Threat Detection in Hostile Environment with Deep Learning based on Drone's Vision

Sufiyan Shaikh[1], Rushikesh Raskar[2], Lajri Pande[3], Zeenat Khan[4], Prof Shweta P.Guja[5]

[1,2,3,4]B.E student, Dept. of computer, NBN Sinhgad School of Engineering, Ambegaon, Pune-411041, Maharashtra, India
[5]Prof. ,Dept of Engineering, NBN Sinhgad School of Engineering, Ambegaon, Pune-411041, Maharashtra, India

---

Abstract - *We consider a video surveillance application, using fa camera mounted on a drone flying over the area to be monitored and sending the video to a control center. The resulting network is composed of a static component, and a moving component (Drone). The video signal will be transferred via radio signal to the control center. The control center will be equipped with AI/Ml enabled application which will be able to detect what is in the frame. This system is very useful in hostile environment where we need to examine the whole area before proceeding and where the involvement of human being is dangerous to his life. This drone will the video feed with minimum to very low latency. The drone will be stealth and will make low noice which will make it difficult to be easily noticed*

1. INTRODUCTION

Nowadays, the interest and demand in a drone are increased very much. Due to this increase in demands, drone merchandise of newer types is being designed and manufactured, so that civilians can afford to buy them for various purposes (i.e., research, leisure, etc.). As for this shift, the commercial drone industry keeps on growing.

There has been a massive production in commercial drones to satisfy the civilian's needs, but this has a drawback. As it became easier to spot drones outdoors, more safety issues have been brought up as concerns. These are not merely about accidents regarding drones harming individuals, but include drones invading government restricted areas. Additionally, considering that a coordinated fleet of drones is capable of more various tasks, drones can be a bigger threat than people could imagine. As there are more drones out in public, it became harder to regulate them legally and safely.

When it comes to the field of detection, the first thing that comes to mind is a radar. Radar can be bigger or smaller in size depending upon the application.

In this paper, we propose a comprehensive drone detection system based on machine learning as an alternative solution for the aforementioned matter. This system applies object detection to capture the possible threats in the image. Our system can be applied in an environment with surveillance drones with the following scenario. During its mission, the surveillance drone will take off from the start position and fly to a specified point and then record its surroundings on camera. If there is any vulnerability or threats, it will be found on the video frame by applying object detection and the system will mark it on the frame. The marked frame is then used further. The system which has learned various threats through machine learning identifies the threat by using the model. Such-system can be applied to another surveillance system such as complement it is applied to the other ability to find unknown threats.

For image processing, the system uses GPU with CUDA programming on the video frames. In each frame, the object detection algorithm is applied to spot the threat in the video frames. Specifically, the YOLO algorithm is adapted for object detection. With the help of machine learning based on various threat images, which are given and processed beforehand, the system learns how to detect the threats accurately from frames. Then, frames with threats detected are processed once more to identify the threat model type

2. RELATED WORK

In this section, we introduce various researches relevant to our research. The content of this paper can be summarized as a drone detection and identification system based on image processing and machine learning. YOLO is applied for object detection. This technique has three key characteristics. One is that by using Integral Image, features used by the detector can be computed very quickly, making its computation complexity O (1).Two is the method constructing the classifier by selecting a small number important feature using a learning algorithm, yielding a very efficient classifier. Three is the method for combining increasingly more complex classifiers in cascade fashion, making it dramatically fast.

Recently, deep learning has been applied, in many studies in image processing field. Among the deep neural network structures, Convolutional Neural Network (CNN) has been gaining great interest because it automatically detects the important features during its learning phase. It has already shown performance breakthroughs in image classification and object detection. The reason why CNN shows great performance in extracting feature is its capability in extracting abstract features. This allows it process with greater accuracy and speed. In this paper, threat identification is based on supervised learning. Supervised learning is a systematically more adjusted CNN, and has been applied. It has been gaining recognition by showing better performance than unsupervised learning.

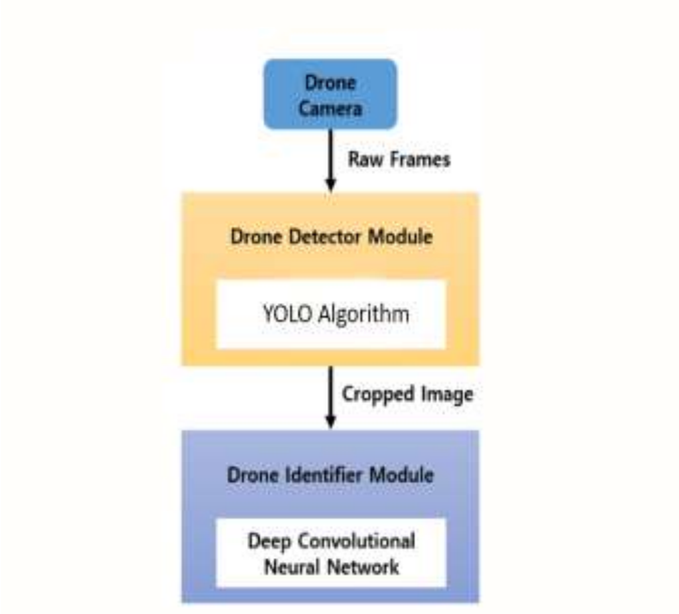


Fig. 1: Process of Drone Detection and Identification System.

Based on the aforementioned researches, we have chosen CNN to increase speed and efficiency. The two modules implemented in this system requires learning based on images. As it heavily relies on machine learning, with careful calibration and sufficient learning, the system can show better performance.

3. SYSTEM DESIGN AND IMPLEMENTATION

In this section, we describe our proposed system and implementation details. We first explain the overall system. Then, we describe each modules in detail.

3.1 Overall System

The proposed system is designed to be on top of a computing board established on each surveillance drones. After going through the process described in Fig.1, the drones are to report to the Ground Control Station (GCS).

3.2 Threat Detector Module

The threat detector is module that detects threats on video frames taken by the camera on the drones. Accurate 2D detection on the frame is key to the overall system. Therefore, to ensure accuracy while maintaining speed, our system adapts YOLO algorithm and OpenCV libraries. For training the classifier, we used 2088 positive examples and 3019 negative examples. We collected the positive examples from Google, and manual cropped the drone from every images. Negative examples were collected fromhttp://face.urtho.net/.By applying image distortion on the positive examples, we expanded our positive example pool to 7000. Based on these positive images, we trained our Classifier. The trained classifier finds the 2D area on the frame, where the threat is located. Then, the area is cropped, and passed on to the drone identifier module.

3.3 Threat Identifier Module

The threat identifier is a module that classifies the unknown threat into vendor models based on the cropped image. To identify the type of the thread, a CNN is constructed. Threat detection and threat identification can be done by other types of deep learning such as ResNet or Raster R-CNN. However, due to its depth, it would take a very long time for training, and the complexity is too great to construct. To overcome this, we have separated the system into a classifier and a CNN for processing. Additionally, this allows us to train faster with fewer image. The CNN consists of two Conv layers and two fully connected layers. Each fully connected layer are implemented with a 30 percent dropout. For the activation function, Rectified Linear Unit (ReLu) is used, and Adaptive Motion Estimation (Adam) is used for the gradient descent algorithm.

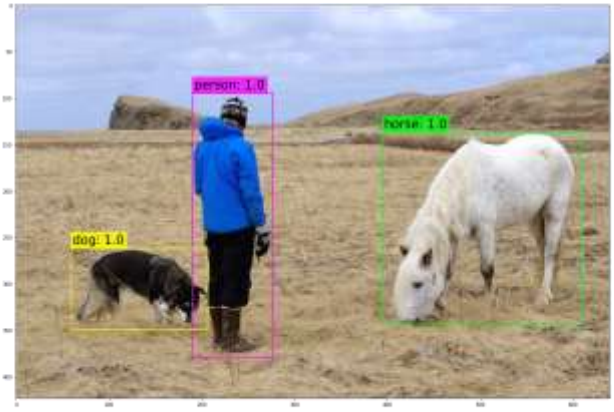
4. PERFOMANCE EVALUAITON

In this section, we show our experiment results and performance analysis. First, we show how the threat detection module works. Then, we analyze the performance of the threat identifier module. In the system, we use the YOLO algorithm for object detection. The YOLO is a very fast real-time multi-object detection algorithm. YOLO stands for "You Only Look Once". This algorithm applies a neural network to an entire image. The image is divided by the network into an S x S grid and gives the result with bounding boxes, which are boxes drawn around images and predicted probabilities for each of these regions. The method used to come up with these probabilities is logistic regression. The bounding boxes are weighted by the associated probabilities. For class prediction, independent logistic classifiers are used.

4.1 Threat Detector Experiment

Before running the threat detector module, it is required to go through a learning phase. For the learning, the images collected from Google has been used for the drone detectors supervised learning. After using the 2088 positive example images, we have prepared 829 test images for testing the performance of the system. These test images were not used for the learning process, and have only been used for performance

evaluation. This is to validate that our system can still detect threats from unfamiliar images that has not been exposed. The detection results of the threat detector. As shown in the image, if the system estimates that there is a threat in the image, then the area, where the threat is expected to be, is marked in a rectangular box.

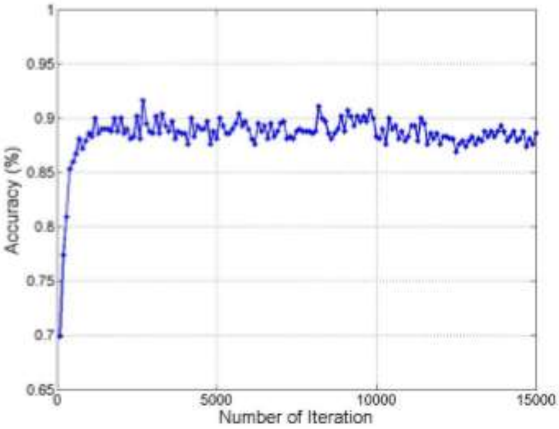


4.2 Threat Identifier Experiment

The threat identifier is to identify the threat model. The 2088 positive example images were all manually labeled. 1777 image among the positive example images were used for training, and 429 were used as evaluation test images. Minibatch with 256 batch size was used for training. The threat identifier shows maximum of 91.6 percent accuracy.

False classification are only made in cases that are quite challanging even for the human eyes. Shows that the system reaches its maximum accuracy within a few iteration, and saturates around 89 percent.

This experiment proves that Deep CNN is very efficient for classifying threat by model using a drone equipped with a camera. Additionally, it shows that it doesnt requires much training data.



5. CONCLUSIONS

In this paper, we propose a threat detection and identification system making decision based on video free obtained from a camera which is placed on a drone. This system has shown that even with simple artificial intelligence, the performance is very promising. All systems were actually implemented, and training data were collected from the web. With small amount of easily collectable training data, the system still showed great accuracy, which makes it more appealing. For future work, we would like to develop a distance estimation module to complement the existing system. Based images, estimating the distance between the surveillance drone and the unknown threats can be valuable information, and can be used for tracking. Although commercial drones have been massively produced to satisfy the civilian's needs, there have been some downsides to this.

As it became easier to spot drones outdoors, more safety issues have been brought up as concerns. These are not merely about accidents regarding drones harming individuals, but include drones invading government restricted areas.

The major need for this type of technology will be in the area of detection. The drone can be used for detecting harmful threads, it can be used for detecting any vulnerability in the crowded areas. The system which will internally, can be trained to detect anything with maximum possible accuracy.

The camera mounted on the drone will feed the video signal to the algorithm which will identify the elements of interest and create a bounding box and also predict the probability of accuracy of the identification.

REFERENCES

[1] S. Yoo, J. Jung, A. Y. Chung, K. Kim, J. Lee, S. Park, S. K. Lee, H. K. Lee, and H. Kim, "Empowering drones teamwork with airborne network," in Advanced Information Networking and Applications (AINA), 2017 IEEE 31st International Conference on. IEEE, 2017, pp. 678–685.

[2] J. Jung, S. Yoo, W. La, D. Lee, M. Bae, and H. Kim, "Avss: Airborne video surveillance system," Sensors, vol. 18, no. 6, p. 1939, 2018.

[3] S. Cook, CUDA programming: a developer's guide to parallel computing with GPUs. Newnes, 2012.

[4] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on, vol. 1. IEEE, 2001, pp. I–I.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Advances in neural information processing systems, 2015, pp. 91–99.

[7] W. Shi, J. Caballero, F. Husz□r, J. Totz, A. P. Aitken, R. Bishop, D. Rueckert, and Z. Wang, "Real-time single image and video superresolution using an efficient sub-pixel convolutional neural network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, pp. 1874–1883.

[8] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[9] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," nature, vol. 521, no. 7553, p. 436, 2015.