`

AN OBJECT DETECTION METHOD USING THE MODIFIED YOLO V4 ALGORITHM - AN ABNORMAL ACTIVITY DETECTION IN ROADWAYS AND RAILWAYS

Ms.V.Megala[1], Dr.S.K.Jayanthi[2]

Abstract –
One of the most extensively used techniques in computer vision is object detection. It enables us to recognize and find an object in an image or video. The main aim is to detect the object which was involved in the abnormal activity in public places like Railway Tracks and National Highway Roads. The abnormal activity denotes the harassment, robbery, suicide and accidents which mostly happen during the night time in Roadways and Railways. This detection methodology uses the modified YOLO v4 algorithm to recognize the object, with the kind of identification and localization, it may be used to count the items in a given image and monitor their suitable place. Object detection distinguishes the sorts of objects present in the picture before locating the precise manifestation of the object. For object labelling, a bounding box surrounding the item can be used to accomplish object localization. As a result, the object detection alerts the control room of the appropriate department about the abnormal activity and helps the officers to take immediate action to prevent the illegal activity occurring in the area. This object detection methodology not only detects human beings but also detects the trapped animals in the railway tracks.

Keywords – Object Detection, YOLO v4, Abnormal Activity Detection, Object Localization

[1]Research Scholar, Department of Computer Science, Vellalar College for Women, Tamil Nadu, India

[2]Associate Professor & Head, Department of Computer Science, Vellalar College for Women, Tamil Nadu, India

DOI: 10.48047/ecb/2023.12.si5a.0560

`

I. INTRODUCTION

According to the evidences, the crime rate against women has increased by roughly 131 percent in May 2020. To avoid such incidents abnormal activities has to be detected in advance during the whole day, in the roadways and railways with Computer Vision methodologies. One of the most extensively used techniques in computer vision is object detection. It enables us to recognize and find an object in an image or video. The You Only Look Once techniques are used for the efficient object detection with minimum stipulated time. In order to identify the object that was involved in anomalous behavior in public areas, Computer Vision uses the YOLO technique to find the object within a short span of time. This detection helps the society to prevent the anti-social activities happening during day to day life.

The objective of this detection method is used to detect the abnormal activities occurring in the National Highways and Railways especially during the night time. Since the original You Only Look Once (YOLO v1) was released in 2015, it has received much attention from the computer vision field and got popular with its effective and efficient techniques.

The subsequent version of this algorithm released one after another with enhanced features compared with the previous version. The YOLO techniques are also used to detect the object that was involved in the abnormal activity in public places like Railway Tracks and National Highway Roads.

By monitoring the railway track the trapped animal can be rescued as well as the train accidents can also be avoided. By using this technology, the suicide attempts occurring in the railway tracks can also be avoided with systems that alert the officers about the detection of any person with unusual activity in the railway track.

Accidents in National Roadways are at times unavoidable due to the negligence of drivers due to which the victims face helpless situations; this technique will be a life saver for those people who face helpless situation in such cases. Such people can be detected using this methodology and proper action would be taken immediately to help those people who face emergency situations.

Accidently, some people get trapped in the railway track and struggle to be released. In such situations the train can be stopped due to prior monitoring of the activity this abundantly leads in saving their life. In this way, this concept is a socio-economic related work which will be useful to the society.

The introduction to object detection, as well as the several techniques and methods applied in the object detection system, is described in Section I. Section II deals with the Literature Survey of various articles. Latest Methodologies used to detect the object in the image or video are covered in Section III. In Section IV, the results and discussion were explained. Section V wraps up conclusion with the object detection approaches and looks ahead to potential advancements in the area.

II. LITERATURE SURVEY

During the last few years, Artificial Intelligence (AI) systems and Machine Learning (ML) have emerged as important tools in a variety of fields, including industry, robotics, healthcare, transportation, business and technology. Deep learning and computer vision methods yield simple tools for achieving commercial sector expansions such as object detection, vehicle tracking, and strengthening safety and security[14], [15].

Najafi Kajabad, Ebrahim, Begen Petr, Nizomutdinov Boris and Ivanov Sergey [17] have presented a method based on YOLO v4 technique and it detects the windows, doors, billboards, ramps and other urban objects. Data collection from the environment is the initial phase, and data augmentation techniques are used to produce data, Non Maximum Suppression (NMS) is used for the selection of an object with the bounding boxes.
Zhi-Hao Chen and Jyh-Ching Juang [18] proposed an idea of Nondestructive Radiographic Testing in Aviation Maintenance Tasks using the YOLO v4 Object Detection Model to perform quality-assured aviation maintenance safety inspections, production, and overhauls to reduce the danger of civic aircraft damage incidents and Non-Destructive Testing (NDT) are used with the Caffe model layers for the detection of an object.

Wu, Han, Chenjie Du, Zhongping Ji, Mingyu Gao and Zhiwei He [19] have proposed the Multi-object tracking (MOT) technique using the YOLO v4 algorithm. Focal Loss and RetinaNet were used to increase the accuracy of the image for the detection. By generating deep features for each bounding box and utilizing their similarity to enter into the tracking logic, Simple Online and

`

Realtime Tracking (SORT) add these deep features. Liu, Tao & Pang, Bo & Zhang, Lei & Yang, Wei & Sun, Xiaoqiang have proposed the object detection system for the sea surface, the Unmanned Surface Vehicles (USVs) have long been engaged in hazardous marine operations.
USVs' environment awareness abilities can be improved by using vision-based sea surface object detection techniques. They used the convolutional block and the maxpool layers for the feature map generation and it enhances the detection ability of the system. The backbone network and feature fusion network of YOLO v4 were designed and implemented using Reverse Depthwise Separable Convolution (RDSC).

Liu, Qi & Fan, Xiaoyu & Xi, Zhipeng & Yin, Zhijian & Yang, Zhen [21] have proposed an idea to improve the object identification precision, an updated BiFPN framework based on YOLO v4-Tiny is provided. Furthermore, the YOLO v4-Tiny network is used as the backbone, and Spatial Pyramid Pooling (SPP) is used to connect and integrate multi-scale zones. Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan [22] suggested a technique that includes the Weighted-Residual-Connections (WRC), Cross-Stage-Partial-Connections (CSP), Cross mini-Batch Normalization (CmBN), Self-Adversarial-Training (SAT), and Mish-activation techniques for the efficient result in the object detection methodology. The DropBlock regularization method mainly used to increase the accuracy of the detection method.

All of the work that has been reviewed, that detects the object, but with compromises in features such as accuracy. Using the effectiveness of the OpenCV and Darknet methodologies, hence this modified YOLO v4 algorithm intends to detect the object in the clumsy image.

III. METHODOLOGY

OpenCV (Open Source Computer Vision Library) is a programming library focused more on real-time computer vision. It was created by Intel and later sponsored by Willow Garage and Itseez. OpenCV is a large open-source library for computer vision, machine learning, and image processing, and it presently plays a vital part in real-time operations, which are essential in today's systems. It will be used to detect the people, items and even human handwriting in photos and movies. Python can process the OpenCV array structure for analysis when it is combined with other modules such as NumPy. Deploy the vector space and execute mathematical operations on these characteristics to identify visual patterns and their different features.

Darknet is a C and CUDA-based open source neural network framework. It's quick and simple to set up, and it works with both CPU and GPU processors. It is a robust and precise framework for real-time object detection. Darknet can be used to create advanced deep neural network implementations. You Only Look Once (YOLO) is used for real-time object detection, recurrent neural networks (RNNs), ImageNet classification and many other implementations are among them.

Object detection can be done using the YOLO technique. It's the algorithm or a strategy that determines how the code will recognize the objects in an image or video. Earlier detection frameworks used image classification techniques to detect things by looking at various regions of the image several times at different sizes. This method is inefficient and stagnant. YOLO, on the other hand, takes a completely different strategy. It merely looks at the full image once and scans the network once before detecting objects.

The modified YOLO v4 can be used to detect the custom object by training the dataset with help of Darknet. Modified YOLO v4 focuses real-time detection and implements training on a single GPU. Here the object can be detected and labeled using the bounding boxes, the object which were trapped in the railway track can be identified by using the modified YOLO v4 technique and the bounding box were created around the object for the labeling process, the object might be an living or non-living thing.

By identifying the object in the railways or roadways the control section of the appropriate department can be made alert and it leads to the avoidance of major accident that occurs in both day time and night time.

The object detection uses various methodologies to find the object. The following flow chart depicts the process of an image processing by using the modified YOLO v4 algorithm. Images are preprocessed before it entered into the detection system so that it avoids the wrong detection and no detection in the system.

`

The Preprocessing helps to achieve the faster result based upon the clarity of the image as an input.

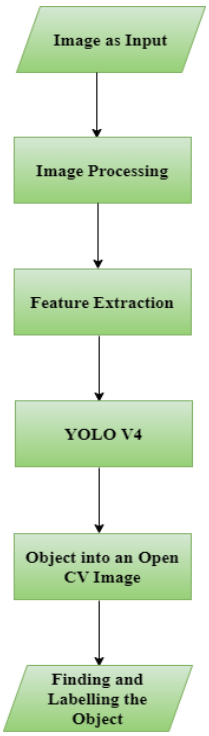

Fig. 1.1 Object Detection Methodology Flow Chart

The input image for the first step is of many categories such as indoor objects, people, vehicles, animals, groceries, birds, and so on. The image is then preprocessed with a modest level of abstraction, which increases the image intensity by conquering adverse characteristics and enhancing it for subsequent processing. The image is scaled, the brightness is normalized, and the contrast impact is reduced to make feature extraction easier. Every alternative pixel is placed in a separate channel by the reorganization layer. The feature extraction involves the maxpool layer; it is used to minimize the size of the representation in order to speed up processing and to improve the robustness of some of the features it detects. The below processes are done by the maxpool layer to achieve the forward pass

- Selection of the maximum values in the receptive fields of the input
- Preservation of the indices
- Creation of volumetric output

To get the ratios of the image, convert the bounding boxes to proper size with the img_width and img_height variables. The rectangle function is used to draw the bounding box around the detected object.The prediction is made with a multinomial probability distribution; the softmax layer is used as the activation function in the output layer. The Dropout layer, which helps prevent overfitting, turns input units to 0 at random with a rate frequency at each step during training time.

The system uses the Batchnorm layer, it is a network layer that is applied between a hidden layer and the following hidden layer. Its main purpose is to normalize the outputs of the first hidden layer before passing them on to the next hidden layer as input. Then the route layer is important for bringing finer-grained information from earlier in the network into the network. A (Recurrent Neural Network) RNN layer iterates over the clock cycles in a sequence and use a for loop while maintaining an internal state that stores information about the clock cycles it has observed so far. The modified YOLO v4 uses OpenCV and Darknet methodology for the faster detection of an object. The scaled modified YOLO v4 weights file that is pre-trained to detect 80 classes in the detection system. The bytes are converted into the numpy array for the processing and the numpy array is converted into the OpenCV BGR image.

By using these methodologies the object is converted into the OpenCV image. The bounding box is used for showing the detected image result. After the bounding box area are marked the system will start the labeling process for the detected object with the library function to compare the object with class which it belongs to, then once the categorization is completed the result will be produced based on the modified YOLO v4 methodology. The Multiple objects are also detected with this system with high accuracy, the Darknet, OpenCV and modified YOLO v4 methodologies helps to attain it.

VOC 2007	R-CNN	YOLO v1	The modified YOLO v1	The modified YOLO v4
Aero	63.5	78	77.9	93.59
Bike	66	74.2	77.6	83.18
Bird	47.9	61.3	63.7	91.72
Boat	37.7	45.7	47.6	91.35
Bus	62.5	68.2	70.7	93.51
Car	70.2	66.8	68.9	62.89
Cat	60.2	80.2	85.3	81.28
Chair	32	40.6	42.2	93.43
Cow	57.9	70.7	71.9	92.43
Table	47	49.8	51.2	66.94
Dog	53.5	79.8	81.9	93.02
Horse	60.1	74.5	77.5	93.46

Table 1: Pascal VOC 2007 Test Result Comparison [7]

`

The results of modified YOLO v1 using R-CNN and YOLO v1 with the same dataset are compared in the preceding report. The YOLO network displays the result in 0.11 seconds. The modified YOLO v4 network detects the item in 0.09 seconds and displays the findings reliably. In comparison to other approaches, the performance time is quite efficient. Because similarity leads to the incorrect recognition of an item, the accurate result is achieved using Multiple Object Detection. The same picture is also tried with other classes. With these circumstances, the outcome is effective and precise.

YOLO is a fast and accurate group of one-stage object detectors. The OpenCV extract the RGB values of the pixels and the train classifier, demo classifier are used for the classification of an object. Dropout is the regularisation technique for eliminating fitting problem in neural networks and it alter the network's structure. During each cycle of training, it loses neurons from the neural network at random. Because various networks will overfit in different ways, dropout will have a net impact of reducing fitting problem, so dropout layer is used for this purpose.

Maximum pooling, often known as max pooling, is a pooling procedure that determines the maximum, or largest, value in each feature map patch. Max pooling chooses the image's lighter pixels. It is helpful when the image's backdrop is dark and concerned with image's lighter pixels. In the case of average pooling, the outputs are down sampled or pooled feature maps that highlight the most prevalent feature in the patch rather than the average presence of the feature, this function can be used as an activation function for a hidden layer in a neural network.

At a bottleneck or concatenation layer, it is applied when the model has to pick or weight multiple different inputs internally. The backbone of the modified YOLO v4 is the CSPDarknet53, this is the reason for the increase of accuracy level of the detection, the network's weight is used for the idetification of an feature of an image in the pre-training process. By using these methodologies the modified YOLO v4 accuracy level is increased and also the result is retrived within short span of time.

IV. RESULTS AND DISCUSSIONS

The abnormal activities in the railways and roadways can be detected using the YOLO techniques with Darknet support. By using the above said methodology various results were retrieved by using different scenario, the object detection would be discussed briefly as follows.

Figure 1. Elephant crossing Railway Track

The fig.1 shows that the elephant were crossing the railway track, the YOLO technique finds the object and also labeled the object as "elephant", if the object detection were not implemented in this scenario its leads to a big accident.

Figure 2. Dog Trapped in the Railway Track

The fig.2 shows that there are two dogs and one of the dog was trapped in the railway track but it couldn't release the leg from the track. The object detection methods identify the object after classifying its category. The results shows that, the trapped dog is accurately detected because the dog are in the lying position but the results shows its 85.34 percent as dog by using their feature. The modified YOLO v4 accurately measures their feature after the feature extraction. The train comes through this track leads to accident which can be avoided using YOLO.

Figure 3. Suicide Attempt by the Person

The above fig.3 illustrates that the person lying down on the track for the suicide attempt, this object detection method can be a life saver for that person if it is detected earlier by the control section. The image or video can be captured by the camera in the front side of the train or may be in the opposite direction of the train and it will help the detection method to work efficiently by detecting the object on the track, here the women laid down on the track for the suicide attempt the YOLO method can detect the train, person on the track and labeled it within in a fraction of seconds. The person can be identified with label as "person" and the percentage of identification is 67.85, even though the person is on the laid down position the YOLO can easily detects the object. The results depict image capturing system in the opposite direction of the train. The detection method not only save the life of that person but also avoid the crashes of the train. The Computer Vision is needed in certain cases like this for the smooth functioning of the systems.

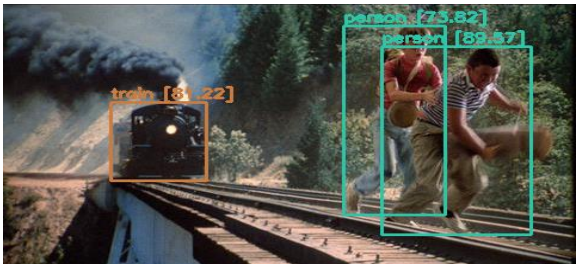

Figure 4. The Person Leg Trapped in the Railway Track

The fig.4 shows that, two children were trying to cross the railway track and one of them were trapped their leg in-between the track and also the train coming too closer to both of them, they have one of two possibilities either they need to release their leg or the train has to be stopped. The previous one have less possibility for the escapism, but the next one has enough possibility for the same. The object detection method detects the person with high accuracy and it helps the officers to save the trapped person's life before the hit.

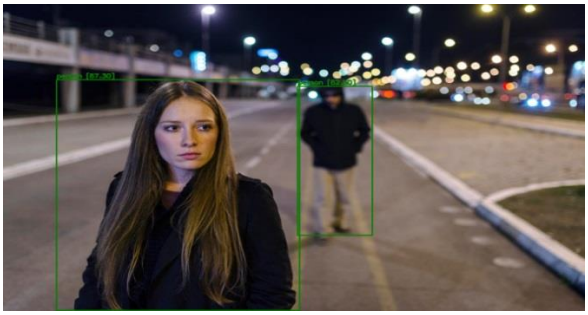

Figure 5. Women followed by Men in the night

The fig.5 denotes that, the women followed by the men at night in the roadways but there is no detection of people other than these two persons so she is in helpless situation. According to research, in May 2020, the crime rate against women surged by nearly 131 percent. To avert such incidents, aberrant behaviours on the highways must be spotted in advance using Computer Vision techniques throughout the day. The above results detects the two person on the road it helps the officers to detect the abnormal activity at night by using these detection method. Then the officers of the control section can take the necessary action to protect that woman from the harassment.

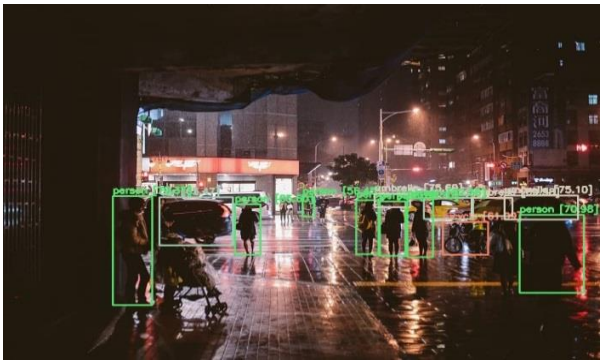

Figure 6. Unwanted Crowd in Night

The fig.6 illustrates that there are more number of peoples were gathered on the roadway at the night time, it is an abnormal activity than the usual day. The detection method can detects the peoples walking on the road very accurately with the percentage of 85.60 as most nearest one and 56.45 as the farthest one. Most of the illegal activities were done at the night time it might be a terror attack or any anti-social activity, this detection helps the officers to avoid any big accidents. The YOLO techniques clearly detect the object with high rate of accuracy and most of the detection methods find difficult to identify the object in the clumsy image or video, but the modified YOLO v4 clearly produce the results with these sample image. The fig.6 has the various kinds of objects such as umbrella, car, bike and person the YOLO easily detects the cars, bike and others with proper labeling without any false detection or no detection. The major problem in other detection is the wrong detection because of the size of the object, for example if the dog is on the standing means its shows the result as dog but the dog is on the laid down position means its shows the results as cat, false detection can be occurred because of the size of the object, its avoided in this modified YOLO v4 algorithm.

V. CONCLUSION

The object detection method helps the relevant department's control centre about the unusual behavior and allowing the officers to take fast action to prevent criminal activities in the region. This method not only saves the life of the people but also save the life of the animals too. The YOLO algorithm easily detects both the living and non-living things within a short span of time. The Computer Vision technology helps the officers to avoid the major accidents and terror attack. The detection time of the modified YOLO v4 is approximately, minimum of 0.02 and maximum of 0.09 seconds. This YOLO approach is more efficient with less detection time and a high percentage of accuracy.

This study gives a process of several object detection techniques that have been proposed by researchers in the past in order to improve the development of an object detection system. All of the strategies discussed have their own set of benefits. The researchers make an attempt to explain the procedures and strategies utilized to discover the item, it will help to understand and compare comparable advancements in other domains.

REFERENCES

1. Mohana, HV Ravish Aradhya, *"Object Detection and Tracking using Deep Learning and Artificial Intelligence for Video Surveillance Applications*", International Journal of Advanced Computer Science and Applications (517-530), Volume 10 No. 12, **2019**.
2. Souhail Guennouni, Anass Mansouri, Ali Ahaitouf,*"Multiple object detection using OpenCV on an embedded platform"*, 2014 Third IEEE International Colloquium in Information Science and Technology (CIST), Tetouan, Morocco, **2014**, pp. 374-377
3. Omkar Masurekar, Omkar Jadhav, Prateek Kulkarni, Shubham Patil4, *"Real Time Object Detection Using YOLOv3",* International Research Journal of Engineering and Technology, Vol. 7 No. 3,pp- 3764 to 3768, March **2020**.
4. Sandeep Kumar, Aman Balyan, Manvi Chawla, *"Object Detection and Recognition in Images"*, International Journal of Engineering Development and Research, Vol. 5 No. 4, pp- 1024 to 1029, 9th **2017**.
5. Ceren Gulra Melek, Elena Battini Sonmez, Songul Albayrak, *"Object Detection in Shelf Images with YOLO",* IEEE EUROCON 2019 -18th International Conference on Smart Technologies, Novi Sad, Serbia, 2019, pp. 1-5.
6. Geethapriya. S, N. Duraimurugan, S.P. Chokkalingam, *"Real-Time Object Detection with YOLO"*, International Journal of Engineering and Advanced Technology, Volume-8, No-3S, February **2019**.
7. Tanvir Ahmad , Yinglong Ma , Muhammad Yahya, Belal Ahmad, Shah Nazir , Amin ulHaq, *"Object Detection through Modified YOLO Neural Network"*, Hindawi Scientific Programming, **2020**.
8. Y. Yoon et all., "*Online Multi-Object Tracking Using Selective Deep Appearance Matching*", IEEE International Conference on Consumer Electronics - Asia (ICCE-Asia), Jeju, **2018**, pp. 206-212.
9. H. S. Bharadwaj, S. Biswas and K. R. Ramakrishnan. "*A large scale dataset for classification of vehicles in urban traffic scenes*", Proceedings of the Tenth Indian Conference on Computer Vision, Graphics and Image Processing, ACM, **2016**.
10. Mohana et.al., "*Performance Evaluattion of background modeling methods for object Detection and Tracking*", International Conference on Inventive systems and Control (ICISC), **2020**.
11. Mohana et.al., "*Elegant and efficient algorithms for real time object detection, counting and classification for video surveil-lance applications from single fixed camera*" International Conference on Circuits, Controls, Communications and Computing (I4C), **2016**.
12. Mohana et.al., "*Simulation of Object Detection Algorithms for Video Survillance Applications*", 2nd International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud), **2018**.

A. Raghunandan et. al., "*Object Detection Algorithms for Video Surveillance Applica-tions*," International Conference on Communi-cation and Signal Processing (ICCSP), **2018**.

13. Z. Zhao, P.S. Zheng, S. Xu and X. Wu, "Object Detection with Deep Learning: A Review", IEEE Transactions on Neural Networks and Learning Systems, vol. 30, no. 11, 2019, pp. 3212–3232,Web: https:// doi. org/10.1109/TNNLS.2018.2876865.
14. Z. Zou, Z. Shi, Y. Guo and J. Ye, "Object Detection in 20 Years: A Survey", 2019, Web: arXiv:1905.05055.

15. L. Liu, W. Ouyang, X. Wang, P. Fieguth, J. Chen, X. Liu and M. Pietikäinen, "Deep Learning for Generic Object Detection: A Survey", International Journal of Computer Vision, no. 128, 2020, pp. 261–318, Web: https://doi.org/10.1007/s11263-019-01247-4
16. Najafi Kajabad, Ebrahim & Begen, Petr & Nizomutdinov, Boris & Ivanov, Sergey. (2021). "YOLO v4 for Urban Object Detection: Case of Electronic Inventory in St. Petersburg". 316-321. 10.23919/ FRUCT 50888.2021.9347622.
17. Zhi-Hao Chen and Jyh-Ching Juang "*YOLO v4 Object Detection Model for Nondestructive Radiographic Testing in Aviation Maintenance Tasks*", AIAA JOURNAL, Vol. 60, No. 1, January 2022, https://doi.org/10.2514/1.J060860.
18. Wu, Han, Chenjie Du, Zhongping Ji, Mingyu Gao and Zhiwei He. "*SORT-YM: An Algorithm of Multi-Object Tracking with YOLOv4-Tiny and Motion Prediction*." Electronics (2021): n. pag.
19. Liu, Tao & Pang, Bo & Zhang, Lei & Yang, Wei & Sun, Xiaoqiang . (2021). Sea Surface Object Detection Algorithm Based on YOLO v4 Fused with Reverse Depthwise Separable Convolution (RDSC) for USV. Journal of Marine Science and Engineering. 9. 753. 10.3390/jmse9070753.
20. Liu, Qi & Fan, Xiaoyu & Xi, Zhipeng & Yin, Zhijian & Yang, Zhen. (2022). Object detection based on Yolo v4-Tiny and Improved Bidirectional feature pyramid network. Journal of Physics: Conference Series. 2209. 012023. 10.1088/1742-6596/2209/1/012023.
21. Bochkovskiy, Alexey & Wang, Chien-Yao & Liao, Hong-yuan. (2020). YOLO v4: Optimal Speed and Accuracy of Object Detection.
22. Kalaivani, R., and C. Manicha. "Object Detection in Video Frames Using Various Approaches.*"International Journal of Advanced Research in Computer and Communication Engineering*" 2, no. 9 (2013): 157-160.
23. Shih, Ya-Fang et al. (2017). "*Deep Co-Occurrence Feature Learning for Visual Object Recognition*" In Proc. Conf. Computer Vision and Pattern Recognition
24. Sanjirani shantaiya et.al, "*A survey on approaches of object detection*", International journal of computer applications, Vol-65, No-18, March 2013
25. John Chiverton, Xianghua Xie and Majid Mirmehdi," *Automatic Bootstrapping and Tracking of Object Contours*", IEEE Transactions on Image Processing, Vol. 21, No. 3, March 2012
26. S. Saravanakumar, A. Vadivel and C.G. Saneem Ahmed, "*Human object tracking in video sequences* "ICTACT Journal on Image and Video Processing, August 2011, Volume: 02, Issue: 01