017**, *69*, S36–S40. [[CrossRef](#)] [[PubMed](#)]
13. Olson, D.L.; Delen, D. *Advanced Data Mining Techniques*; Springer Science & Business Media: Berlin/Heidelberg, Germany, 2008. [[CrossRef](#)]
14. Liu, B.; Liu, B. *Supervised Learning. Web Data Mining: Exploring Hyperlinks, Contents, and Usage Data*; Springer: Berlin, Germany, 2011; pp. 63–132. [[CrossRef](#)]

15. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a convolutional neural network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6. [\[CrossRef\]](#)
16. Van, O.M.; Wiering, M. *Reinforcement learning and markov decision processes* In *Reinforcement Learning*; Springer: Berlin/Heidelberg, Germany, 2012; pp. 3–42. [\[CrossRef\]](#)
17. Mnih, V.; Kavukcuoglu, K.; Silver, D.; Graves, A.; Antonoglou, I.; Wierstra, D.; Riedmiller, M. Playing atari with deep reinforcement learning. *arXiv* **2013**, arXiv:1312.5602. [\[CrossRef\]](#)
18. Dosilovic, F.K.; Brcic, M.; Hlupic, N. Explainable artificial intelligence: A survey. In Proceedings of the 2018 41st International Convention on Information and Communication Technology, Electronics and Microelectronics (MIPRO), IEEE, Opatija, Croatia, 21–25 May 2018; pp. 0210–0215. [\[CrossRef\]](#)
19. Ayamga, M.; Akaba, S.; Nyaaba, A.A. Multifaceted applicability of drones: A review. *Technol. Forecast. Soc. Chang.* **2021**, *167*, 120677. [\[CrossRef\]](#)
20. Shahmoradi, J.; Talebi, E.; Roghanchi, P.; Hassanalian, M. A Comprehensive Review of Applications of Drone Technology in the Mining Industry. *Drones* **2020**, *4*, 34. [\[CrossRef\]](#)
21. Tang, L.; Shao, G. Drone remote sensing for forestry research and practices. *J. For. Res.* **2015**, *26*, 791–797. [\[CrossRef\]](#)
22. Wang, X.; Chowdhery, A.; Chiang, M. SkyEyes: Adaptive video streaming from UAVs. In Proceedings of the 3rd Workshop on Hot Topics in Wireless 2016, New York, NY, USA, 3–7 October 2016; pp. 2–6. [\[CrossRef\]](#)
23. Zhu, X.; Lyu, S.; Wang, X.; Zhao, Q. TPH-YOLOv5: Improved YOLOv5 Based on Transformer Prediction Head for Object Detection on Drone-captured Scenarios. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops, Montreal, BC, Canada, 11–17 October 2021; pp. 2778–2788. [\[CrossRef\]](#)
24. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 7263–7271. [\[CrossRef\]](#)
25. Redmon, J.; Farhadi, A. Yolo3: An incremental improvement. *arXiv* **2018**, arXiv:1804.02767. [\[CrossRef\]](#)
26. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. Scaled-yolov4: Scaling cross stage partial network. In Proceedings of the IEEE/cvf Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 13029–13038.
27. Liu, Z.; Lin, T.; Cao, Y.; Hu, H.; Wei, Y.; Zhang, Z.; Lin, S.; Guo, B. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, BC, Canada, 11–17 October 2021; pp. 10012–10022. [\[CrossRef\]](#)
28. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.; Kaiser, L.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*.
29. Woo, S.; Park, J.C.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19. [\[CrossRef\]](#)
30. Wang, Q.; Wu, B.; Li, P.; Zuo, W.; Hu, Q. Spatial Channel Attention for Deep Convolutional Neural Networks. *Mathematics* **2022**, *10*, 1750. [\[CrossRef\]](#)
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *37*, 1904–1916. [\[CrossRef\]](#)
32. Tan, M.; Pang, R.; Le, Q.V. EfficientDet: Scalable and efficient object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 10781–10790. [\[CrossRef\]](#)
33. Kingma, D.P.; Ba, J.A. A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980. [\[CrossRef\]](#)
34. Stehman, S.V. Selecting and interpreting measures of thematic classification accuracy. *Remote Sens. Environ.* **1997**, *62*, 77–89. [\[CrossRef\]](#)
35. Zhao, L.; Zhu, M. MS-YOLOv7: YOLOv7 Based on Multi-Scale for Object Detection on UAV Aerial Photography. *Drones* **2023**, *7*, 188. [\[CrossRef\]](#)
36. Zeng, Y.; Zhang, T.; He, W.; Zhang, Z. YOLOv7-UAV: An Unmanned Aerial Vehicle Image Object Detection Algorithm Based on Improved YOLOv7. *Electronics* **2023**, *12*, 3141. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.