March - April 2020
ISSN: 0193-4120 Page No. 2064 - 2069

Moving Object Detection Using Tensor Flow

Archana Rajesh Date
Department of Electronics and Telecommunication, Sinhgad College of engineering Pune, India
archanadate@gmail.com

Sanjeevani Kiran Shah
Department of Electronics and Telecommunication Smt. KashibaiNavale College of engineering Pune, India
sanshah2608@gmail.com

Article Info
Volume 83
Page Number: 2064 - 2069
Publication Issue:
March - April 2020

Article History
Article Received*: 24 July 2019*
Revised*: 12 September 2019*
Accepted*: 15 February 2020*
Publication*: 18 March 2020*

Abstract

In video surveillance, detection of moving object is crucial and basic task as its output is an input to the next-level such as tracking and behavior analysis. Detecting camouflage object in a video is the biggest challenge. Computer vision methods are already largely used. With recent development in deep Learning, development of object detection application has become easier. TensorFlow's Object Detection API is an open source model that is developed and pretrained on the datasets. Thus, in this article for detecting camouflage object under different environment tensor flow is used. Quantitative results were obtained on standard datasets and compared with the other algorithms. TensorFlow gives excellent results as compared to other algorithms.

Keywords – *Object, video, camouflage, moving, tensor-flow..*

I. Introduction

Initial step in any video surveillance system is the detection of a moving object. Moving object detection seperates moving target objects from the background objects. It also gives the position of the single or multiple objects which is helpful for further analysis. It mainly used in applications like video surveillance, video retrieval and driver assistance systems. It is the dynamic research topic with various challenging situations like dynamic background, bootstrap, occlusion, illumination changes, shadows etc.[1]. Camouflage is the very big challenge as the object and background share similar color and hence are difficult to differentiate.Camouflage concept is very useful for the soldiers in military.Camouflaging is very useful to the soldiers as it help them to protect themselves against enemies. The current scenario in video surveillance systems is to detect motion by taking into consideration the information about each pixel. Gray level and color of the pixel is considered while detection. Once the object is detected, it serves as an input to different tracking

and behavioral analysis system. Hence detection of camouflage moving object in an video has become an important field of research. Though Camouflage appears to a simple challenge, the moving camouflage objects are very difficult to detect and hence many more methods have to be invented.

Animals and insects make use of camouflage to conceal from their enemy. When the object is moving it may hide itself in the similar color background. It is said to be motion camouflage. Soldiers in the bbattlefield wear similar texture patterns clothes as that of the background to conceal themselves from the enemy. In case of animals it is known as natural camouflage whereas others are the examples of artificial camouflage.

As many developments were made in neural network, it becomes easy with the help of neural

March - April 2020
ISSN: 0193-4120 Page No. 2064 - 2069

network to develop the algorithm for moving object detection. In this article, TensorFlow application programmable interface was tested for detection of moving humans. The model used within the application programmable interface was a single shot detector with Mobile net. The pretrained model was used and tested with standard datasets for object detection.

II. LITERATURE REVIEW

In computer vision object detection is nothing but searching, finding position and classifying the object into different classes. Initially, the parametric background models were created by averaging the background frames and then current frame and background is compared for object detection. This method overcomes the changes in illumination and also noise but gives poor results for variations in background regions. To overcome this problem Gaussian Mixture Models (GMM) [2] were introduced which models each multimodal background. In this each pixel is modeled on their color intensities using mixture of gaussian. The probability density functions are used. Advancement in GMM were introduced later [3]– [6] that models' various dynamic elements in the background. Kernel Density Estimation (KDE) [7] is more advanced than other methods [8], [9] as it does not depend on any parameter. In Nonparametric model's background probability density functions is calculated directly from current pixel intensity values at each pixel point.

Nowadays the mostly used object detection methods are based on convolutional neural networks [10]. Mostly used methods are RCNN [11], fast RCNN [12], Faster R-CNN [13], R-FCN, Multibox Single Shot Detector (SSD) and YOLO (You Only Look Once). Basic R-CNN system executes a neural net classifier on the samples taken from images by drawing different rectangle boxes on it (features are extracted from all the cropped samples). This method was computationally expensive as many samples are required. Fast RCNN decreases the overhead by doing feature extraction only once from the overall image and then warping is done. In case of Faster RCNN extracted features are used to create different proposals on a single layer. In R-FCN unlike Faster R-CNN feature warping is done with the help of various layers and efficiency is increased [13]. YOLO (You Only Look Once) method uses a single convolutional network on the input (once) and gives rectangular boxes along with the class classification percentage for every class parallelly. Thus, the YOLO model is fast in comparison with above explained method and it generalize representation of all objects. YOLO gives bad results for the images having varying dimensions or group of small objects but advantage is increased true positive rate [14]. SSD [15] is used in real time applications due to their higher speed of operation. But precision goes on reducing. Many new detection methods require high speed for detecting an object. Methods such as You look only once or single shot detector has high speed of operation, but accuracy of prediction goes on decreasing. Methods such has, faster R-CNN gives best perfection but costly in terms of computation.

III. METHODOLOGY

In Object Detection, algorithm uses basic concept of feature extraction from the input frame. The features that are extracted used for classifications of objects into different classes. Deep Learning is an art of the learning with the help of methods, which try to show huge level reflections in information, using many functioning layers. These methods are based on two basic concepts, the neural network and the algorithm that does backward propagation of error.

The very first system tries to map the human brain structure with codes and data structure union. The prototype is first trained on a very huge quantity of information. It then learns the relationship between the data. The prototype is then then tested on the functions that are learned such as detecting an object. The construction of the system is generally consisting of three firm layer structures: Preprocessing, middle and Output Layer. The layer that actually does functioning are middle layers, which can be in any number. The length of the model depends on the number of these middle layers. Depending on number of middle layers it is known as deep.

Then the error is back propagated through the network that trains the model so that the error can be reduced and results are improved. The error function is calculated by using gradient descent method with respect to weights. These is then passed to the next step that updated all the weights in the network and decreases error.

TensorFlow is Google's Open Source Machine Learning [16] Framework for dataflow programming for various of tasks. Vertices gives algebraic operations, and the lines act as the complex data arrays (tensors) which are used for communication between vertices. Tensors are just complex arrays, a supplement of 2-dimensional data tables to higher size. Due to different features of tensor flow it is mainly appropriate for deep learning.

TensorFlow Object Detection is an application programmable interface that built, train and deploy the object detection structures [16]. It gives user a readymade model trained on various object detection datasets using machine instructions and program cipher and hence the models can be used for object detection job. In this work, a single shot detector model with Mobilenet (ssd_mobilenet_v1_coco) was chosen. TensorFlow has been trained using MSCOCO Dataset [17]. In this dataset there are 2.5 million instances of data in 328000 images, and 91 different object classes s such as "person", "car" or "cat" .

Algorithm

1. Import necessary libraries.
2.Initalise detection graph and load configuration from trained model.
3. Initialse input and output tensors.
4. Read image.
5.Run tensor flow session.
6. Extract object location.
7. Get ground truth location.
8. Compare and decide.

IV. Results and Discussions

The algorithm is tested on standard dataset for quantitative analysis. Standard databases containing camouflage are used to test the developed algorithm. The database used are changedetection.net and LASIESTA (Labeled and annotated sequences for integral evaluation of segmentation algorithms).

A.Change Detection.net (CD.NET)

Changdetection.net provides video databases with various challenges so that the developed algorithms can be tested for object detection. Two datasets are available: 2012 Dataset and 2014 Dataset. The sequences are taken in both indoor and outdoor environment. Thus, these datasets cover various detection in different scenarios. Dataset 2012 consists of videos for baseline, dynamic background, and camera jitter, thermal, shadow and intermittent moving object. Out of this from baseline category high way sequence is selected and from intermittent object motion Sofa sequences is selected. Sofa consists of 2750 frames with size 320x240 and highway contains 1700 frames with size 320x240.

B.LASIESTA

LASIESTA is composed by huge range of data sequences which are taken in indoor and outdoor environment. These are divided into various categories. Every category covers different challenges in moving object detection. Every sequence is completely annotated at both pixel and region level. Thus, it is appropriate for all the challenges that occurs during moving objects detection.

1. Qualitative Analysis

TEST Engineering & Management

March - April 2020
ISSN: 0193-4120 Page No. 2064 - 2069

1. ICA_O1 sequence

(a) (b) (c)

2. ICA_O2 sequence

(a) (b) (c)

3. Sofa sequence

(a)

(b) (c)

4. Highway sequence

(a) (b) (c)

Fig. 1:(a) Input frame (b) Ground-truth frame (c)Output

Above figure shows the qualitative analysis of four sequences namely I_CA_01, I_CA_02 which are from LASIESTA database and Sofa, highway sequences are from changedetection.net database. Starting from left first is the input frame, ground truth frame and output frame. In all four sequences tensor-flow detects camouflage object completely with the help of rectangular box.

2. Quantitative Analysis

The parameters have been evaluated for performance analysis. Performance evaluation tells how correctly the method detects the moving target with less faulty detections.

1. Recall (Detection rate) compares the percentage of pixels that are classified correctly as foreground to the entire number of foreground pixels that are present in the ground truth:

$$Recall = \frac{Number\ of\ correctly\ identifying\ pixel}{Number\ of\ foreground\ pixels\ in\ ground\ truth.}$$

2. Precision compares the percentage of the pixels that are correctly classified as foreground by the algorithm with the total pixels classified as foreground by the method:

$$Precision = \frac{Number\ of\ correctly\ identifying\ pixel}{Number\ of\ foreground\ pixels\ detected\ by\ algorithm}$$

A good performance is obtained when the Detection Rate is high without altering the Precision.

3. The F-Measure is calculated as given below:

$$F\ measure = \frac{2 * recall * precision}{recall + precision}$$

The F-Measure denotes the achievement of detection in Precision-Detection Rate space. The maximum value of is one.

Table 1: Quantitative analysis

Database	Precision	Recall	F measure
I_CA_01	0.9263	0.9670	0.94623
I_CA_02	0.8341	0.9583	0.8919
Sofa	0.9804	0.9719	0.9144
Highway	0.9411	0.8986	0.8020

As seen from the results Tensor-flow gives best results for detecting camouflage object. Recall is near about 98%, precision also 98% and f-measure near about 96 %.

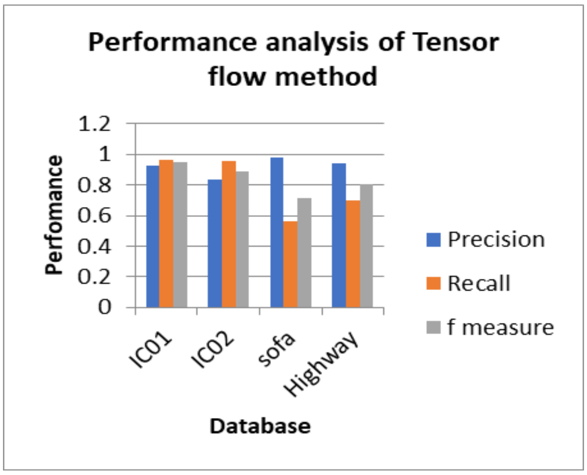

Fig. 2: Graph of quantitative measures

V. Conclusion

Camouflaged moving object is detected by implementing tensor flow application programmable interface. The testing of implemented algorithm is done in different environment which consists of indoor as well as outdoor camouflage sequences from standard datasets. Both quantitative and qualitative results were obtained and compared. Tensor-flow method gives very good accuracy of 95%. Thus, from above it is cleared that tensor-flow approach gives very good results for detecting camouflaged moving object.

VI. Acknowledgment

I express my sincere thanks to Dr. Sanjeevani S. Shah for her valuable advice, excellent guidance, support, encouragement and confidence towards me.

VII. References

[1] K. Toyama, J. Krumm, B. Brumitt and B. Meyers, “Wallflower: Principles and practice of background maintenance,” Proc. IEEE Int’l Conf.Computer Vision, vol. 1, pp. 255 - 261, 1999.

March - April 2020
ISSN: 0193-4120 Page No. 2064 - 2069

[2] S. Brutzer, B. Hoferlin and G. Heidemann," Evaluation of background subtraction techniques for video surveillance," Proc. IEEE International Conference on Computer Vision., pp. 1937 - 1944, Nov. 2011.
[3] C. Stauffer and W. E. L. Grimsson, "Adaptive background mixture models for real-time tracking," in Proc. IEEE Conf. Computer. Vis. Pattern Recognition., vol. 2, 1999, pp. –252 Vol. 2.
[4] Z. Zivkovic, "Improved adaptive gaussian mixture model for background subtraction," in Proc. IEEE Int. Conf. Pattern Recognition., vol. 2, 2004, pp. 28–31 Vol.2.
[5] Z. Zivkovic and F. van der Heijden, "Efficient adaptive density estimation per image pixel for the task of background subtraction," Pattern Recognition. Lett., vol. 27, no. 7, pp. 773–780, 2006.
[6] D.-S. Lee, "Effective gaussian mixture learning for video background subtraction," IEEE Trans. Pattern Anal. Mach. Intel., vol. 27, no. 5, pp.827–832, May 2005.
[7] A. M. Elgammal, D. Harwood, and L. S. Davis, "Non-parametric model for background subtraction," in Proc. 6th European Conf. on Computer. Vis., 2000, pp. 751–767.
[8] A. Mittal and N. Paragios, "Motion-based background subtraction using adaptive kernel density estimation," in Proc. IEEE Conf. Computer. Vis. Pattern Recognition., vol. 2, 2004, pp. II–302–II–309 Vol.2.
[9] Y. Sheikh and M. Shah, "Bayesian modeling of dynamic scenes for object detection," IEEE Trans. Pattern Anal. Mach. Intel., vol. 27, no. 11, pp. 1778–1792, Nov 2005.
[10] A. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification with deep convolutional neural networks. In NIPS, 2012.
[11] R. Girshick, J. Donahue, T. Darrell, and J. "Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In CVPR, 2014".
[12] R. Girshick, "Fast R-CNN," in IEEE International Conference on Computer Vision (ICCV), 2015.
[13] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", TPAMI 2016.
[14] Redmon J, Divvala S, Girshick R & Farhadi A (2016) You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.: 779-788.
[15] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C & Berg AC (2016) SSD: Single shot multibox detector. European conference on computer vision., Springer: 21-37.
[16] The TensorFlow Authors (2017) TensorFlow Object Detection API (readme). URI: https://github.com/TensorFlow/models/tree/master/research/object_detection. Cited November 23, 2017.
[17] Lin T, Maire M, Belongie S, Hays J, Perona P, Ramanan D, Dollár P & Zitnick CL (2014) Microsoft coco: Common objects in context. European conference on computer vision., Springer: 740-755.