IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

DISTRIBUTED VISUAL SEARCH USING DATA ANALYTICS

[1]Kavya J, [2]Deepika K

[1]PG Student, [2]Assistant Professor

[1]MCA,

[1]RV College of Engineering, Banglore, India

Abstract: Object detection effectively and accurately has been an important topic in the advancement of computer vision systems. The project aims to provide high precision identification of objects in real-time output, video streaming, and image. A comprehensively deep learning-based approach to solving an end-to-end object detection problem. In this case, the system input will be an image or a video, and the output will be a bounding box corresponding to all the objects in the image or video, together with the object class in each box

Index Terms **– Image/ video object detection, Real-time object, YOLO**

I. INTRODUCTION

Precise detection of objects has been a vital topic in advancing vision systems. The accuracy of object detection has increased drastically with the advent of deep learning techniques. The field of image/video acquisition is getting aided in the present days and storing devices have determined an amazing growth in the quantity of visual-audio content exchanged, transmitted, and shared over the internet. In the past years, the handiest method of searching statistics in multimedia databases was primarily based on textual annotation, which includes associating a collection of keywords to each individual item. Such a procedure requires an enormous quantity of human interplay and is intractable in the case of huge multimedia databases[1].

Since the human brain and visual system actively seeks for exciting regions through paying more attention to some elements. Visual saliency may be described as the perceptual fine that allows an object, human being, or pixel to stand out and apprehend a scene. visual attention is more given to the region of interest in images and interesting actions in video sequences

A major challenge in several object detection systems is that it will reassign classifier to perform detection depending on alternative vision techniques to serve the deep learning-based approach, resulting in slow and non-optimal performance. In this project, a completely deep learning-based approach to solve the classification problem has performed. A propensity to view target identification as a downside regression of spatially isolated bounding boxes and associated class probabilities. In one analysis the neural network predicts objects in bounding boxes with sophistication possibilities. Since the entire pipeline detection can be optimized directly on the detection performance in an end-to-end fashion in a single network[1].

II. RELATED WORK

In this paper[2]the author presents an approach to object detection which is used to replace linear convolution with ordinal convolution by using an efficient LSH scheme. for a variety of applications involving object detection methods they proposed this approach. Large empirical tests on DPM detectors have been conducted and the result showed that the system performs comparably to the original DPM detectors and they also mentioned about performance degrades gracefully as the number of object classes is increased, and the final output was up to 100,000 object classes can be simultaneously detected on a single machine in under 20 seconds.

The paper [3] discusses the video segmentation and the Background subtraction methods. Background subtraction method, is used to identify the moving object in the video stream. Background subtraction finds moving objects information by subtracting the background model. Performing the segmentation of the video will help in identifying various objects and their features present. And The comparison between Background subtraction and segmentation is performed.

In this paper[4] a spatiotemporal attention detection framework. is proposed for automatic detection and tracking of salient objects in video streams. within the salient object detection the more concentration is given on the objects instead of its surrounding. The video is segmented into shots supported scale-space filtering graph partition method. for every detected shot the associated static summary is developed employing a leap keyframe extraction method. this is often a bottom-up approach and applied recursively RANSAC algorithm.

In this paper[5], a video-based objection detection method is proposed for the traceability system with a deep learning method. The surveillance video is collected first, from which an annotated image database of target objects such as people or vehicles was constructed to train convolutional neural network model off-line. With the trained model, real-time target detection and recognition system are designed and implemented. The proposed method mainly includes three aspects: video processing, target detection, and object recognition. It provides a variety of video interfaces to support the downloaded video and real-time video stream. The experimental results indicate that the proposed deep learning-based detection method is efficient for the traceability application.

In this paper, [6] the author describes to the latter philosophy and propose to train a detector, called "DeepMultiBox", which generates a small number of bounding boxes as object candidates. These boxes are created in a class agnostic manner by a single Deep Neural

Network (DNN) using a single Deep Neural Network, which generates the boxes. Localization of artifacts using an image approach that uses a deep convolutionary neural network that predicts several boundary boxes at a time as an abstraction and learning model base function. It formulates a multiple box localization cost, which can take advantage of a variable number of ground truth locations of interest in a given image and learn to predate

III. PROPOSED METHODOLOGY

The main objective of this project is to detect the object present in the video or image with high accuracy. Here the algorithm used is YOLO algorithm which is the best algorithm when compared to CNN in terms of accuracy. Many of object detection systems which have to be compelled to undergo the image over just once to be able to detect all the objects within the image, or it has to undergo two stages to detect the objects. YOLO doesn't need to go through these processes as the name itself brief that YOLO needs to look at the image/video only once to recognize the object in the image/video. To train the model the related dataset has been collected from "Kaggle". The images are trained in Darknet which is an open-source framework. The algorithm is based on regression. Initially pretrained weights file should be present, loads the weights file. Layer by layer the configuration file and the weights are performed. Finally after the completion of training the model the dataset will be passed and tested.

To define the model the information needed is

- configuration file (*.cfg) which contains the details of the model
- pre-trained weight file should be loaded
- In order to train the batch how many data is needed
- number of iterations to train
- training a dataset
- directory to store the annotation files
- directory called dataset where the image files are stored

IV. CHALLENGES

From a significant test the issue is that of the variable component of the output which is caused because of the variable number of items that can be available in some random info picture or video. Any broad machine learning task requires a fixed element of information and output for the model to be prepared. Another important obstacle for the widespread adoption of object detection systems is the requirement of real-time while being precise in identification. The more unpredictable the model is, the additional time it requires for inference; and the less intricate the model is, the less is the exactness. This exchange off among exactness and execution should be picked according to the application. The issue includes classification as well as regression, driving the model to be adapted all the while. This adds to the intricacy of the issue.

V. RESULT

In this project a set of objects has been trained under the video image and real-time objects. This project has three modules called real-time object detection module, image object detection module and video object detection module.

Figure 1

The above Figure is the output screen from the video where the model has detected the object that is person in the video and text has been displayed on top of the bounding box

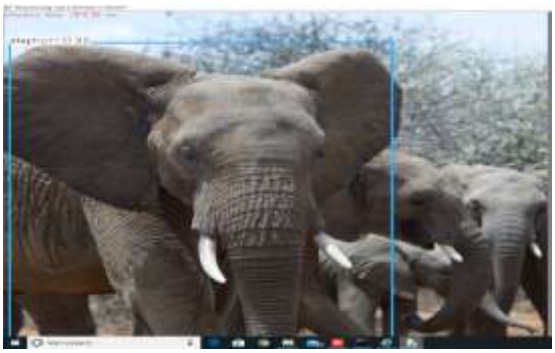

Figure 2

The above Figure is the output screen from the image where the model has detected the object in the image and text has been displayed on top of the bounding box

Figure 3

The above Figure is the output screen from the real-time object detection. where the model has detected the object in the image and text has been displayed on top of the bounding box.

VI. CONCLUSION

Results showed that the proposed technique is able to extract moving objects from different sequences with success. Sometimes the boundaries of the extracted object are not sufficiently accurate to place them in various scenes that require an almost perfect boundary location. Developing a model was challenging, and can be trained on complete images directly. Unlike classifier-based approaches, YOLO a unified Object Detection Model is trained on a loss function that corresponds directly to the detection performance and the whole model is trained together. The project result is the target has been defined with high precision in video, images, and in real-time.

REFERENCES

[1]. Chengji Liu, Yufan Tao, Jiawei Liang, Kai Li, Yihang Chen, "Object Detection Based on YOLO Network", 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC)
[2]. Dean, T., Ruzon, M. A., Segal, M., Shlens, J., Vijayanarasimhan, S., & Yagnik, J," Fast, Accurate Detection of 100,000 Object Classes on a Single Machine", 2013 IEEE Conference on Computer Vision and Pattern Recognition
[3]. Mohan, A. S., & Resmi, "Video image processing for moving object detection and segmentation using background subtraction". 2014 First International Conference on Computational Systems and Communications (ICCSC). IEEE
[4]. Tapu, R., Mocanu, B., & Tapu, E. (2012)." Salient object detection in video streams", 2012 10th International Symposium on Electronics and Telecommunications
[5]. Sobti, A., Arora, C., & Balakrishnan, M. (2018), "Object Detection in Real-Time Systems: Going Beyond Precision", 2018 IEEE Winter Conference on Applications of Computer Vision (WACV).
[6]. Erhan, D., Szegedy, C., Toshev, A., & Anguelov, " Scalable Object Detection Using Deep Neural Networks", 2014 IEEE Conference on Computer Vision and Pattern Recognition.
[7]. Koay, S. Y., Ramli, A. R., Lew, Y. P., Prakash, V., & Ali, R. (n.d.), " A motion region estimation technique for Web camera application", Student Conference on Research and Development. IEEE
[8]. Cucchiara, R., Piccardi, M., & Mello, P. (n.d.)," Image analysis and rule-based reasoning for a traffic monitoring system", Proceedings 199 IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems
[9]. Haritaoglu, I., Harwood, D., & Davis, L. S,"real-time surveillance of people and their activities", IEEE Transactions on Pattern Analysis and Machine Intelligence
[10]. Cucchiara, R., Grana, C., Piccardi, M., & Prati, A. (n.d.), " Statistic and knowledge-based moving object detection in traffic scenes", 2000 IEEE Intelligent Transportation Systems. Proceedings
[11]. Lau, B., Arras, K. O., & Burgard, W, "Tracking groups of people with a multi-model hypothesis tracker", 2009 IEEE International Conference on Robotics and Automation.
[12]. Tian, B., Li, L., Qu, Y., & Yan, L. (2017), " Video Object Detection for Tractability with Deep Learning Method", 2017 Fifth International Conference on Advanced Cloud and Big Data (CBD).