

## **Novel Method for Object Detection in real time using AI and Deep Learning**

**Digambar Jadhav<sup>1</sup>, Dr. Soumitra Das<sup>2</sup>, Vaibhav Muddebihalkar<sup>3</sup>, Laxman Khandare<sup>4</sup>**

Dept. of Computer Engineering,

<sup>1,3,4</sup>Dr. D. Y. Institute of Technology, Pimpri, Maharashtra, India

<sup>2</sup>Indira College of Engineering and Management, Pune

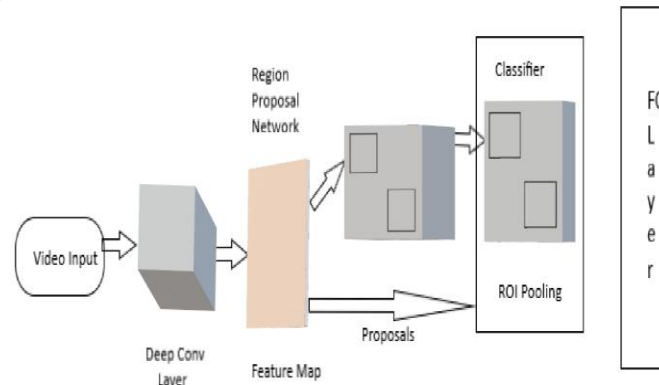
**Abstract:** *In object detection, it is a very challenging task to track movements of objects and generate high-efficiency results. We can generate better applications with the help of computer vision technology. Our main focus area of object detection is to enhance the E-commerce business and to provide a better experience to customers by recommending products. We are trying to make a model that will be beneficial for the user as well as the provider. It is complex to detect shopping items in a video. We are using Faster CNN to generate accurate results so that we can detect small objects in video to provide recommendation links.*

**Key Words:** Object Tracking, Computer vision, Faster R-CNN, Image Retrieval, Recommendation, E-commerce etc.

### **1. INTRODUCTION**

In today's world, most of people are using social media and watch different videos. If anyone likes any object or product in a video we can directly provide a link of the similar products which are available on internet through recommendation links. Faster RCNN algorithm provides better efficiency and performance than that of RCNN algorithm. It uses convolution neural networks like YOLO and SSD. As the processing data is huge so we can use a cloud-based system [7]. In the case of video data, the background is constantly changing and it is a challenging task to detect the object [6]. To deal with this we used novel based approach [6, 8]. To solve the problem of silent object detection in video, we used the concept of virtual border and guided filter and embedding topological features into a deep neural network for extracting semantics [3, 4]. Sometimes the data is very complex or crowded and for detecting those complex objects we can use video interlacing to improve multi-object tracking [9]. In many cases, a high amount of time may be required for detecting objects. To deal with this issue we use YOLO (You Look Only Once) algorithm [1].

## 2. FASTER R-CNN



*Fig -1: Faster R-CNN Block Diagram*

Ross Girshick, Shaoqing Ren, Kaiming presented Faster RCNN is an object detection architecture in 2015, and is one of the famous object detection architectures that uses convolution neural networks like YOLO and SSD.

### Step 1: Anchors

By using a system of anchors Faster R-CNN allow the operator to define the possible regions that will be fed into the Region Prediction Network. An anchor is a box. We give stride of 16 so that each of the anchors will slide over the image skipping 16 pixels at a time, there will be almost 18,000 possible regions.

### Step 2: Region Proposal Network (RPN)

The possible regions generated by the algorithm are feed by the algorithm in the previous step, into the RPN, a special CNN used for predicting regions with objects of interest. The RPN predicts the possibility of an anchor being background or foreground and refines the anchor or bounding box.

The training data of the RPN is the anchors and a set of ground-truth boxes. Anchors that have a higher overlap with ground-truth boxes should be labeled as foreground, while others should be labeled as background. The RPN convolves the image into features and considers each feature using the 9 anchors, with two possible labels for each. To predict the labels for each anchor, the output is fed into a Softmax or logistic regression activation function.

### Step 3: Region of Interest (RoI) pooling

Regions of different sizes are provided by the RPN. Each of these is a CNN feature map with a different size. Now the algorithm applies Region of Interest (RoI) pooling to reduce all the feature maps to the same size.

Faster R-CNN performs RoI pooling using the original R-CNN architecture. It takes the feature map for each region proposal, flattens it, and passes it through two fully-connected layers with ReLU activation. It then uses two different fully-connected layers to generate a prediction for each of the objects.

### **3. LITERATURE SURVEY AND RELATED WORK**

Based on the YOLO network author propose a real-time object detection algorithm for videos. Here, the author trains the fast YOLO model by eliminating the influence of the image background by image pre-processing [1]. It is challenging to detect salient objects.

In [2] author paper describes a high-speed video salient object detection method at 0.5s each frame. In [2, 8] author makes use of two models, the initial spatiotemporal saliency module and filter based salient temporal propagation module.

We can also use [3] embedding topological features into a deep neural network for extracting semantics which the author used for salient object detection. Segmentation of input image and compute the weight for each region with low-level features. Here, the weighted segmentation result is called a topological map and it provides an additional channel for CNN. By making use of virtual border and [4] guided filter author trying to propose a novel method for salient object detection in videos.

The author used a novel multi-task framework [5] for object detection. A novel multi-task framework uses multi-label classification as an auxiliary task which will improve object detection and can be trained and tested end-to-end. In some cases, there may be moving cameras that result in a variable background.

In [6] the author is using a novel approach for detecting and tracking objects in videos that are captured by cameras. It is a challenging task to separate actual moving objects from the background as both background and foreground changes in each frame of the image sequence. In object detection, we need to handle data that may be huge or small.

In [7] the author makes use of an automated video analysis system to process a large number of video streams. On can get access to a huge amount of data using a cloud-based system. Cloud provides us unlimited storage which results in saving hardware cost. Some video data frames consist of complex and multiple objects which are sometimes difficult to track.

In [9] the author is using Multi-Object Tracking-by Detection which is based on a Spatio-temporal interlaced encoding video model and specialized DCNN.

### **4. EXISTING SYSTEMS AND ITS LIMITATIONS**

1. In the case of existing systems, we make use of different methods and algorithms like YOLO, MOT-bD, topological maps, novel framework, and guided filter.
2. By making use of these techniques and methods we can solve many problems but still, there is a scope to improve to the existing system in some areas.
3. We can improve our system in multi-class detection. Sometimes there are many classes present in the image of video and it is challenging to detect multiple classes.

4. The efficiency of the methods we are using in a real-time system is less in the present situation which can be improved by trying some different approaches.
5. Existing object detection systems also contain some limitations.
6. In a mean shift when the background is similar to target then tracking problems may arise.
7. In contour tracking, it is difficult to handle the entry and exit of objects.

## 5. NEED FOR OBJECT DETECTION

Object detection techniques are useful to make processing faster and simpler. We need object detection in a variety of areas including surveillance, medical image analysis, human-computer interaction, and robotics. There is a big demand for object detection where we use 3D space.

It is necessary to detect or track all movements in that 3D space. We can depend on object detection algorithms to not only detect objects in an image but to do that with the speed and accuracy. We need object detection to improve workplaces, increase security, and automated vehicle systems. Object detection can be useful in case of traffic management. Government can use object detection techniques to keep track of abnormal activities in various regions to maintain security.

## 6. PROJECT DIAGRAM NOVEL ARCHITECTURE

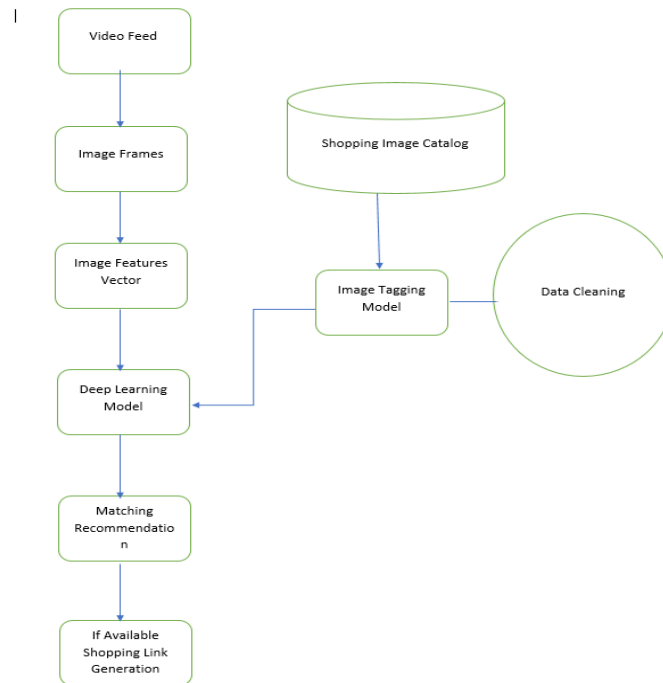


Fig -2: Novel Architecture

## 7. ADVANTAGES

Object detection is advantageous in computer vision and image processing. It is useful in detecting various objects and tracking movements.

1. In object detection, a Kalman filter is a method by which we can track points in noisy images.
2. Particle filter gives us optimal results when the evaluation of the image takes place at the hypothesis object position.
3. In object detection mean shift is applicable for situations with dominant colors.
4. Using CamShift we can apply resizable search windows.
5. We can use the KLT tracker for time-efficient and robust occlusions.
6. By using object detection techniques we can handle complex models for rigid and non-rigid objects.

## 8. MODEL EVALUATION

By using Faster RCNN, we have generated some results which are detecting shoppable objects in the video by labeling them and showing the accuracy.

## 9. CONCLUSIONS

In this paper, we applied the Faster R-CNN algorithm which is useful in tracking objects and their movements. We used different types of video data which includes more shoppable items to generate better results. We implemented the idea of using object detection techniques by giving video input which can be used for recommendation purposes.

In the future, we can improve the performance and efficiency of this model using the implementation of various algorithms. Also, we can try to integrate this model with an e-commerce website.

## REFERENCES

- [1] Shengyu Lu, Beizhan Wang, Hongji Wang et.al. *A real-time object detection algorithm for video*, *Computer and Electrical Engineering* 77 (2019) 398-408.
- [2] Qi Qi, Sanyuan Zhao, Wenjun Zhao, Zhengchao Lei et.al. *High-speed video salient object detection with temporal propagation using correlation filter*, *Neurocomputing* 356 (2019) 107-118.
- [3] Lecheng Zhou, Xiaodong Gu et.al. *Embedding topological features into convolutional neural network salient object detection*, *Neural Networks* 121 (2020) 308-318.
- [4] Qiong Wang, Lu Zhang, Wenbin Zou, Kidiyo Kpalma et.al. *Video salient object detection using a virtual border and guided filter*, *Pattern Recognition* 97 (2020) 106998.
- [5] Tao Gong, Bin Liu, Qi Chu, Nenghai Yu et.al. *Using multi-label classification to improve object detection*, *Neurocomputing* xxx (2019) xxx.
- [6] Kumar S. Ray, Soma Chakraborty et.al, *Object Detection by Spatio-Temporal Analysis and Tracking of the Detected Objects in a Video with Variable Background*, *J. Vis. Commun. Image R.* S1047-3203 (2018) 30328-6.
- [7] Muhammad Usman Yaseen, Ashiq Anjum, Omer Rana, Richard Hill et.al, *Cloud based scalable object detection and classification in video streams*, *Future Generation Computer Systems* S0167-739X (2017) 30192-9.
- [8] Zhigang Tu, Zuwei Guo, Wei Xie, Mengjia Yan et.al, *Fusing disparate object signatures for salient object detection in video*, *Pattern Recognition* S0031-3203(2017)30299-6.
- [9] Ala Mhalla, Thierry Chateau, Najoua Essoukri, Ben Amara et.al, *Spatio-temporal object detection by deep learning: Video-interlacing to improve multi-object tracking*, *Image and Vision Computing* 88 (2019) 120-131.

- [10] *Sohom Mukherjee, Sk. Arif Ahmed, Debi Prosad Dogra, Samarjit Kar et.al, Fingertip detection and tracking for recognition of air-writing in videos, Expert Systems with Applications 136 (2019) 217-229.*
- [11] *Digambar Jadhav, Vaibhav Muddebhalkar, Ashok Korke et.al, Dust Cleaner System for PV Panel using IoT, International Journal of Computer Applications (0975 – 8887) Volume 178 – No. 10, May 2019.*
- [12] *Digambar Jadhav, S.V. Chobe, M. Vaibhav, Laxman Khandare et.al, Missing Person Detection System in IoT, 2017 International Conference of Computing, Communication, Control and Automation (ICCUBEA) 18080236.*
- [13] *Digambar Jadhav, Swati Nikam et.al, Need for Resource Management in IoT, International Journal of Computer Applications (0975-8887) Volume 134-No.16, January 2016.*
- [14] *Digambar Jadhav, Swati Nikam et.al, Resource Management new System Architecture, 2016 International Conference on Computing, Communication, Control and Automation(ICCUBEA) 16585562.*
- [15] *Piyush Sanjay Zope, Suhas Khatal, Nishikant Bahalkar, Rushikesh Gontlewar, Digambar Jadhav, Object Detection in Real Time using AI and Deep Learning, IRJET Volume: 06 Issue: 12 | Dec 2019.*