IJARSCT

ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

IJARSCT
Impact Factor: 4.819

Volume 1, Issue 1, January 2021

Long Distance Human Detection in UAV Images Using Improved Faster R-CNN

Rajendra Prakash Ghorpade
Lecturer, Department of Electronics and Telecommunication Engineering
Bharati Vidyapeeth Institute of Technology, Navi Mumbai, India

Abstract: *Recently, using consumer Unmanned Aerial Vehicles (UAV) for aerial photography has become a trend. However, the images captured from the UAV raise a challenge to the existing pedestrian detection algorithms, because the humans in the image are too blur and too low-resolution resulted from the long distance between the UAV and pedestrians. The problem of detecting long distance humans in an image has always been overlooked, so even the performance of the state-of-the-art detection algorithms are not satisfactory when used on UAV pedestrian detection. In this paper, we improve the Faster R-CNN algorithm by proposing an improved Region Proposal Network (RPN) and utilizing object context information to improve the detection performance. The experimental results exhibits that the extended algorithm improves the performance of detecting pedestrians captured by UAV.*

Keywords: CNN, UAV, RPN (Regional Proposal Network), LCCNet

I. Introduction

Recently, most of peoples began to choose the UAV for out do or shooting activities. Compared to traditional capture methods, UAV gives us a new perspective to understand the world, but also brings a challenge to the video analysis algorithms such as pedestrian detection. This challenge is mainly resulted from the UAV shooting distance is generally faraway, resulting in the target in the image too blur and low resolution, which greatly disgrace the performance of detection algorithm. However, when we are increasingly dependent on the data provided by UAV, we must upgrade the performance of long distance human detection algorithm suited for UAV data. The state-of-the-art Faster R-CNN [1] and other algorithm such as Fast R-CNN[2], do work well when the target object covering large region and with high resolution. But when the method is used to detect human captured by UAV, it usually tends to fail. UAV pedestrian detection is more challenging than normal pedestriandetection. There are still few ways studying on the problem. pedestrian detection on UAV data has two main difficulties:

1. Humans are more difficult to locate due to a larger perspective.
2. Humans have less pixel information due to low-resolution, so their associated features obtained from the convolution network are less, which will increase the difficulty of classifying them. This paper extend the Faster R-CNN algorithm in two aspects to solve the two difficulties described above respectively. In the third section, we will describe in detail the improvements we use in our method.

II. Existing Systems

According to Hazim Shakhatreh et al. [1] UAVs are expect to be a outstanding deliverer of civil services in many fields including aerial photography, transportation, farming, disaster management and surveillance. The recent research literature on UAVs basis on vertical applications without considering the challenges facing UAVs within specific vertical domains and across application domains. Also, these studies do not discuss practical ways to solve challenges that have the prospective to contribute to multiple application domains.

Unmanned aerial vehicles (UAV) give us a new perspective to understand the world, but also bring a challenge to the video analysis algorithms such as pedestrian detection. This challenge is mainly resulted from the UAV shooting distance is generally faraway, resulting in the target in the image too blur and low- resolution, which greatly degrade

IJARSCT
ISSN (Online) 2581-9429
ISSN (Print) 2581-XXXX
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Impact Factor: 4.819
Volume 1, Issue 1, January 2021

the performance of detection algorithm. However, when we are increasingly dependent on the data provided by UAV. We must improve the performance of long distance human detection algorithm suited for UAV data.(R-CNN).

As a key use of image processing, object detection has boomed along with the exceptional improvement of Convolution Neural Network (CNN) and it upgrades since 2012. When CNN series expand to Faster Region with CNN(R-CNN). Humans rebound at an image and abruptly know what objects are in the image, where they are, and how they interact. If algorithms for image processing could be fast enough and accurate so the computers would be able to drive cars without specialized dependable devices and sensors, would be able to transmit real-time scene information to users. Thus, the basic tasks of image processing are a serial of recognition, classification, localization and object detection and the key challenges are cost, accuracy, and complexity and speed.

III. Proposed System Architecture

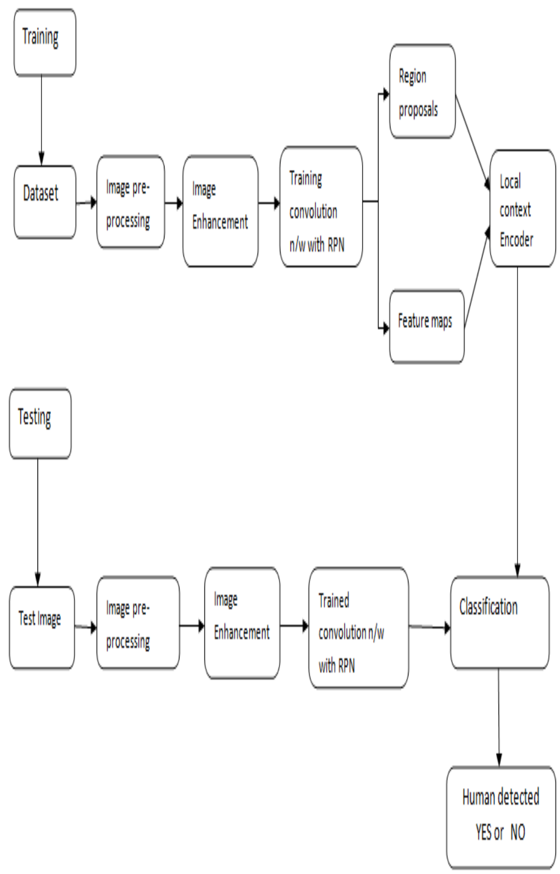

Figure 1: Architecture of the System

A) Image Pre-processing

Image Pre -processing means working on images in order to change it in a suitable form on so it can be trained and tested upon acceptable algorithm. These steps include subtracting a mean pixel value and scaling the image, in this stage the captured images will be cropped and be resized so that it can be effectively tested. In Digital image processing, pc algorithms are applied to perform image process on digital pictures. Pre-processing consists of many processes that include:

- Resize Image
- Filter Image

B) Image Enhancement

Image enhancement is the process in which a stored image is digitally handled to change its brightness contrast or of an image or manipulate the red-green-blue color patterns or grayscale of an image by using image enhancement tools. Image enhancement tools also contain actual de-blurring of images and other complex resources for restoring or clarifying images that may be in poor condition.

C) Convolutional Network with PRN

Inception-RPN sliding on the convolutional feature maps to predict the candidate target position. Then each candidate proposal generated by the RPN is expanded with context information to two proposals.

Inception-RPN Inspired by the proposal of the inception module in GoogLeNet[3], our proposed Inception-RPN based on inception module which consists of a different convolution, and max pooling layers, which is fully connected to the input feature maps. To search for region proposals, we slide our Inception-RPN over the top of convolution feature maps and associate a set of prior bounding boxes with each sliding position.

D) Local Context Encode Network (LCCNet)

It is fundamentally challenging to find pedestrians in long distance because there is little information can be obtained from the target object. Hence we must have to use image evidence beyond the object extent. We call this as context. Ramanan et al. [6] uses ground plane estimation as contextual features, which improves the performance on detecting small instances.

Our proposed work makes use of local context extended from original ROI. We show that context is very useful for finding low-resolution human. It is possible that the feature extracted from the proposal region contain insufficient information to correctly discriminate it.

E) Testing and Classification Phase

In the testing phase the test image are taken which is pre-processed, enhanced and then it is passed through trained convolutional network to extract region proposals. And classification phase takes output feature map produced by pooling layer passes a series of fully connected layers which can clearly determine whether it is a positive sample or a negative one for us humans, Loss function for an image is defined as:

$$L(\{pi\},\{ti\}) = \frac{1}{Ncls} \sum_i Lcls\,(pi, pi^*) + \frac{1}{Nreg} \sum_i Lreg\,(ti, ti^*)$$

G) Data

Our proposed method is going to use latest PASCAL VOC dataset which, the images in our dataset are collected from the Internet called as standard dataset and the pedestrian's relative areas are only 0.001 to 0.004. This corresponds to 20x45 to 35x85 pixel areas in an image.

H) Results

We will compare final output result of our proposed system with output result of previously designed system given below

1. Faster R-CNN.
2. Faster R-CNN + LCCNet.
3. Faster R-CNN + LCCNet + Inception-RPN.

We will also evaluate the performance of our proposed system in following terms.

1. Accuracy.
2. Sensitivity
3. False positive rate.

IV. Conclusion

In this paper, to resolve the problem of long distance pedestrian detection suited for UAV, we enhance the state-of-the art Faster R-CNN algorithm by Inception-RPN and LCCNet. We have prepared a dataset for research. Through detailed experimental verification, we found that the use of context information and Inception-RPN can improve the performance of detecting pedestrians in long distance captured by UAV.

References

[1]. G. McGraw and G. Morisett, "Attacking malicious code: A report to the Infosec Research Council," IEEE Software, 2000.

[2]. R. Benzmüller, "Malware trends 2017," GData Security Blog, 10 04 2017.

[3]. J. Yonts, "Attributes of Malicious Files," SANS Inst. InfoSec Read. Room, 2012.

[4]. J. Z. Kolter and M. A. Maloof, "Learning to Detect Malicious Executables in the Wild," in Proceedings of the International Conference on Knowledge Discovery and Data Mining, Seattle, WA, USA, 2004.

[5]. M. G. Schultz, E. Eskin, F. Zadok and S. J. Stolfo, "Data mining methods for detection of new malicious executables,", " in Proceedings 2001 IEEE Symposium on Security and Privacy. S&P 2001, Oakland, CA, 2001.

[6]. K. Nigam, A. McCallum, S. Thrun and T. Mitchell, "Learning to Classify Text from Labeled and Unlabled Documents," AAAI, 1998.

[7]. VX-Heavens: http://vxer.org/