Flying Object Detection and Classification using Deep Neural Networks

Suresh Arunachalam T¹, Shahana R², Vijayasri R³, Kavitha T⁴

B.Tech (CSE) Final Year, Periyar Maniammai Institute of Science & Technology, Thanjavur, India
tasuresh97@gmail.com, shahanajanani1997@gmail.com, vijayasrirajendren@gmail.com
Assistant Professor, CSE⁵, Periyar Maniammai Institute of Science & Technology, Thanjavur, India
tkavitha07@yahoo.com

Abstract— Unmanned Arial Vehicles (UAVs) are extensively used everywhere in commercial applications such as delivering goods and medicines, taking photographs, and to monitor crowded areas. Sometimes these drones are used for capturing our private information without our knowledge. To avoid misuse of UAVs, we need to detect them in advance before entering into the protected areas. Detecting the UAV is a complex task because it is supplemented by birds, aircrafts, moving clouds, and swaying trees. To prevent this, we will detect the drones by video camera. In this paper, we compare the existing computer vision methods such as background subtraction, frame differencing, optical flow and edge detection for object detection. In our work, we will use Convolutional Neural Network for both object detection and classification to enhance its performance.

Keywords— Unmanned Arial Vehicles, Computer Vision, Background Subtraction, Frame Differencing, Optical Flow, Edge detection, Convolutional Neural Networks.

I. INTRODUCTION

Flying object detection is done by extracting the moving objects from a video sequence. There are various methods have been proposed such as, the background subtraction, the frame differencing, optical flow method, and edge detection.

Background subtraction is a method which is widely used for detecting foreground objects from the background in video sequence. It is calculated by subtracting the current frame from the reference frame, called “background image”. Its performance always depends on the quality of background modeling.

The advantage of this method is fast and accurate detection and easy to implement. It is not applicable for moving camera because each and every frame will have a different background.

Frame differencing is used to extract the moving object which is based on calculating the difference between current frame and previous frame in video sequence. The advantages of the frame difference algorithm is quick to implement and highly adaptive to dynamic scene changes. relatively low computational complexity. However, it has some drawbacks, it is generally not an efficient approach for extracting all relevant pixels of moving regions and it is unable to find the internal pixels of enormous size, with uniformly distributed intensity values.

Optical flow method is a common method for objects detection. The optical flow method detects the objects based on the relative velocity of objects in the scene. Although this method adapts to moving objects detection, the complicated calculation makes it not applicable for real-time detection.

Edge detection method converts the original images into edge images by examining the changes of intensity values in the image. The edges are extracted from the video based on the features like edge, curve, straight lines and this is able to detect the objects.

II. RELATED WORKS

In paper[1], they discussed UAV detection framework is based on video images and they capture either by static camera or moving camera. Detection by static camera involves background subtraction. Moving camera based detection done by Region Proposal Neural Network. UAV is detected based on confident value of UAV detection. By using RPN, the number of false positive will be reduced which leads to be an early warning detection system. In future work, they will use temporal information to increase the accuracy.

Paper[2], focuses on detecting moving objects in image sequences produced by stationary camera. An edge tracker is a method which is used to extract edge traces on the drone. It is robust and fast method for detecting moving objects. Moving objects are detected by computing the difference between the edge-trace flows in consecutive frames. Pairs of consecutive images are processed and perceptual edge segments are found piece by piece.

YOLO detection method was discussed in [3] to detect objects. This unified architecture is ultimately fast and YOLO model processes images in real-time at 45 frames per second. Fast YOLO, can process an astonishing 155
When comparing with other detection systems, YOLO makes more localization errors but is less likely to predict false positives on background. In paper[4], the authors used convolutional neural networks (CNN) to detect the moving objects. By combining with background subtraction algorithm they have created the artificial dataset which contains the real images. By this approach we can get high precision and recall value at the same time. In future, they will use time domain to improve the performance.

In paper[4], the authors used convolutional neural networks (CNN) to detect the moving objects. By combining with background subtraction algorithm they have created the artificial dataset which contains the real images. By this approach we can get high precision and recall value at the same time. In future, they will use time domain to improve the performance.

CNN method which are capable of solving the optical flow estimation problem are discussed in [5]. Since existing sets are not sufficiently large to train a CNN, they generated a large synthetic Flying Chairs dataset. The networks trained on this paper is unrealistic data set. When comparing with Deep flow and Epic flow CNN performed well on this data set. In future work, they will try for realistic training data set.

Background subtraction method is handled in [6]. Objects are detected by PTZ camera. By using the data of two consecutive images previous and upcoming rotation new background will be created for the detection of moving objects. This method achieved the object detection for single moving object. In future they will choose a multiple moving object detection.

III. COMPARISON OF VARIOUS METHODS

<table>
<thead>
<tr>
<th>Reference</th>
<th>Methodology</th>
<th>Description</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>[1]</td>
<td>Background Subtraction + Region Proposal Network</td>
<td>Background Subtraction is used to identify moving objects. It is done by finding the difference between the current frame and background key model. Region Proposal Network consists of set of candidate regions that consists of object are produced for every individual frames.</td>
<td>Background Subtraction is only applicable for short sequence length. It is only suitable for static camera.</td>
</tr>
<tr>
<td>[2]</td>
<td>Edge Detection</td>
<td>Edge detection reduces the data and filters the useless information present in an image. It is implemented by GCS and CPP algorithm. GCS consists of curve points.</td>
<td>We must have prior knowledge about patterns o implement edge linking which is impossible in every moment.</td>
</tr>
<tr>
<td>[3]</td>
<td>YOLO Detection System</td>
<td>YOLO can process an images at the speed of 45 frames per second and Fast YOLO can process an image astonishing 155 frames per second.</td>
<td>YOLO unable to find small objects that appear in group like flocks of birds. It learn to predict bounding box from the data, it is not suitable to identify objects in unusual configuration.</td>
</tr>
<tr>
<td>[4]</td>
<td>Deep Networks</td>
<td>YOLOv2 attempt to design regression for object detection. Fully convolutional architecture are trained to extract the high level features in an image. It aims to find bounding boxes of an object in the video sequence directly, replacing the dealing with regions in the video sequence.</td>
<td>It require larger dataset that can include many scenario for high accuracy.</td>
</tr>
<tr>
<td>Reference</td>
<td>Methodology</td>
<td>Description</td>
<td>Limitations</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>[5]</td>
<td>Convolutional Neural Networks</td>
<td>They introduce correlation layer that works by multiplicative patch comparison in between feature maps.</td>
<td>It is only applicable for GPU.</td>
</tr>
<tr>
<td>[6]</td>
<td>Background Subtraction</td>
<td>PTZ (Pan, Tilt, Zoom) camera is used to detect the moving object because it provide broad vision and more information. Mean filter algorithm is used to build the initial background model.</td>
<td>It is not applicable for moving cameras due to initialization of background model.</td>
</tr>
<tr>
<td>[7]</td>
<td>Region Shrinking Algorithm</td>
<td>They introduced power transformation method before region shrinking to solve limitation of situation that there is a single object in an image.</td>
<td>___</td>
</tr>
<tr>
<td>[8]</td>
<td>Background Subtraction</td>
<td>The main role of background modelling is to track the changes in background scene and extract the foreground information.</td>
<td>It is unable to process the moving camera videos due to the background key frames.</td>
</tr>
<tr>
<td>[9]</td>
<td>Faster RCNN</td>
<td>In this method Caffe deep learning library is used to detect the objects. These are publically available for most of the models and it requires less amount of data set for deep learning system to train from basic. They have used ZF and VGG16 architecture to train their system.</td>
<td>They have used only drone dataset so there is a possible for false positive.</td>
</tr>
<tr>
<td>[10]</td>
<td>Background subtraction + frame differencing Algorithm</td>
<td>For getting background model its uses surendra background algorithm and uses frame differencing algorithm to update it, finally the moving object can be extracted from the images from these two methods.</td>
<td>It depending on The time interval between frames.</td>
</tr>
</tbody>
</table>

IV. BLOCK DIAGRAM

![Block Diagram]

V. SUMMARY

We have analysed the existing methods such as Background Subtraction, Frame Differencing, Optical Flow and Edge Detection algorithms for object detection. Background Subtraction is only suitable for static camera due to initialization of background model. Frame Differencing, Optical Flow and Edge Detection are applicable to both static and moving cameras. YOLO based detection can process an images at the speed of 45 frames per second and Fast YOLO can process an image at the speed of 155 frames per second, but Fast YOLO will not give an accurate result compared to YOLO detection method.

VI. CONCLUSION

The existing object detection methods such as Background Subtraction, Frame Differencing, Optical Flow, and Edge Detection algorithms are suitable for detecting objects flying at a particular distance, either in a clear or in a slightly blurred background. In our proposed UAV detection and classification system, we will use CNN both to detect and to classify the flying objects even at a far away distance and with unclear background. The performance of our system will be improved by training the network with sufficient data set.
REFERENCES


