ISSN (e): 2250 – 3005 || Volume, 09 || Issue, 7 || July– 2019 ||
International Journal of Computational Engineering Research (IJCER)

Research on Object Detection in Video Streaming Using Deep Learning

Shaba Irram, Sheikh Fahad Ahmad
Department of Computer Science and Engineering, Integral University, Lucknow, U.P.
Corresponding Author: Shaba Irram

ABSTRACT: *Machine learning researchers have analyse that to acquire more accuracy in object detection using deep learning techniques, the key is efficiently optimized feature extraction. So, we focus on the feature extraction part of the whole process. Mainly in Deep learning we have CNN for feature extraction. From many research papers it is concluded that best result is obtained when feature extraction is done by CNN. Many authors have been tried to obtain good accuracy in object detection but it becomes very difficult when the objects are moving and they proposed many techniques for object detection and tracking. Different parameters like throughput, accuracy and speed are used for measuring the performance of algorithm. The proposed approach in this paper increases the accuracy and speed of tracking the objects in a video stream by using MCNN for feature extraction. This new approach is based on Depth CNN and RGB CNN both.*

Keywords: *Deep Learning, Object Detection, Video Streaming, CNN, SVM, D-CNN*

Date of Submission: 07-07-2019 Date of Acceptance: 25-07-2019

I. INTRODUCTION

Detecting and tracking moving objects in video files has wide selection of applications in real life. If an individual simply watches a video file, human eyes won't be able to provide specific details about the time and also the method of fixing location of objects, particularly those with a quick change method (such as detection of face, smile detection once taking footage, cars on the road, the trail of a moving ball and others), objects with sophisticated orbits (such as football players running on the pitch), or objects with a method of modification slowly compared to the background (objects between the ocean and sky, daylight changes, and tides). Additionally, the object detection within the video file at the instant can facilitate lots in reality, as an example, to determine if there was a goal or not (in football), if a court game balls is in/out of court (in tennis), or that contestant has finished initial (in speed races). Many methodologies are proposed to perform this task in many research papers, some of them are:

Z. Kalal et al. [1] shows a novel system for long trailing of a person's face in unconstrained videos is constructed on Tracking-Learning-Detection (TLD) approach. The system extends TLD with the idea of a generic detector and a validator that is intended for real-time face trailing resistent to occlusions and look changes. The off-line trained detector localizes frontal faces and the on-line trained validator decides that faces correspond to the half-tracked subject. many methods for building the validator throughout trailing are quantitatively evaluated. The system is valid on a broadcast episode (23 min.) and a monitoring (8 min.) video. In each cases the system detect stracks the face and mechanically learns a multi-view model from one example and an unlabelled video.

J. Chai et al. [2] proposed methodology to recover a transparent image from one motion blurred image which has long been a difficult open downside in digital imaging. A regularization-based approach is projected to get rid of motion blurring from the image by regularizing the sparseness of each the first image and therefore the motion- blur kernel beneath tight rippling frame systems. moreover, an tailored version of the split Bregman methodology is projected to expeditiously solve the ensuing diminution downside. The experiments on each synthesized pictures and real images show that our rule can effectively take away complicated motion blurring from natural images while not requiring any previous info of the motion- blur kernel.

A.Dimou et al. [3] proposed a methodology to improve the overall performance of CCTV footage to detect many objects at a time. The data set with different objects for motion blur are researched to train thedetector. The Blurred content was detected with high performance during PTZ (pan, tilt, zoom) operation. The spatial transformation of the object is used to train the RNN in each frame of the video. It is observed that the functioning of the detector can be improved by real time scaling comparing with static scale operation.

L.H.Jadhav et al. [4] proposes a methodology to detect the removed object from a scene in video. The hybrid model is used to detect the non moving objects. When a non moving object is detected its features are taken out to classify the object. The features are extracted in the form of height, width, size, colour and time. Many parameters are used to improve accuracy like human, luggage, attitudes etc.

K. Simonyan et al. [5] In this work we tend to investigate the impact of the convolutional network depth on its accuracy within the large-scale image recognition setting. Our main contribution is a thorough analysis of networks of accelerating depth using a design with very tiny (3×3) convolution filters, that shows that a major improvement on the prior-art configurations are often achieved by pushing the depth to 16–19 weight layers. We tend to conjointly show that our representations generalise well to different datasets, wherever they win progressive results. We have created our 2 best-performing ConvNet models in public out there to facilitate further analysis on the employment of deep visual representations in laptop vision.

R. Girshick et al. [6] This paper proposes a quick Region-based Convolutional methodology (Fast R-CNN) for object detection. Fast R-CNN builds on previous work to with efficiency classify object proposals victimisation deep convolutional networks. Compared to previous work, quick R-CNN employs many innovations to improve coaching and testing speed whereas conjointly increasing detection accuracy. quick R-CNN trains the terribly deep VGG16 network 9× quicker than R-CNN, is 213× quicker at test-time, and achieves the next mAP on PASCAL VOC 2012. Compared to SPPnet, quick R-CNN trains VGG16 3× faster, tests 10× quicker, and is additional correct.

L. Zhang et al. [7] proposed methodology for detecting pedestrian as a special topic on the far side general object detection. Though recent deep learning object detectors like Fast/Faster R-CNN have shown wonderful performance for general object detection, they need restricted success for investigating pedestrian, and former leading pedestrian detectors were normally hybrid ways combining handmade and deep convolutional options. In this paper, we tend to investigate problems involving quicker R-CNN for pedestrian detection. we tend to discover that the Region Proposal Network (RPN) in Faster R-CNN so performs well as a complete pedestrian detector, but amazingly, the downstream classifier degrades the results. For accuracy two main reasons are: (i) low resolution of feature maps for handling little instances, and (ii) lack of any bootstrapping strategy for mining exhausting negative examples.

S. Ren et al. [8] State-of-the-art object detection networks rely upon region proposal algorithms to expect object locations. Advances like SPPnet [1] and fast R-CNN [2] have reduced the time period of those detection networks, exposing region proposal computation as a bottleneck. during this work, we tend to introduce a region Proposal Network (RPN) that shares full-image convolutional options with the detection network, therefore sanctionative nearly cost-free region proposals. an RPN may be a totally convolutional network that at the same time predicts object bounds and objectness scores at every position. The RPN is trained end-to-end to generate high-quality region proposals, that are utilized by quick R-CNN for detection. we tend to more merge RPN and fast R-CNN into one network by sharing their convolutional features—using the recently widespread language of neural networks with 'attention' mechanisms, the RPN part tells the unified network wherever to seem.

K.Tahboub et al. [9] proposed a two-stage quality-adaptive convolution neural network to address the challenges of changing video data rates. Video compression is defined on a person's foot detection state-of - the-art and is examined for training in compressed images. Video compression can produce poor quality that affects video processing accuracy rates. In this paper, we explore the problem of changing video object rates and study how it involves the workings of video processing, especially pedestrian detection, using a two-stage quality-adaptive neural network.

L. Deng et al. [10] proposed methodology that provides an outline of general deep learning methodology and its applications to a range of signal and data process tasks. the applying areas are chosen with the subsequent 3 criteria in mind: (1) experience or information of the authors; (2) the application areas that have already been reworked by the productive use of deep learning technology, like speech recognition and laptop vision; and (3) the applying areas that have the potential to be impacted considerably by deep learning, both tongue and text process, information retrieval, and multimodal informatics authorized by multi-task deep learning.

B. Hio et al. [11] proposed a background subtraction methodology for moving target detection from the image used. In this paper, we suggest the next moving target location method that applies a deep understanding of how even in a workable framework to achieve a wholehearted act. The methodology recommended follows qualities of visual point of view just as features of movement. To this end, we are reasoning for a deep learning engineering that has accumulated two systems: a visual perspective system and organizing a movement. The two systems are made by applying the presence of the objective item despite the movement contrast to powerfully discover the moving target to the foundation movement. The examination shows that the technique proposed achieves a speed of 50fps in GPU.

Christian S et al. [12] During this paper we tend to go one step any and address the problem of object detection using DNNs, that's not solely classifying however conjointly precisely localizing objects of varied categories. we tend to gift an easy and nevertheless powerful formulation of object detection as a regression drawback to object bounding box masks. we tend to outline a multi-scale reasoning procedure that is ready to supply high-resolution object detections at a low value by some network applications.

S. Hayat et al. [13] Proposed an machine learning methodology to distinguish convolution neural network (CNN) evidence and research as a multi-class target. The convolution neural network system is generated with standardized grade instability and primed with the preparation of test group images from 9 different target classes as well as test images using generally changed data set. All results in the python tensor stream system are completed. We disintegrate and contrast CNN results and last mark vectors detached from different BOW techniques depending on the direct classifier L2-SVM.
C. Li et al. [14] demonstrates deep learning of small samples assisted by a target detector. As a subject of first interest, to evaluate the issue of few examples, the computation can increase greatly the preparation of tests by imported examples generator. The generator of processed examples is expected by exchanging target regions in different scenes. In the deep convolution neural network system, deeply supervised learning and dense form of expectation are improvised along these lines.
S. Rosi et al. [15] proposed to survey on moving object detection and chase strategies is conferred by classifying them into completely different classes and establish new trends. This survey shows moving object detection and chase using totally different and efficient methodologies. Object detection and object tracking is used to trace the article type(such as human, vehicles) and notice the movement of the object(such as moving, standing).This survey shows numerous methodologies for object detection and tracking like background subtraction, background modeling, intensity vary based mostly background subtraction. The simulated result shows that used methodologies for effective object detection has higher accuracy and with less interval consumption instead of existing strategies.
L.Wu et al. [16] proposes program which is based on embedded video surveillance which support moving object detection. The Servfox streaming media is used on ARM Linux based work frame and which make possible to get the videos by using USB cameras. It follows method of background model with inter frame difference to find the object. The images which are the background can be found with more accuracy. To find the moving object background difference technique is used and to find the average object speed inter frame difference technique is used[17, 18, 19].
A.Ucar et al . [20] proposed a auto-directed driving technique involving accurate target identification in real driving environments. In this paper, we offer a Convolutional Neural Network (CNNs) with a new hybrid Local Multiple System (LM-CNNSVM) and Support Vector Machines (SVMs), respectively, with their strong feature extraction technology and robust classifying feature. We divide the whole image into local domains in the proposed system and use multiple CNNs to acquire local target features. Second, by applying Principal Component Analysis, we take features. Then we introduce multiple SVMs that practice real and operational risk reduction instead of using a direct CNN to increase the classifier system's generalization strength.
B. Tian et al. [21] proposed a footage-supported technique for deep learning target detection. The observation video is gathered right from the start, from which a remarked point-target image database with the ultimate goal that was worked as individuals or vehicles to disconnect convolution neural network structure. An ongoing item location and ID framework is anticipated and linked to the prepared structure. For the most part, the proposed system includes three perspectives: video monitoring, object recognition and proof-differentiating target. It provides a type of video interfaces to help the video clips extracted and the video stream continuously. The content-based results show that for the recognizability application, the proposed deep training-based location system is practicable.
H. S. G. Supreeth et al. [22] suggested methodology for Gaussian Mixture Framework (GMM) represented target detection, deep learning neural network-based categorization and target detection using correlation filter is proposed, which can manage phony discoveries and improve productivity. When the work is investigated using "True Positive Rate" (TPR) and "False Alarm Rate" (FAR) as probabilistic measurements, the computation is expected to discover only vehicles and people.
N. Paragios et al. [19] proposed a brand new variational framework for detection and tracking multiple moving objects in image sequences. Motion detection is performed employing a statistical framework that the discovered interframe distinction density operate is approximated employing a mixture model. This model consists of 2 parts, namely, the static (background) and also the mobile (moving objects) one. This statistical framework is employed to provide the motion detection boundaries. Then, the detection and also the chase drawback are addressed during a common framework that employs a geodesic active contour objective operate. This operate is reduced employing a gradient descent methodology, wherever a flow deforms the initial curve towards the minimum of the objective operate, below the influence of internal and external image dependent forces.
N. K. Patil et al. [23] proposes a completely unique technique of illumination standardisation supported bar graph of a picture and scaling perform. It helps in construction of a best international lighting area from these pictures that improve accuracy of face recognition system. The projected technique helps in recognition of sparsely sampled pictures with completely different lighting too. Also, most valuable data of a picture, i.e. grey

value, is not discarded and person's discriminative data in face image is strengthened. thus recognition will be applied using preserved illumination invariant options.
J. Pan et al. [24] proposes a methodology to devise an area best match authentication (LBMA) algorithmic rule to handle complete occlusions, so a way more trustworthy detection of the tip of an indiscriminately long complete occlusion is achieved. In each the real-world and synthetic video sequences show that our projected resolution tracks targets dependably and accurately irrespective of after they are under: short-run, long-term, partial or complete occlusions.
Zaimbashi et al. [25] In a clutter-dominant target detection problem, an adaptive detector must either discolour or null the muddle previous to reception. To discolour muddle, it's needed to estimate the variance matrix of a received muddle, whereas a correct set or estimate of the parameters of muddle topological space model is needed to null the received muddle. Our simulation results show that the subspace-based detector not solely will reach a preset chance of warning (fully nulling clutter) but additionally offers higher detection performance as compared with its counterpart. during this case, the SCM-based detector solely offers better detection performance for detection targets with Doppler frequency resided in muddle region as compared therewith of subspace-based one.
R. Xu et al. [27] In the examination method of the helicopter, the airborne video camera can turn out a small or severe noise inevitably. so as to get a stable, clear review video, video noise has to be paid and estimation of the camera motion is essential to motion compensation. During this paper, we found that there are 3 sure modes of capturing of the airborne video camera, together with dynamic capturing mode along the craft line, static capturing mode in hover, and the switch capturing mode of the 2 modes.
Meanwhile, there are three shaking mode consequently, together with dynamic shaking mode, static shaking mode and switch shaking mode. In this paper, a video motion compensation methodology is projected by analysis of various shaking modes of mobile video camera in the light of continuous video frames. Experimental result showed the projected methodology is admire Kalman filter and mean filter.
L. Snidaro et al. [28] In this paper, an out of doors multi-camera video investigation system operative below ever-changing climatic conditions is conferred. a replacement confidence live, look magnitude relation (AR), is outlined to mechanically judge the sensors' performance for each time instant. By comparison their ARs, the system will choose the foremost acceptable cameras to perform specific tasks. once redundant measurements are accessible for a target, the square measure measures are accustomed perform a weighted fusion of them. Experimental results are conferred on out of doors scenes below completely different climatic conditions.

II. PROPOSED METHODOLOGY

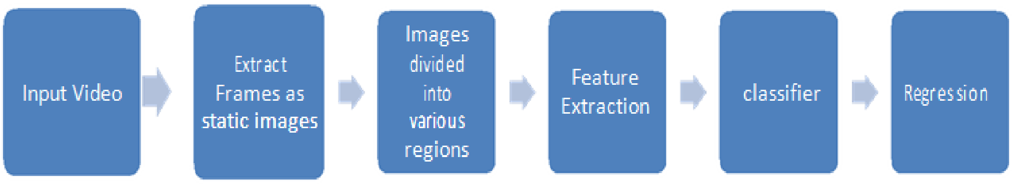


Fig1 : Block Diagram

1. First we take video and extract frames from videos.
2. The Image will be divided into different reasons.
3. We will then consider each reason as a separate image.
4. Pass all these reasons images to the CNN and classify them into various classes.
5. Once we have divided each reason into its corresponding class we can combine all the Seasons to get the original image with the detect object.
6. Extracting 2000 regions for each image based on selected search.
7. Features are extracted using CNN for each reason of image, suppose we have K images then the number of CNN features will be K*2,000.
8. The entire process of object detection using a (modified)MCNN has three models:

1) CNNs for feature extraction: Depth CNN and RGB CNN
2) SVM(support vector classifier) for classification
3) Logistic Regression for prediction

III. ALGORITHM

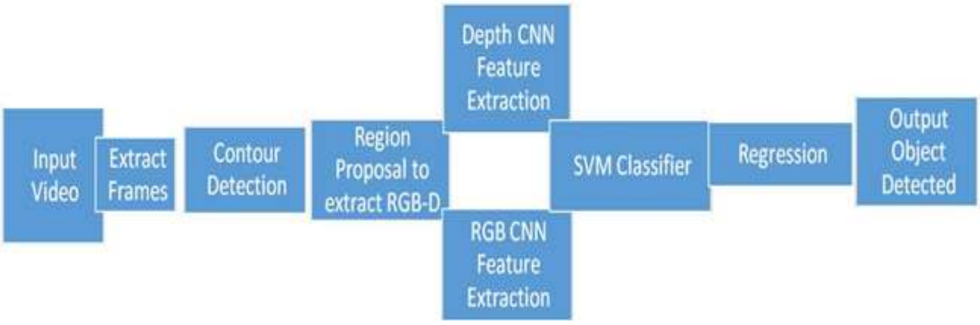


Fig2: MCNN Approach for Object Detection

1. Contour Detection using HED

It is a contour detection technique which uses holistically nested edge detection. It uses the side outputs of intermediate layers. It performs image to image prediction based on Deep learning. It automatically represents radical structure in order to achieve human ability to solve the ambiguity in edge and object boundary detection [29].

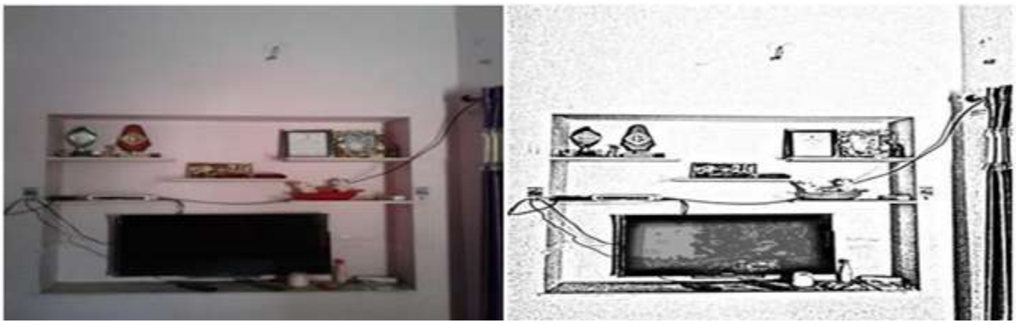


Fig3 : Contour Detection using HED

2. Region Proposal using Boundary Segmentation

It uses segmentation to identify the target in an image and form a group of adjacent regions which are similar in colour, shape, texture etc. It is a hierarchical model that divide into regions on the basis of pixel, regions, object in the image. The pixels are labelled at the region level depends on background classes or foreground class[30].

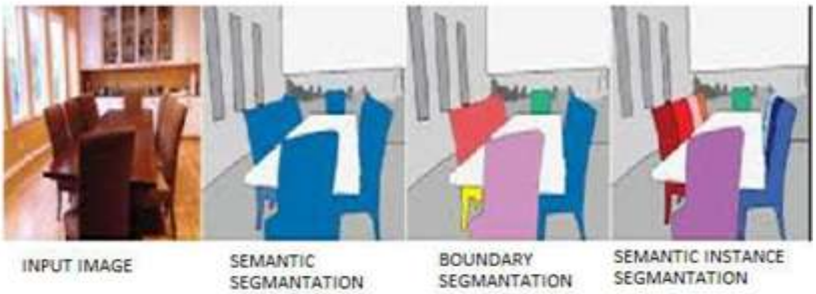


Fig4 : Region Proposal using Boundary Segmentation

3. Depth CNN Feature Extraction

It works on patches rather than whole image. Weakly-supervised pre-training CNN is used to extract multiple patches from a single image. We include weakly-supervised training because patches only cover few part not whole scene. After training from weakly-supervised pre training CNN we transfer the weight to CNN for fine tuning[31].

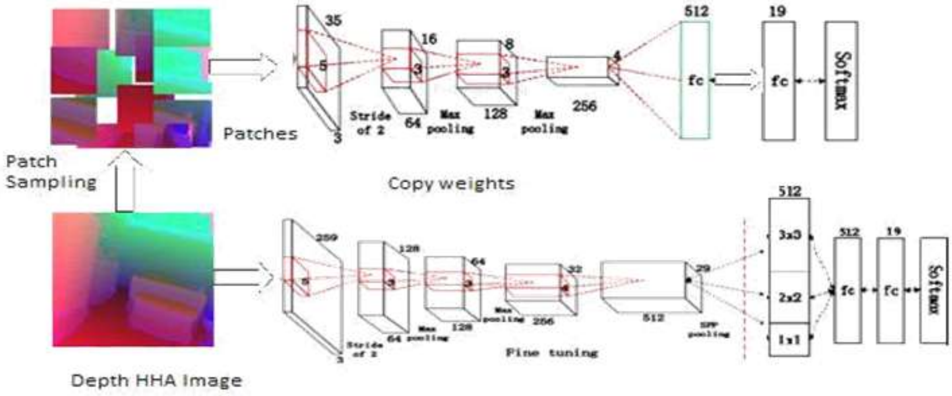


Fig5 : Depth CNN Feature Extraction

4. RGB CNN Feature Extraction

It is the basic colour pattern for any image which is used by CNN model. Before depth images features are extracted using CNN model pre trained on RGB images. Depth images pre trained by RGB CNN model encoded by HHA method have proved in appreciated result in object detection[32, 33, 34].

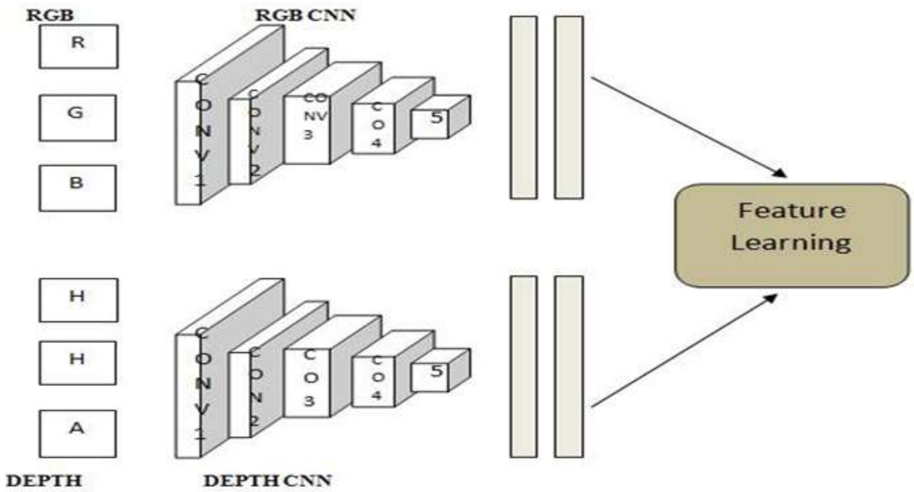


Fig6 : RGB CNN Feature Extraction

5. SVM Classifier

It is a support vector machine classifier it is used to find a hyperplane in an n dimensional space that classify the data points. It is a supervised learning model use for object classification. It uses decision boundaries for dividing the objects and decide whether the object belongs to which class. It separate the object by dividing into categories by a clear gap[35].

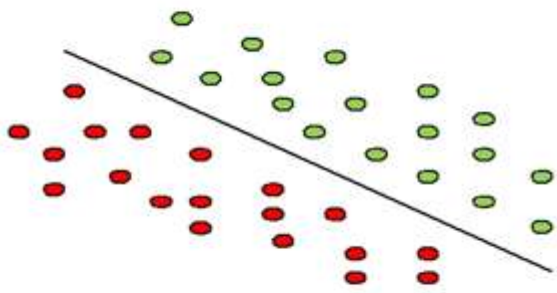


Fig7 : SVM Classifier

6. Logistic Regression for Prediction

It is a type of regression analysis used for prediction in object detection using deep learning. It is used to predict the outcomes of the dependent variable which is based on one or more independent variables. Response model is used to estimate the expectation values. It is used for the problem in which the dependent variable is in binary. Regression analysis is used to check every feature if it is meaningful or not[36].

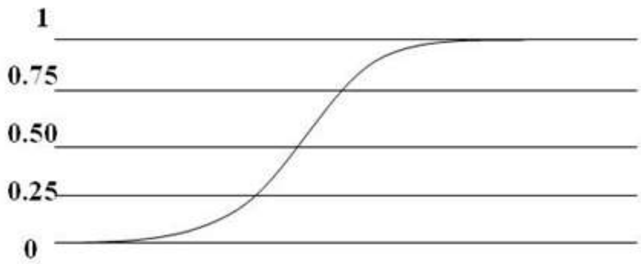
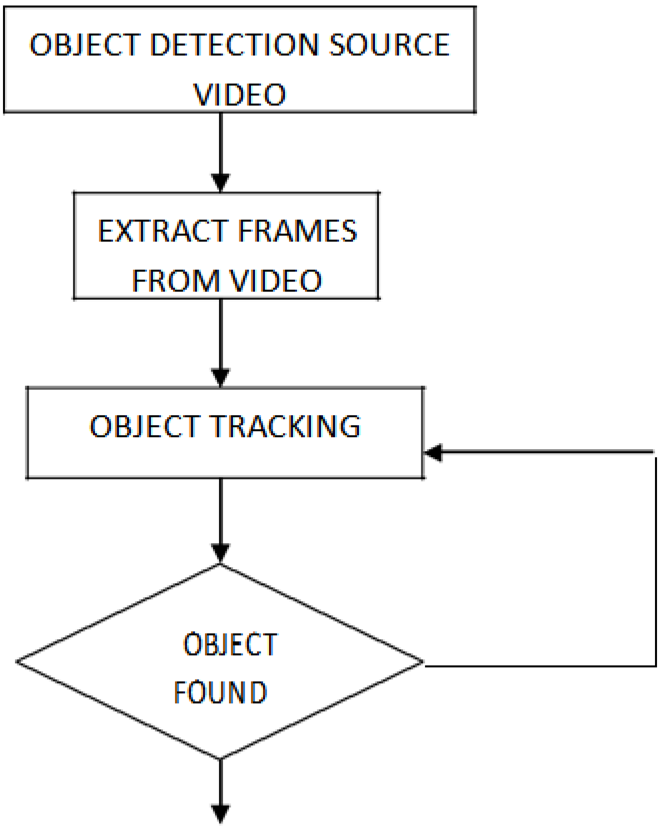


Fig8 : Logistic Regression for Prediction

T**able1** : Overview of Various Techniques used in Algorithm

Various Techniques used in Algorithm	**Researchers**
Contour Detection using HED	S. Xie et al. [29]
Region Proposal using Boundary Segmentation	Gould et al. [30]
Depth CNN Feature Extraction	Song et al. [31]
RGB CNN Feature Extraction	W. Zhang et al. [32]
SVM Classifier	Tao Q-Q et al [35]
Logistic Regression for Prediction	J. Kim et al. [36]

IV. FLOW CHART



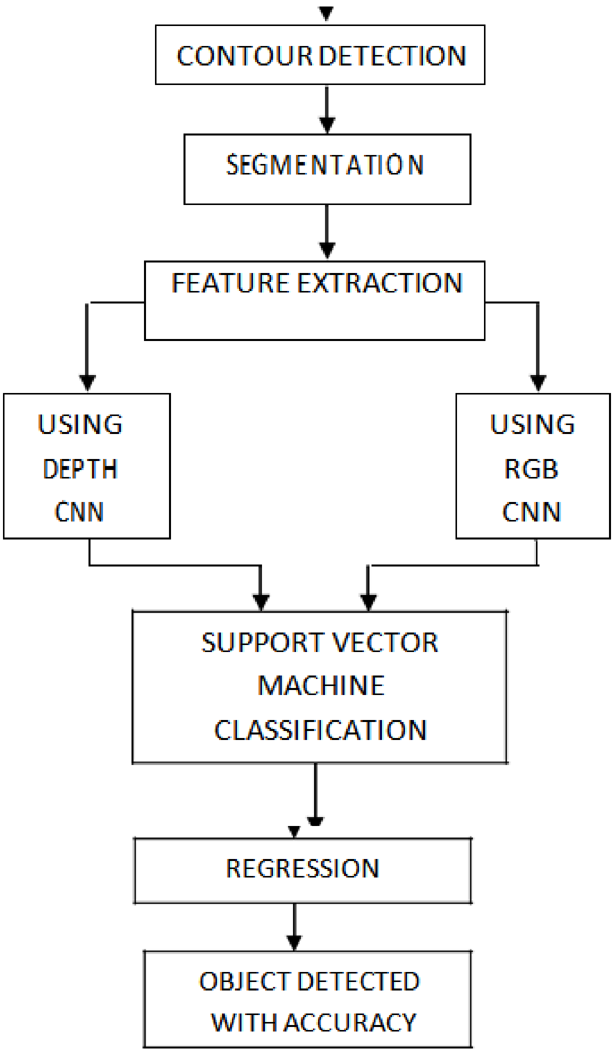


Fig9 : Flow Chart

V. RESULT ANALYSIS

Many machine learning practitioners believe that properly optimized feature extraction is the key to effective model with more accuracy. So, we focus on the feature extraction part of the whole process. Mainly in Deep learning we have CNN for feature extraction. From many research papers it is concluded that best result is obtained when feature extraction is done by CNN. Based on the RGB image and the Depth Map, detect the outline in the image and generate region proposals. Using CNN for feature extraction, the network here includes two CNNs:

1) Depth CNN learning features on the depth map
2) RGB CNN learning features on the 2D image
3) Finally using SVM for classification.

Depth + RGB = D(Depth) Map(HHA) + RGB Color Image

Result for object detection in video streaming is implemented in python code using Tensorflow. The result of this approach compares with various model of CNN result from previous literature review which calculated by using MatplotLib in python environment. Results for object detection and tracking is measure at different objects.

Comparative Study Analysis between MCNN and various previous approaches:

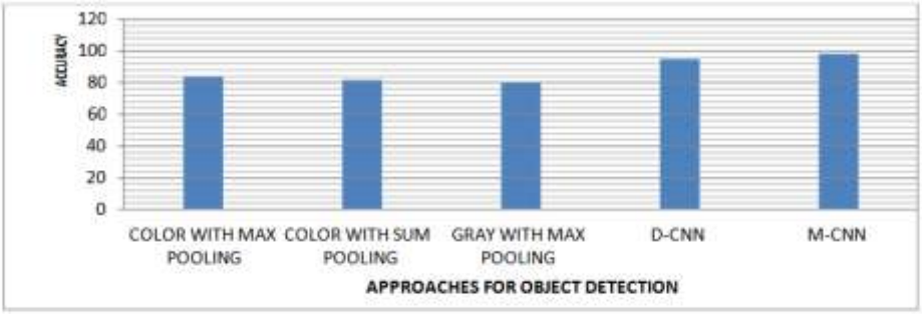


Fig10 : Comparative Study Analysis

In our proposed native approach, for measuring the performance of the object detector, we have used following values for different parameters:

Number of Objects: 80 objects
Models: CNN, SVM and Regression model
Here we can see that the result obtain by the new approach of CNN i.e. MCNN is showing better result than any other models of object detection. MCNN has achieved more than 98% of accuracy in detecting objects from videos.

VI. CONCLUSION

A Research process consists of series of procedures or steps necessary to successfully carry out research and the preferred sequencing of these steps to produce new knowledge, or to offer a new manner of accepting present knowledge.
We have practical a quantitative approach towards the study. In distinct phases we have designed this work, each stage having its own importance. After identification of the problem, we have done a literature survey in a detailed, focusing on the accuracy in detecting objects from given video streaming.
This literature review was followed by a implementation modeling. The results were gathered, analyzed and conclusions were drawn on the basis of the results obtained from the implementation.

REFERENCES

[1]. Zdenek Kalal, Krystian Mikolajczyk, and Jiri Matas," Tracking-Learning-Detection", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 6, NO. 1, JANUARY 2010
[2]. Jian-Feng Cai, Hui Ji, Chaoqiang Liu, and Zuowei Shen, "Framelet-based blind motion deblurring from a single image," Image Processing, IEEE Transactions on, vol. 21,no.2,pp.562–572,2012
[3]. A. Dimou, P. Medentzidou, F. Á. García and P. Daras, "Multi-target detection in CCTV footage for tracking applications using deep learning techniques," 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 928-932.
[4]. L. H. Jadhav and B. F. Momin, "Detection and identification of unattended/removed objects in video surveillance," 2016 IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), Bangalore, 2016, pp. 1770-1773
[5]. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Proceedings of the International Conference on Learning Representations (arXiv:1409.1556), May 2015, San Diego, CA
[6]. R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448, December 2015, Santiago, Chile.
[7]. L. Zhang, L. Lin, X. Liang and K. He, "Is faster R-CNN doing well for pedestrian detection?" Proceedings of the IEEE European Conference on Computer Vision, pp. 443–457, October 2016, Amsterdam, Netherlands.
[8]. S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," Proceedings of the Advances in Neural Information Processing Systems Conference, pp. 91–99, December 2015, Montr´eal, Canada.
[9]. K. Tahboub, D. Güera, A. R. Reibman and E. J. Delp, "Quality-adaptive deep learning for pedestrian detection," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 4187-4191.
[10]. Deng L and Yu D. Deep learning: methods and applications. Foundat Trend Signal Proc 2014; 7: 3–4.
[11]. B. Heo, K. Yun and J. Y. Choi, "Appearance and motion based deep learning architecture for moving object detection in moving camera," 2017 IEEE International Conference on Image Processing (ICIP), Beijing, 2017, pp. 1827-1831.
[12]. Christian S, Toshev A and Erhan D. Deep neural networks for object detection. In: Proceedings of advances in neural information processing systems (NIPS), Lake Tahoe, Nevada, 5–10 December 2013, pp.2553–2561. Red Hook, NY: Curran Associates Inc.
[13]. S. Hayat, S. Kun, Z. Tengtao, Y. Yu, T. Tu and Y. Du, "A Deep Learning Framework Using Convolutional Neural Network for Multi-Class Object Recognition," 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC), Chongqing, 2018, pp. 194-198.
[14]. C. Li, Y. Zhang and Y. Qu, "Object detection based on deep learning of small samples," 2018 Tenth International Conference on Advanced Computational Intelligence (ICACI), Xiamen, 2018, pp. 449-454.
[15]. 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 ISSN 2250-3153

[16]. L. Wu, Z. Liu and Y. Li, "Moving objects detection based on embedded video surveillance," 2012 International Conference on Systems and Informatics (ICSAI2012), Yantai, 2012, pp. 2042-2045.
[17]. ZH. J. Zhao, X.ZH. Lin, J.Y. Zhang, "The moving object detection algorithm Based on the background reconstruction," Journal of Beijing Institute of Petro-chemical Technology. Beijing, vol. 16,pp 27-29, February 2008.
[18]. Z. D. Ma, ZH. B. Zhang, "Software Realization of Video Surveillance Terminal Based on ARM- Linux," Computer Measurement & Control .Beijing ,vol.19,pp.456458, February 2011.
[19]. N. Paragios , R. Deriche," Geodesic Active Contours and Level Sets for the Detection and Tracking of Moving Objects," IEEE T PATTERN ANAL.USA, vol. 22,pp. 266-280, March 2000.
[20]. A. Uçar, Y. Demir and C. Güzeliş, "Moving towards in object recognition with deep learning for autonomous driving applications," 2016 International Symposium on INnovations in Intelligent SysTems and Applications (INISTA), Sinaia, 2016, pp. 1-5.
[21]. B. Tian, L. Li, Y. Qu and L. Yan, "Video Object Detection for Tractability with Deep Learning Method," 2017 Fifth International Conference on Advanced Cloud and Big Data (CBD), Shanghai, 2017, pp. 397-401.
[22]. H. S. G. Supreeth and C. M. Patil, "Moving object detection and tracking using deep learning neural network and correlation filter," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, 2018, pp. 1775-1780.
[23]. N. K. Patil, S. Vasudha and L. R. Boregowda, "A Novel Method for Illumination Normalization for Performance Improvement of Face Recognition System," 2013 International Symposium on Electronic System Design, Singapore, 2013, pp. 148-152
[24]. J. Pan, B. Hu and J. Q. Zhang, "Robust and Accurate Object Tracking Under Various Types of Occlusions," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 2, pp. 223-236, Feb. 2008
[25]. A. Zaimbashi, "Target detection in clutter background: Null or whiten the clutter," 2016 24th Iranian Conference on Electrical Engineering (ICEE), Shiraz, 2016, pp. 1963-1966.
[26]. R. Guo, Q. Dai and D. Hoiem, "Paired Regions for Shadow Detection and Removal," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 12, pp. 2956-2967, Dec. 2013.
[27]. R. Xu, Z. Zhang, F. Qin, Z. Zhou, H. Yang and R. Li, "Research on video motion compensation method based on the analysis of airborne video camera shaking mode of continuous video frames," 2016 IEEE PES 13th International Conference on Transmission & Distribution Construction, Operation & Live-Line Maintenance (ESMO), Columbus, OH, 2016, pp. 1-4.
[28]. L. Snidaro, R. Niu, P. K. Varshney and G. L. Foresti, "Automatic camera selection and fusion for outdoor surveillance under changing weather conditions," Proceedings of the IEEE Conference on Advanced Video and Signal Based Surveillance, 2003., Miami, FL, USA, 2003, pp. 364-369.
[29]. S. Xie and Z. Tu, "Holistically-Nested Edge Detection," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, 2015, pp. 1395-1403.
[30]. Gould, Stephen & Gao, Tianshi & Koller, Daphne. (2009). Region-based Segmentation and Object Detection.. 655-663.
[31]. Song, Xinhang & Herranz, Luis & Jiang, Shuqiang. (2018). Depth CNNs for RGB-D scene recognition: learning from scratch better than transferring from RGB-CNNs.
[32]. W. Zhang, G. Liu and G. Tian, "HHA-based CNN image features for indoor loop closure detection," 2017 Chinese Automation Congress (CAC), Jinan, 2017, pp. 4634-4639.
[33]. S. Gupta, R. Girshick, P. Arbelaez, and J. Malik, ´ Learning Rich Features from RGB-D Images for Object Detection and Segmentation. Cham: Springer International Publishing, 2014, pp. 345–360.
[34]. S. Gupta, P. Arbelez, and J. Malik, "Perceptual organization and recognition of indoor scenes from rgb-d.
[35]. Tao Q-Q, Zhan S, Li X-H, et al. Robust face detection using local CNN and SVM based on kernel combination. Neuro computing 2016; 211: 98–105
[36]. J. Kim et al., "Optimal Feature Selection for Pedestrian Detection Based on Logistic Regression Analysis," 2013 IEEE International Conference on Systems, Man, and Cybernetics, Manchester, 2013, pp. 239-242.

Shaba Irram" Research on Object Detection in Video Streaming Using Deep Learning" International Journal of Computational Engineering Research (IJCER), vol. 09, no.7, 2019, pp 34-43