Object Detection Techniques in Videos

Kusuma S., M. V. Sudhamani

Abstract: Object detection in videos has increased its popularity because of its wider applications. It has gained more research attention now days as it is applicable in real time situations like pedestrian detection, anomaly detection, Self moving cars, sports, counting of people etc. This paper begins with the introduction of object detection and briefs the basic steps in the process. It also provides a review of various techniques and approaches used for object detection in videos. Discussion of every approach and limitations will provide several promising directions and guidelines for future work.

I. INTRODUCTION

In recent years, object detection in video applications has become broadly developing research domain such as video-surveillance, sports and traffic surveil- lance etc. any application may contain static and moving objects. The main task is to identify the objects and the physical displacement of a moving object in area. This can be determined by using different techniques.

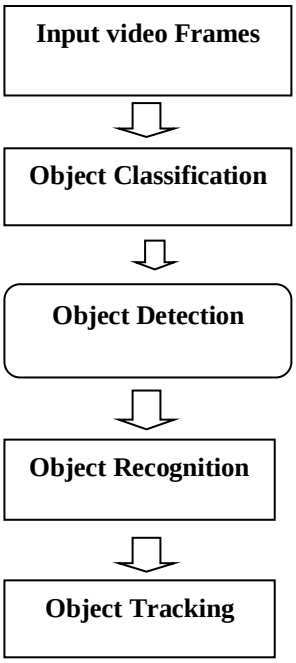

Figure 1. The basic flow diagram of Object Detection System

Earlier methods proposed by many researchers were more useful as well as has some disadvantages. The objective of this paper is to provide an idea of the various approaches in detecting the objects by specifying both advantages and disadvantages.

Revised Manuscript Received on December 15, 2019.
Mrs. Kusuma S., Assistant Professor, Dept. of ISE, RNSIT.
Dr. M. V. Sudhamani, Professor and HoD, Dept. of ISE, RNSIT.

Any video scene contains objects that can be determined by object detection technique which is called as object detection. Object tracking also very useful technique which is used to locate the object which is available in each frame for different applications.

Identified objects can belong into various classes. There are such as animals, cars, ball and other moving objects. In any video sequence, detection of an object plays a very prominent role.

The basic steps in object detection is shown in Figure 1, which receives the video frames as an input, classify the input image objects in to particular group, then apply object detection techniques for detecting object of our interest and recognize the object and track the object for its behavior as a final step. Various methods in each main step can be adopted for the varied criterions to gain accuracy.

II. PROCESS OF OBJECT DETECTION

One of the understandings on object detection is to find the object in the selected/inputted video frames. In this process, firstly, where the object is found need to located and next finding out the group to which it belongs to, should be categorized. As the role of this process is to find where the objects are found and category of each object, there are in general three steps specified for this process as shown in Figure 2. These three steps are: Region or location selection, Feature extraction, and classification.

A. Region /location selection: The object of interest may be seen in any region or location inside the image. That object can vary in its size, shape, color etc

B. Feature extraction: Here, appearance of object and its related attributes need to be identified for achieving the robustness in object detection. Describing the perfect features of object in diverse applications based on attributes like illumination, shape etc. is very difficult.

C. Classification: use of a best classifier helps to categorize the objects among all other existing objects in the video frame.

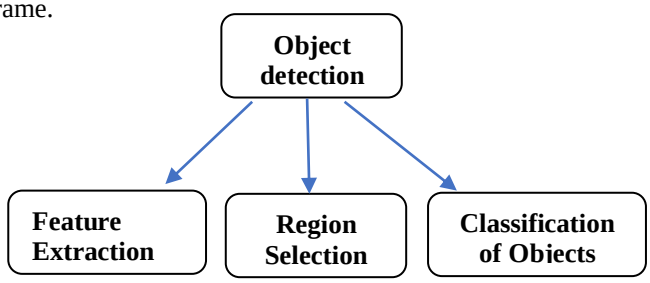

Figure 2. General Steps in Object Detection

III. LITERATURE REVIEW

Researchers over the years have done lot of contribution towards object detection in videos. Research directions are varied in object detection itself with respect to so many attributes like identification of single object or multiple objects, object of interest depending on the application, object varying with its nature like size, shape, color etc. Positioning of the object in a scene, illumination etc. such factors affect object identification accuracy. Always detecting object has been challenging considering above factors. paper focuses on identifying the major contributions in object detection, which further helps to understand the drawbacks of the existing methodologies and lays foundation for future research.

In [1], object detection work is done on the shape fragments-based model for complex environment. This paper also proposes a framework of two-stages for object detection. This is mainly because this method is based on the resident shape patches without conditions covering the key point instead of common contour groups. Also, object center image was successfully achieved.

In [2], detail survey on object detection is proposed. Comparision of the various techniques and its features are discussed here. According to that the feature based approaches are not so effective in detection of many objects and also it affects on the color aspect also. It supports the research by its rich survey and references.

In [3], the focus is on classifying the input images and finds out the object concepts and locations. Different subtasks associated in object detection listing are done here, which are face detection, pedestrian detection and skeleton detection. It provides the useful information for understanding the semantics of images and videos belonging to many applications. This paper also helps researchers to understand an object detection based on deep learning- that consider varied sub problems, such as obstruction, disorder, and low resolution, with changed degrees of modifications on R-CNN. This review also provides a useful insight on behavior analysis, face recognition and self-directed driving.

In [4], the different methods of object detection as well as tracking in the video frames through different phases were analyzed. Here, a review of different object detection, tracking, recognition techniques and feature descriptors, segmentation methods are described which are based on the various detection and tracking technologies.

In [5], the various techniques for object detection and tracking are listed. Some of the techniques are background subtraction, modelling, etc. This paper discusses the simulated result showing the effective comparison of effective object detection and states that has better correctness by taking less processing time compared to existing methods. For moving object this method uses the complex wavelet transform domain. This comparing gives the best result with all the above-mentioned techniques.

In [6], discussion on the approaches in video and classification of approaches in to different categories are achieved. It states that recurring target detection is hard to find with template-based method. the still image shape template is one of the major restrictions till now. classifier will not locate object properly because of the classification error. An approach for motion is very active for object detection. It is observed so far, researchers are working on single object.

In [7], a method based on the deep learning for objection detection in videos for traceability system is achieved. This method also specifies details on video processing, detection and recognition of objects. The major contribution of this paper is that it provides a diversity of video interfaces which assist for downloading any video and realtime video frames. The results of this paper indicate that this method is quite effective for the application based on traceability.

In [8], a review of several methods used in video investigation applications for threat detection and classification is proposed. Usage of these methods in wireless networks helps to reduce energy used by the devices, as it decreases the quantity of data transmitted through the network. And also, the focus is on methods like SVM using the shape-based features and kNN along with SURF features that are used in both applications. A good solution to reduce the amount transmitted wireless sensor network data can also be achieved and hence delaying the network lifetime.

In [9], discussion on the steps of multi camera video analytics system for detection across multi-camera, and re identification is done. It proves that in smart city video analytics can be established according to the correctness of society such as intelligent surveillance smart parking, traffic monitoring, vehicle navigation, smart healthcare etc. This paper analyzes and reviews various approaches on multiple object detection and tracking.

In [10], it is stated that challenging in detecting objects is due to various issues. It analyses the same based on application to any specific smart city video analytics system. This work also recommends many object tracking algorithms to handle above mentioned issues. Finally, it states that occlusion can be overcome with multiple cameras.

In [11], The work explored on deep learning technique reputation which is built on CNN for object detection methods. Different techniques for object detection systems are also compared and assessed here.

Future research directions are also given here. As human beings have limitations to handle large surveillance data, need to bring data closer. This might help into real time object identification.

In existing approach this have 1-20 nodes of groups having GPUs. Systems need to be expanded to deal with realtime video producing frames at 30 - 60 per second.

The other issue addressed here is to know how to incorporate processing into a centralized, robust GPU for processing data obtained from various servers simultaneously and performs near real time detection analysis. large datasets over the size of 100 terabytes are essential. 100 million images or more is required to guide the self driven cars.

In [12], discussion on the Object detection and its advancement based on Convolution Neural Network (CNN) and its alternatives is done and states that the quick necessity of object detection upgrading is to speed up the speed. This paper also exhibits You Only Look Once (YOLO), which gives the new direction to the CNN group and supports for the new mode of solving the object detection problem with most effective way.

In [13], a successful way is described to take the advantage of sequential contexts by connecting the same object to form tubelets is proposed in videos. And also summing of classification scores in the tubelets is done. In this paper, the focus is on finding excellence object linking results for improved classification.

In [14], methods for the improvement by accelerating the speed are presented. Based on the contextual and the core solution CNN. This paper proposes method of CNN methods You Only Look Once (YOLO). Its fastest speeds have achieved the exciting unparalleled result., YOLOv2 is the new version, achieves a best between accuracy.

The goal of paper [15], is to survey a recent work done and to analyze the open challenges in visual surveillance systems. This paper also describes the moving object detection and processing of videos approaches. Different methods for motion detection including background subtraction, frame differencing and optical flow methods are mentioned here. Also, an overview of recent developments in human motion analysis is also described. And it suggests that the statistical methods are better choice in more unrestrained situations and shows the future direction as a need for more reckoning time to get better results.

Paper [16], delves in to vision-enabled autonomous systems such as robots and autonomous cars. Speed and accuracy are the two aspects to address in image scaling. This paper results have achieved re-scaling of the image to a lower resolution, which will sometimes produce better accuracy. An approach of dubbed Ada Scale, which selects the input image improves both accuracy and speed for video object detection.

This paper contributions are three-folded: (i) the use of down-sampled images demonstration is done for improving both speed and accuracy for video object detection (ii) provided consolidated results that demonstrate improvement in both ImageNet VID as well as mini YouTube-BB datasets (iii) combination of proposed technique with video object detection acceleration techniques is done to further improve the speed by an additional 25% and benefit also of slightly higher accuracy.

In [17], the activist in the detection of future object pattern in video object is discussed. While the Long Short-Term Memory (LSTM) cannot necessarily model object connection between successive frames. So, this paper, propose the group LSTM. This technique reverts and classifies directly on object locations and categories of each output object.

Association of objects in two consecutive frames was also addressed by minimizing the matching error Additionally, this method works in an online style, which will be useful for video tasks. This approach outperforms on standard video datasets efficiently. A Long Short-Term Memory (LSTM), plays an important role in deep RNN architecture as it has the capability to admittance long-range context in sequence.

In [18], application-based object detection method is proposed. Application chosen is of underwater sonar images. It lists the main difficulties like object sequence, misunderstanding from false targets and complicated backgrounds. And also, the key problems for underwater object recognition and proposes prior knowledge based on transformable template matching for underwater object recognition. It also locates the objective state in the objective image through the fast-unique detection techniques based on FFT, which will avoid finding the entire image for the objec

IV. RESULTS AND DISCUSSIONS

Authors	Techniques	Results	Limitations
Chen Guodong ZeyangXia, et al., 2011, [1]	Model based Object Detection.	Achieved the Comparison of the shape fragments of detected image to the codebook.	Background clutter need to be detected for better accuracy.
Sanjivani Shantaiya et al 2013, [2]	Object detection approaches (Shape, color, Motion, Frame difference)	Single object research is only given prominence is identified.	Multiple objects detection on feature-based approaches are not effective.
S. Rosi, W. Thamba Meshach, et al, 2014, [3]	Different methods of moving object detection and tracking.	Method called wavelet transform domain is preferred for moving object detection for gaining accuracy.	Accuracy can be more in moving object detection.
Muhsin civelek et al , 2017, [4]	Methods in wired and wireless wireless surveillance networks using SVM.	Shape-based features are quite satisfactory for wireless multimedia.	The results that only HOG will not be appropriate for this type of an application.
Yongyi Lu HKUST, et al, IEEE, 2017 [6]	The method for association LSTM.	Method runs on twelve FPS on both datasets in the testing stage. The result is almost real time.	CNN structures are not updated in response to the outputs of the LSTM.
BingTian,Liang Li,IEEE 2017 [7]	Deep learning based new method for traceability system.	Introduce two video collection services, downloading video from the NVR servers and the other is real-time live video stream.	Need to find a method of target deception for motion detection
Fancy Joy, et al , 2018, [9]	New Mask R-CNN method.	More appropriate for footages of sports teams compared to YOLO.	Inability to reliably detect the sports ball. Existence and position of an object Shadow is not considered much.
Rahul Dutt Sharma1, et al, 2018 , [10]	Multiple object tracking	Many tracking algorithms are proposed.	The surveillance area can be extended with multiple cameras.
Elena Luna , et al, 2018, [11]	Various methods, background subtraction, frame differencing and optical flow are proposed for motion detection.	The statistical methods are better in more free situations.	Takes of more computation time.
AjeetRam Pathaka, et al, ICCIDS 2018 [12]	Abandoned Object Detection (AOD)	The impact of each stage on the AOD performance is discussed.	Hand-crafted feature-based approaches are less efficient, in complex scenarios.
OPEN ACCESS,Science Direct Journal, 2018. [14]	New object detection frameworks.	Salient object detection, face detection, and pedestrian detection are also briefly reviewed.	Detection accuracy varies for application
Zhong-QiuZhao, IEEE 2019. [15]	Frameworks on modifications on R-CNN.	Promising future directions in neural networks and in other learning systems	Sub problems handled applications are limited.
Ting-WuChin,et al, IEEE 2019. [16]	Method towards rea time object detection Adaptive scaling	improving both speed and accuracy in video object detection with adaptive scaling.	Accuracy need to be improved.

V.CONCLUSION AND FUTURE SCOPE

The paper presents an elaborated discussion on already used techniques by various authors about object detection in videos and tracking of object in various approaches.
This paper helps to understand the details about the work carried out and yet to be used techniques for further improvement in object detection.

Various new hybrid approaches can be developed by using the ideas discussed in this paper. Research scope is to improve more accuracy inspite of many affecting attributes, consideration of single or multiple objects and application

specific object detection for achieving real time results without much delay.

REFERENCES

1. Chen Guodong , ZeyangXia RongchuanSun, ZhenhuaWang, ZhiwuRen and LiningSun, "A Learning Algorithm for Model based Object Detection", The 8th International Conference on Ubiquitous IEEE, 2011.
2. Sanjivani Shantaiya Keshri Verma Kamal Mehta , "A Survey on Approaches of Object Detection", International Journal of Computer Applications (0975 – 8887) Volume 65– No.18, March 2013.
3. S. Rosi, W. Thamba Meshach, J.Surya Prakash, "A Survey on Object detection and Object tracking in Videos " ,International Journal of Scientific and Research Publications, Volume 4, Issue 11, November 2014 .
4. Muhsin civelek and Adnan yazici," Object Extraction and Classification in Video Surveillance Applications",2017.
5. Mukesh Tiwari, Dr. Rakesh Singhai, "A Review of Detection and Tracking of Object from Image and Video Sequences ", International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 5 (2017), pp. 745-765.2017.
6. Yongyi Lu HKUST, Cewu Lu Shanghai Jiao, "Online Video Object Detection using Association LSTM ",2017 .
7. Bing Tian ,Liang Li ، Yansheng Qu ، Li Yan, "Video Object Detection for Tractability with Deep Learning Method", 2017 IEEE .
8. M.Buric,mpolar,M.Ivassickos,Rijkca", MIPRO 2018, May 21-25, 2018.
9. Fancy Joy, V. Vijaya kumar," A Review on Multiple Object Detection and Tracking in Smart City Video Analytics", (IJITEE) ISSN: 2278-3075, Volume-8 Issue-2S2 December 2018.
10. Rahul Dutt Sharma1, Subham Kumar Gupta, "A Survey on Moving Object Detection and Tracking Based On Background Subtraction", OJIDS Oxford Journal of Intelligent Decision and Data Science 2018 .
11. Elena Luna , Juan Carlos San Miguel , Diego Ortego and José María Martíne , "Abandoned Object Detection in Video-Surveillance: Survey and Comparison", : 5 December 2018.
12. Ajeet Ram Pathaka, Manjusha Pandeya, Siddharth Rautaraya, "Application of Deep Learning for Object Detection "Published by Elsevier, the International Conference on Computational Intelligence and Data Science (ICCIDS 2018). Published by Elsevier Ltd.
13. Journal of Physics: Conference Series PAPER • OPEN ACCESS "Understanding of Object Detection Based on CNN Family and YOLO"Juan Du 2018 J. Phys.: Conf. Ser. 1004 012029.
14. OPEN ACCESS "Understanding of Object Detection Based on CNN Family and YOLO" To cite this article: Juan Du 2018 .
15. Zhong-Qiu Zhao , Member, IEEE, Peng Zheng, Shou-Tao Xu, and Xindong Wu , Fellow, IEEE, "object detection with deep learning: a review", IEEE transactions on neural networks and learning systems 1,ieee transactions,2019.
16. Ting-WuChin, RuizhouDing Diana Marculescu," Ada scale: towards real-time video object detection using adaptive scaling", IEEE-2019.
17. Peng Tang, Chunyu Wang, Xinggang Wang, Wenyu Liu, Wenjun Zeng, and Jingdong Wang," Object Detection in Videos by High Quality Object Linking",2019.
18. JianjiangZhu,1 SiquanYu,2,3 ZhiHan ,2 YandongTang,2andChengdongWu3," Underwater Object Recognition Using Transformable Template Matching Based on Prior Knowledge", Engineering Volume 2019.

AUTHORS PROFILE

Mrs. Kusuma S., currently working as Assistant Professor, Dept. of ISE, RNSIT. She is having Teaching experience of 15 years and pursuing Ph.D in the area of Video Processing.

Dr. M. V. Sudhamani, currently working as Dean-R&D, Professor and HoD, Dept. of ISE, RNSIT. She is having Teaching, Research and Industrial experience of 25 years. She has specialization in Image Processing, Content-based Image Retrieval, Advanced Algorithms and Databases. Guided and guiding candidates for Ph. D degree. She has carried out two research projects from VTU and AICTE. She has served as member of Board of Examiners (BOE) and Board of Study (BOS) member in VTU and other autonomous institutions across India. She has organized two international conferences ICDECS 2011 and 2015, and one more in December 2019.