iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 9 **Issue:** I **Month of publication:** January 2021

DOI: https://doi.org/10.22214/ijraset.2021.32765

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue I Jan 2021- Available at www.ijraset.com

Implementation of Real-Time Object Detection System using Machine Learning Algorithm

Mr. Shismohammad Mulla[1], Dr. Mahesh Chavan[2]
[1]*M.Tech. in Electronics and Telecommunication Engineering,* [2]*Professor in Electronics Engineering, Department of Electronics Engineering, KIT College of Engineering (Autonomous), Kolhapur*

Abstract: ***Object detection is generally important capability required by a variety of computer vision robots and systems. Object detection is one of these fields which are observing enormous success in domain of computer vision. In Object detection system images have diverse types of object this want to be accurately identified by computer vision system. This is very elementary problems of computer vision system. Classical object detection techniques built on simple architectures and handcrafted features. This type of method is very slow and less accurate. However, due to rise of machine learning methods, performance and accuracy of such types of problems is significantly improved. New machine learning method able to greatly increase effectiveness of such systems. Compared with traditional object detection techniques, the deep learning based on objects detection method is incredibly fast and also more accurate. Due to lots of development in high computing power and improved powerful tools, therefore developed system can learn semantic segmentation, high-level understanding and complex deeper features of image to address existing problems in traditional architectures. Real-time object detection and object tracking systems encompass range of applications such as autonomous vehicle, surveillance systems and recognition of license plates. This paper introduces the implementation of the real-time object detection and objects counting system using YOLO algorithm and real-time object counting with their class-names. YOLO algorithm works in real-time. This algorithm is extremely fast and tremendously accurate. The convolutional neural networks (CNN) are basis for these types of algorithms. In this paper, the case study of the YOLO algorithm and how this algorithm can be applied to detection of daily objects for various task and test accuracy of the object detection system.***

Keywords*:* ***Digital image processing, computer vision, machine vision, Convolutional Neural Network, Real-time Object Detection, Object Recognition, YOLO algorithm, Real-time object counting, Artificial intelligence, Deep Learning, Machine Learning .***

I. INTRODUCTION

Object detection is a technique associated with the problem of computer vision and image processing which deals with detecting objects in images or videos of a specific category. There are many uses of object detection and recognition systems, including object counting, facial recognition, character recognition, and autonomous driving in a surveillance camera. The task of such a system is not only to classify various images, but also to classify approximate object positions found in each image [1], in order to provide a full understanding of images and videos. The system must be able to detect, localize and classify several objects using given image or video feed. This is a harder task than only classifying images into distinct classes. Objects detecting and tracking is challenging tasks in surveillance system in which system has to accomplish task to determine meaningful events, suspicious activities, and automatically detect intrusion and provide a summary. Lots of applications of object detection, recognition and tracking are listed below:

A. Crowd counting
B. Video surveillance
C. Human detection
D. Action Recognition
E. Segmentation
F. Vehicles Detection
G. Autonomous Vehicles
H. Objects Counting

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue I Jan 2021- Available at www.ijraset.com

Object detection and recognition is one of the most significant problems in computer vision. As a significant problem in computer vision system, object detection task can provide essential information for semantic perception of images and videos. Several issues in the world of computer vision have saturated their accurateness for two decades. Thanks to the growth of deep learning algorithms, such object detection area is fast growing. Object detection is a common essential problem in computer vision, and it has been deemed a very difficult problem to overcome for a long period of time. Today, the accuracy of object detection has dramatically improved with the advancement of deep learning methods; a convolutional neuronal network (CNN) is another type of deep learning neural networks structure, it has many hidden layers. CNN is influenced by structures that are found in the visual cortex of humans. In image recognition problems, CNN is able to overtake humans. This technical advancement attempts to automate tasks that can be performed by humans. All of this development is not due to a just too very powerful hardware, a ton of datasets and a better model, but also modern methods, algorithms and enhanced network architecture. The old Fast R-CNN algorithm by girshick et al [2] is best case of algorithm in computer vision. This detection algorithm similar to YOLO performs remarkably fast and accurate. YOLO is a novel deep learning approach to object detection task to perform detection. In the YOLO algorithm, the authors reframe the object detection problem as a regression problem to spatially separate bounding boxes and an associated class of probabilities. This algorithm is extremely speedy, efficient and accurate. YOLO model can processes images in real-time at 45 FPS [3The YOLO network's lighter model is known as fast-YOLO, which processes 155 frames per second. We are going to present the implementation of the real-time object detection method in this paper, and objects counting using YOLO algorithm and object counting with their class names. The case study of the YOLO architecture as well as how the YOLO algorithm is used to detect everyday objects for different tasks and measure the reliability of the object detection system is further discussed in the paper.

II. LIERATURE REVIEW

There are a large number of authors, who has used different types of object detection and recognition methods; from which YOLO outperform other methods in terms of speed and accuracy.

Paul Viola and Michael Jones published a research paper in 2001, on "Rapid object detection using a boosted cascade of simple features". They established a method for the task of object detection in this research. They archived object detection architecture in this design, which minimizes processing time while achieving high accuracy of detection. The researcher has built an algorithm in this project to build a face detection system that is 15 times faster than almost any prior approach [4]. The model developed by the researchers produces new representations of the system and observations. This is much more popular in image processing and computer vision that had extensive application.

Their study is split into three main parts. The first element is an integral image that facilitates very fast calculation of the characteristics used by the detector. The second part is a learning-based algorithm typically based on ADA-boost. In third part is a cascade method is used for combining major complex classifiers.

Navneet Dalal, Bill Trigs et al presented a research paper in 2005, on "Histograms of Oriented Gradients for Human Detection". In this paper, researchers created a methodology that shows that istograms of oriented gradients (HOG) descriptors significantly outperform current human detection feature sets. They had used human detection model based on a linear Support Vector Machine (SVM). The system model developed in this research can operate flawlessly on the MIT dataset. Throughout this study, the researcher used locally normalized histogram of the gradient orientations features like SIFT such as SIFT descriptors in a dense overlapping grid that provides very great person detection performance. Authors concluded that fine orientation binning, relatively coarse spatial binning, fine-scale gradients, and high-quality local contrast normalization in overlapping descriptor blocks are similarly significant for good results [5].

Alex Krizhevsky and Ilya Sutskever presented a research paper on "Image Net Classification with Deep Convolutional Neural Networks" in 2012. A supervised machine learning model was developed by them. Throughout this algorithm, they majorly used Convolutional Neural networks in their method. Including 60 million parameters and 650K neurons, that model employed convolutional including 60 million parameters and 650K neurons, that model employed convolutional A total of five convolutional layers are composed of this convolutional neural network model. Few of them is followed by max-pooling layers as well as the other three were followed by fully-connected (FC) layers with something like a soft-max layer within final 1K way[6]. They also use non-saturating neurons as well as an enormously potent GPU for its development of the convolution operation to allow training of such a neural network easier. A neural network which gets 1 week to train on dual GTX 580 3GB GPU is included in this developed model. A Convolutional Neural Networks model built by them can lead to a large convolutional neural network able to achieve record breaking performance.

Ross Girshick, Jeff Donahue et al have published a research paper on presented a paper on "Rich feature hierarchies for accurate object detection and semantic segmentation". A supervised machine learning algorithm was developed by them. They developed regions with convolutional networks (R-CNN) model within is machine learning model [7]. On the PASCAL VOC dataset, the authors evaluated the R-CNN model. They have implemented a simple and scalable detection algorithm that, with more than 30 percent mAP, can increase mean average accuracy (mAP). The deconvolutional neural network was used to take 227x227 resolution images only at input nodes. The system developed is very precise and incredibly effective. R-CNN can also be expanded to hundreds and thousands of groups of objects. R-CNN model Pre-training on open source CAFFE CNN open library has been evaluated and the results. The optimizer like stochastic gradient descent (SGD) is used for RCNN model. Trained and optimized SVMs on the VOC 2012 dataset following the implementation of the RCNN model tested on the PASCAL VOC 2007 dataset and fine-tuned the CNN model on the VOC 2012 dataset [7].

Joseph Redmon, Ali Farhadi et al, published a research paper on "Real-Time object detection a state of the art algorithm based on convolutional neural networks". They also made different changes to original version of YOLO object detection. The latest enhanced YOLO algorithm is incredibly fast and precise. The YOLO machine learning algorithm proposed by the authors can recognize and detect over 9,000 object categories. They frame the problem of object detection as a regression problem in this design, straight from image pixels to bounding box coordinates and their class probabilities. While training and testing period, YOLO examines the whole picture, which fully encodes contextual information about classes and its appearance [3]. To training and testing techniques, they used the NVIDIA Titan X GPU. Convolutional layers of neural net select features from images in the designed process. There are 24 convolutional layers (CL) of their Neural Network followed by 2 fully connected (FC) layers. Up to 45 FPS can be processed in real time with YOLO architecture images, which is a state of the art speeds. It defeats other well-known models of detection, such as the algorithm for Fast R-CNN and DPM. They trained the YOLO model on the COCO dataset and the Image-Net dataset using the optimization method for stochastic gradient descent (SGD).

Aysegul Ucar, Yakup Demir et al. in 2017 has published research paper on "object recognition and detection with deep learning for autonomous driving applications". In this study, the author developed an autonomous vehicle model and presented details on how to use real-time object detection system in autonomous vehicles and their issues associated. In an autonomous vehicle, occlusion induces incorrect identification and detection in object detection system. A new Local Multiple Method and Support Vector Machines algorithm has been proposed by authors that can easily manage these all difficulties. A new model called hybrid Local Multiple System and Support Vector Machines based on deep convolutional Neural Networks for object recognition and detection of pedestrians is proposed in this paper [8]. In this process, the entire image is first divided into local regions and multiple fully convolutional Neural Networks are used to extract discriminatory features and then pick unique characteristics used by the main component study (PCA).

III. OLD METHOD

In object detection task images contains diverse types of objects which can be accurately identified by the computer vision systems. This is the fundamental problems of the computer vision systems. Currently, traditional object detection methods are built on handcrafted features and simple architectures. This is very slow and less accurate. Many of the techniques in computer vision have been saturating on their accuracy last a decade [9].

IV. NEW METHODOLOGY

The accuracy of computer vision problems has improved significantly due to advancement of deep learning models and computing performance. The usefulness of a real-time object detection method can be exponentially increased by many different deep learning algorithms in the computer vision field.

This kind of DL-based approach is extremely fast, extremely precise and highly adaptable. In addition, YOLO is a related algorithm based on deep learning applied to the detection task for real-time object detection. This seems to be an extremely speedy, effective and precise algorithm.

The YOLO is an algorithm for real-time object detection that uses a convolutional architecture of the neural network beneath. YOLO algorithm train on full image and directly optimizes detection performance. YOLO algorithm sees the complete image during the training process and testing process, thus it will inherently encode the contextual information concerning classes and also their appearance [3]. The YOLO algorithm can learn generalized object descriptions, such that the algorithm defeats many successful detection systems when the model is trained on natural images/video frames and evaluated on the new data set.

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.429
Volume 9 Issue I Jan 2021- Available at www.ijraset.com

A. Block Diagram

Figure 1 shows the complete working block diagram of YOLO algorithm. At the input side of the block, image acquisition takes place. The input may be in form of a single image or a video frame. In the pre-processing block, the image is resized to a specific resolution. Then this image cropped.

Thus therefore using the YOLO algorithm divides the input image/video frame into an S×S grid (S = n). If centre of the object falls on particular grid cell then this grid cell is accountable for detection of that object. If the centre of the bounding box did not fall in a specific grid cell, then that cell will not responsible for it. Each grid cell predicts only one object [13]. For example, in Figure 2 red cell tries to predict the “car” object whose centre (the red dot) falls inside the grid cell. Every grid cell predicts a fixed number of boundary boxes.

In Figure 4, the grid cell makes the two bounding boxes (blue boxes) for object prediction to locate where the "Car" object is. In the YOLO algorithm, nevertheless, due to one object rule, there is some restrictions to how objects are close to each other. In algorithm for every grid cell, algorithm will predict B boundary boxes and every boundary box has one box confidence score. After bounding boxes are drawn, algorithm will predicts (C) conditional class probabilities. There are 5 components: (x, y, w, h, class confidence) to bounding box prediction.

The x, y coordinates represent the centre of the bounding box, relative to the grid cell location. The w and h are widths and heights respectively.

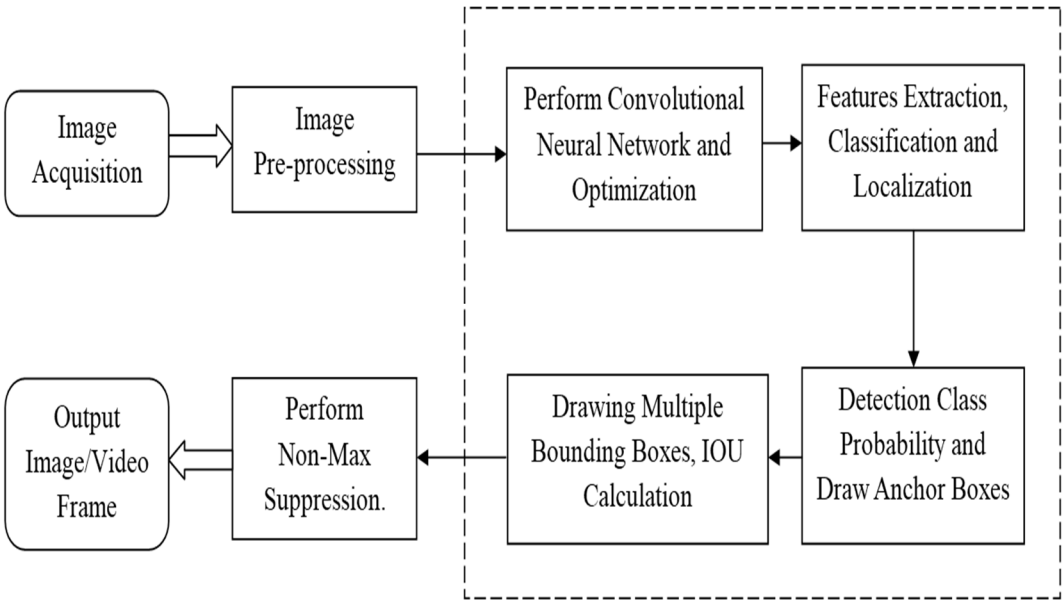


Fig. 1 Block diagram

B. Flowchart

Figure 2 show the flowchart of algorithm. At the beginning of the algorithm, image or video frame is feed. Then input image resized to NxN resolution. Then convolutional neural network is utilized. Algorithm normally have 24 convolutional layers (CL) succeeded by 2 fully connected (FC) layers.

This will perform convolution operation input image at a particular instances. After that YOLO algorithm divide the input image/video frame into an S×S grid. One grid cell predicts only one object [13]. Then the image classification and localization operations are performed to obtain information about objects which are present in given image.

If any object is detected in a grid then class probability is calculated for that object. It represents probability of existence of an object within the bounding box [14].

After that, algorithms will predict bounding boxes for every object and all bounding box has a box confidence score. Subsequent to this non maximum suppression is performed to remove unwanted bounding boxes. Thus detections are made, after that the counting of objects is done.

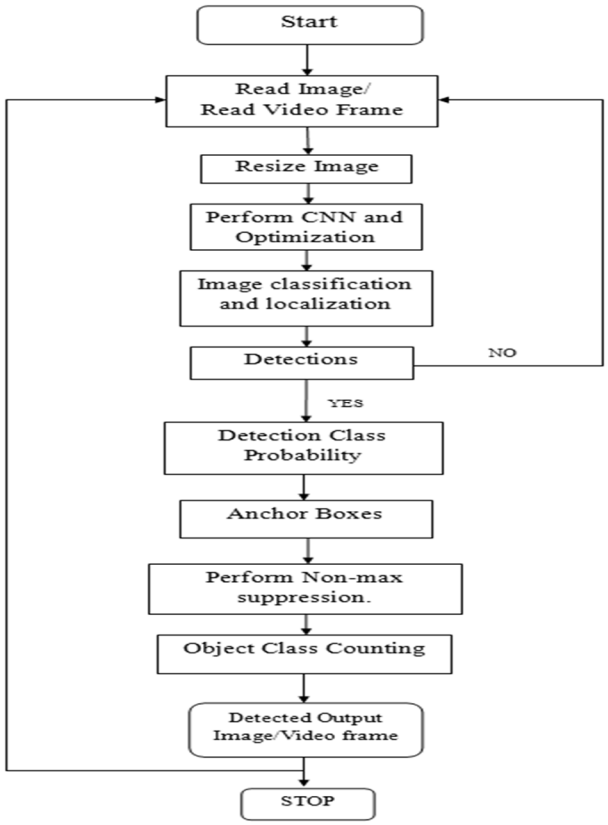


Fig.2 Flowchart

C. Class Confidence Score

This will measure confidence on both the classification and the localization (where an object is located) [13]. Class confidence score can be calculated as:

$$\text{Box Confidence Score} = P_r(\textit{Object}) \times \textit{IoU}$$

$$\text{Conditional Class Probability} = P_r(\textit{Class}_i \mid \textit{Object})$$

$$\text{Class Confidence Score} = Pr(\textit{Class}_i) \times \textit{IoU}$$

$$\text{Class Confidence Score} = \textit{Box confidence score} \times \textit{Conditional class probability}$$

$$\text{Class confidence score} = P_r(\textit{Object}) \times \textit{IoU} \times P_r(\textit{Class}_i \mid \textit{Object}) = Pr(\textit{Class}_i) \times \textit{IoU}.$$

Where,

$P_r(Object)$ is probability that the box contains an object.

IoU is intersection over union between the predicted box and the ground truth.

$P_r(Class_i \mid Object)$ is the probability that the object belongs to $class_i$ given that an object is present.

$Pr(Class_i)$ is the probability that object belongs to $Class_i$

The following formula describes how YOLO algorithm gets bounding box and prediction scores:

$$b_x = \boldsymbol{\sigma}(t) + C_x$$

$$b_y = \boldsymbol{\sigma}(t_y) + C_y$$

$$b_w = Pw\, e^{tw}$$

$$b_h = Pw\, e^{th}$$

Where b_x, b_y, b_w, b_h are x, y center coordinates, width and height of prediction box. t_x, t_y, t_w, t_h is the network outputs. c_x and c_y is top left co-ordinates of the grid cell. p_w And p_h are anchors dimensions for the box [13].

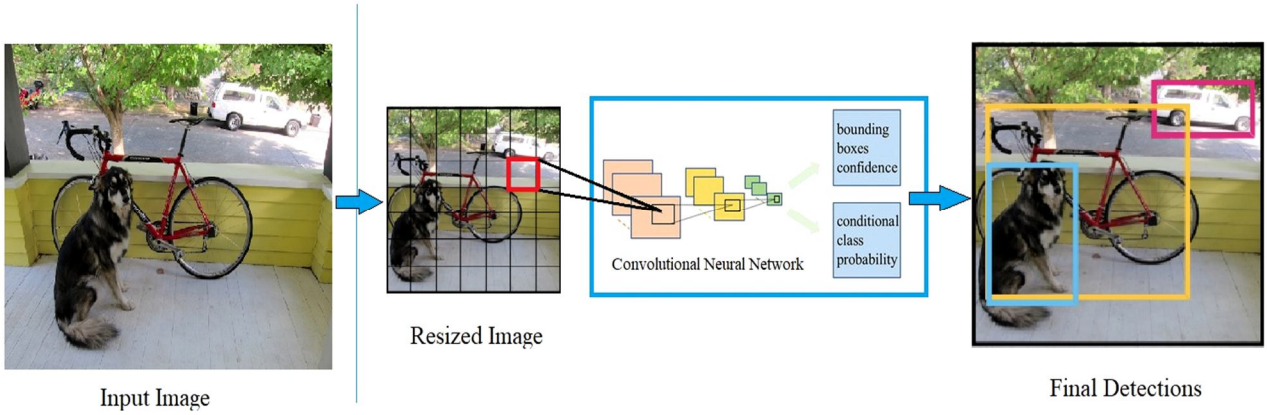


Fig. 3 Working Diagram

The ultimate predicted value is confidence (Pc). It represents the probability of presence of an object in bounding box. The figure 5 shows how intersection of union (IOU) is works. Anchor boxes are drawn to reduce computing time. Predicted bounding boxes look like the figure 6. If the confidence score is max then thicker bounding box is drawn:

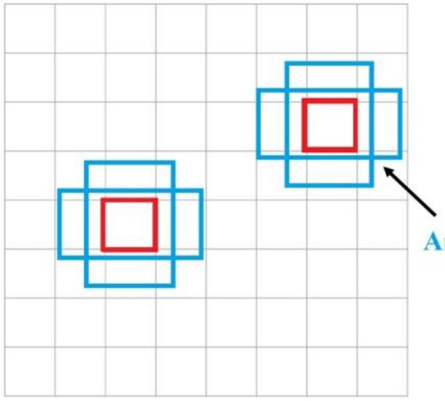


Fig. 4 Anchor boxes

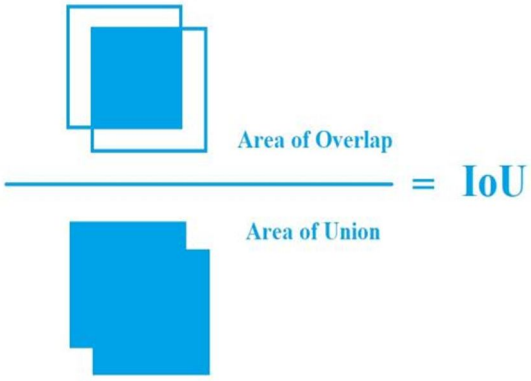


Fig.5 IOU calculation

The confidence score is max then thicker bounding box is drawn:

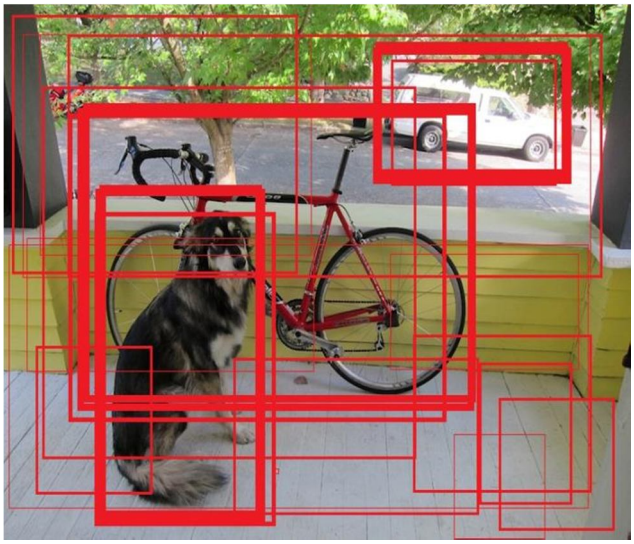


Fig. 6 bounding box configuration

D. Non-Maximum Suppression (NMS)

YOLO makes duplicate detections for same object. Most of these boxes not contain any object also make lots of boxes to same object. To fix this problem, generally in algorithm uses non-max suppression (NMS) to remove bounding boxes with very low object probability.

Fig.7 Non-maximum Suppression

V. RESULTS AND DISCUSSION

A. Output on Python Command Shell

Following figure shows output in anaconda prompt. The created anaconda environment have python=3.7 version, Tensor-flow 2.2 GPU version, Scipy, Keras, PIL, OpenCV and other libraries which required by our project. At output of command prompt python shell our program detect objects, count total number of object and average FPS as following fig 8 shows:

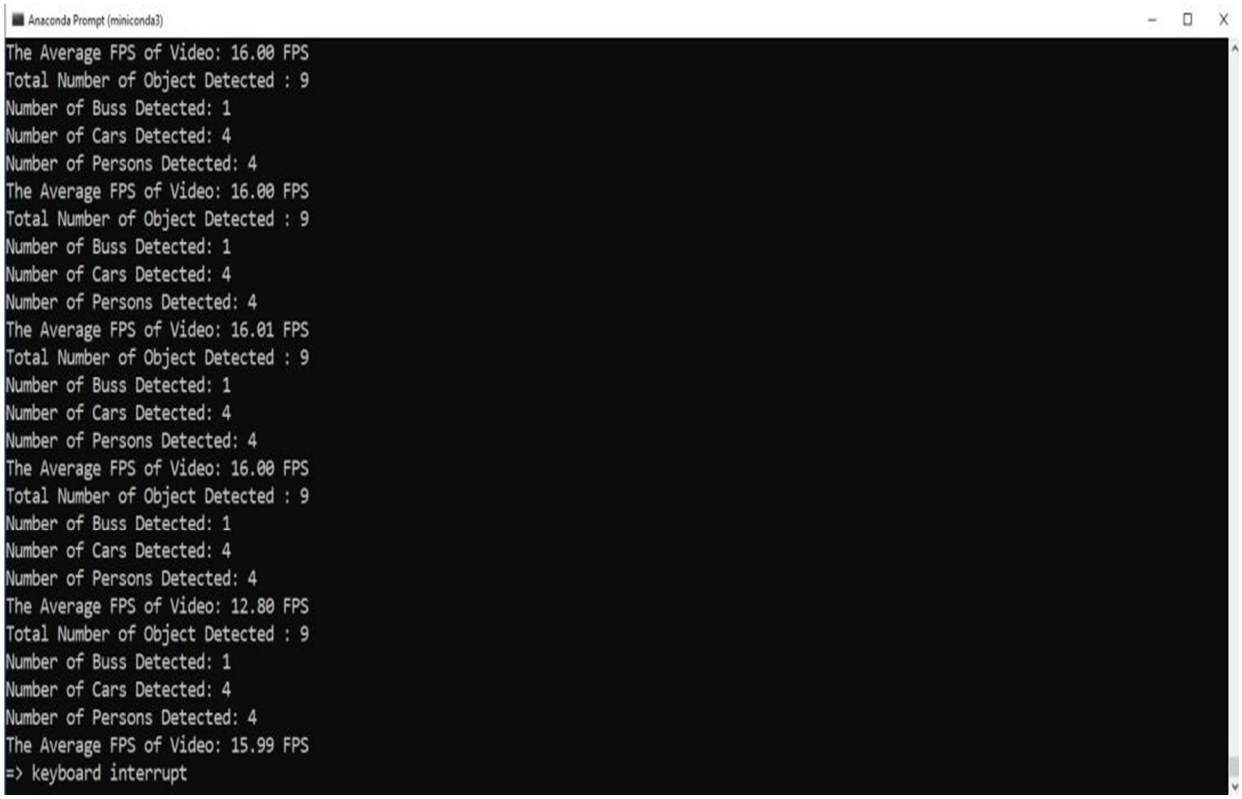


Fig.8 Output on anaconda conda command prompt

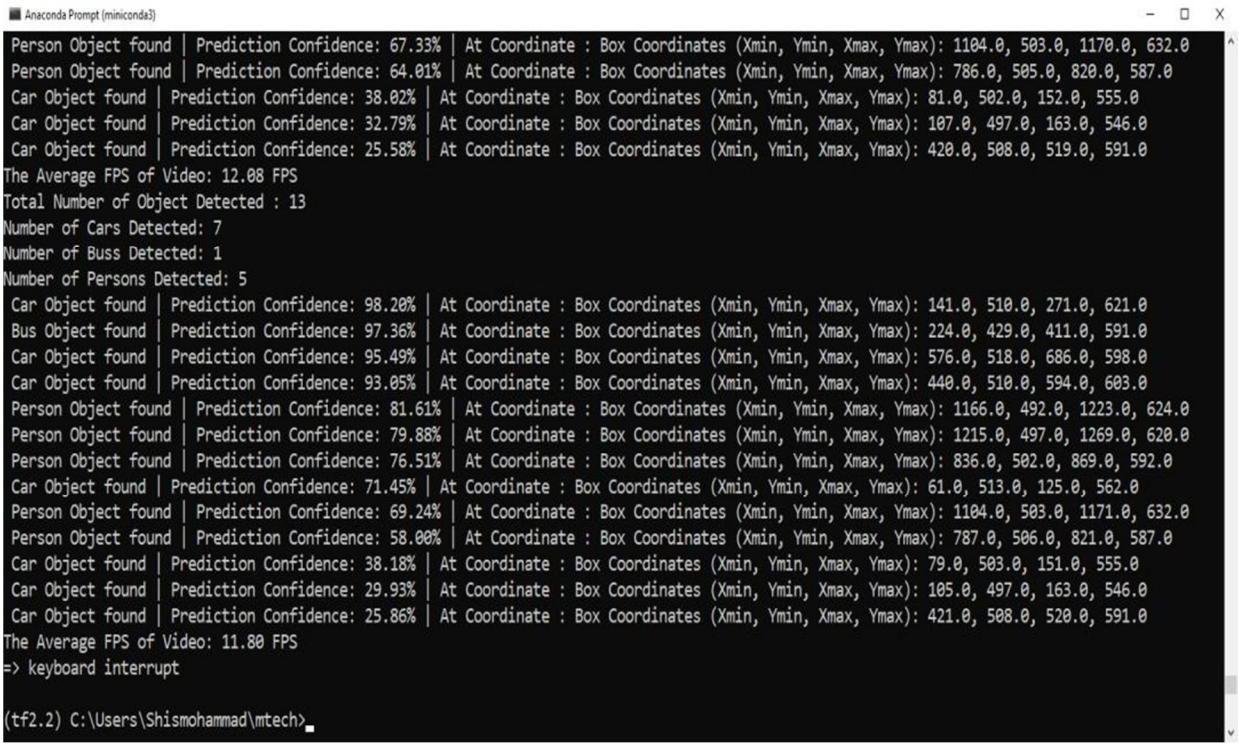


Fig.9 Output of detected object box coordinates

B. Output Video Window

The figure 10 and 11 shows that output of the object detection and counted object program. The program will print counted object with their classes.

Fig.10 Detected Output in video

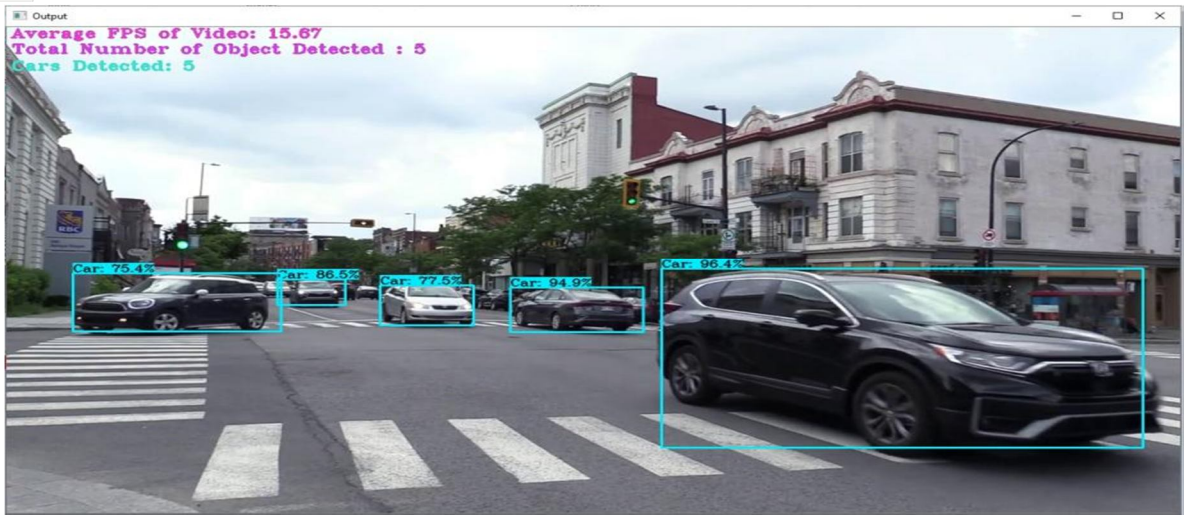


Fig.11 Output of Cars Detected in video

VI. CONCLUSION AND FUTURE SCOPE

A. *Conclusion*

This project is based on the implementation of YOLO algorithm which is machine learning based real-time object detection technique. We implemented real time object detection and object counting with their class system in python with use of Tensorflow and Open-CV library. This system works in real time. In this project, we counted the total number of objects which are present in image or current video frame. We utilises laptop's NVIDIA 940M GPU to execute tiny YOLO algorithm which can gives an average of 15 frames per seconds. This paper gives comprehensive implementation of deep learning based YOLO object detection algorithm which can handle diverse sorts of nuisance of such as occlusion, clutter and low resolution. Most important objective of this project is to study and understanding of problem in contexts of real time object detection system for implementation for autonomous vehicles.

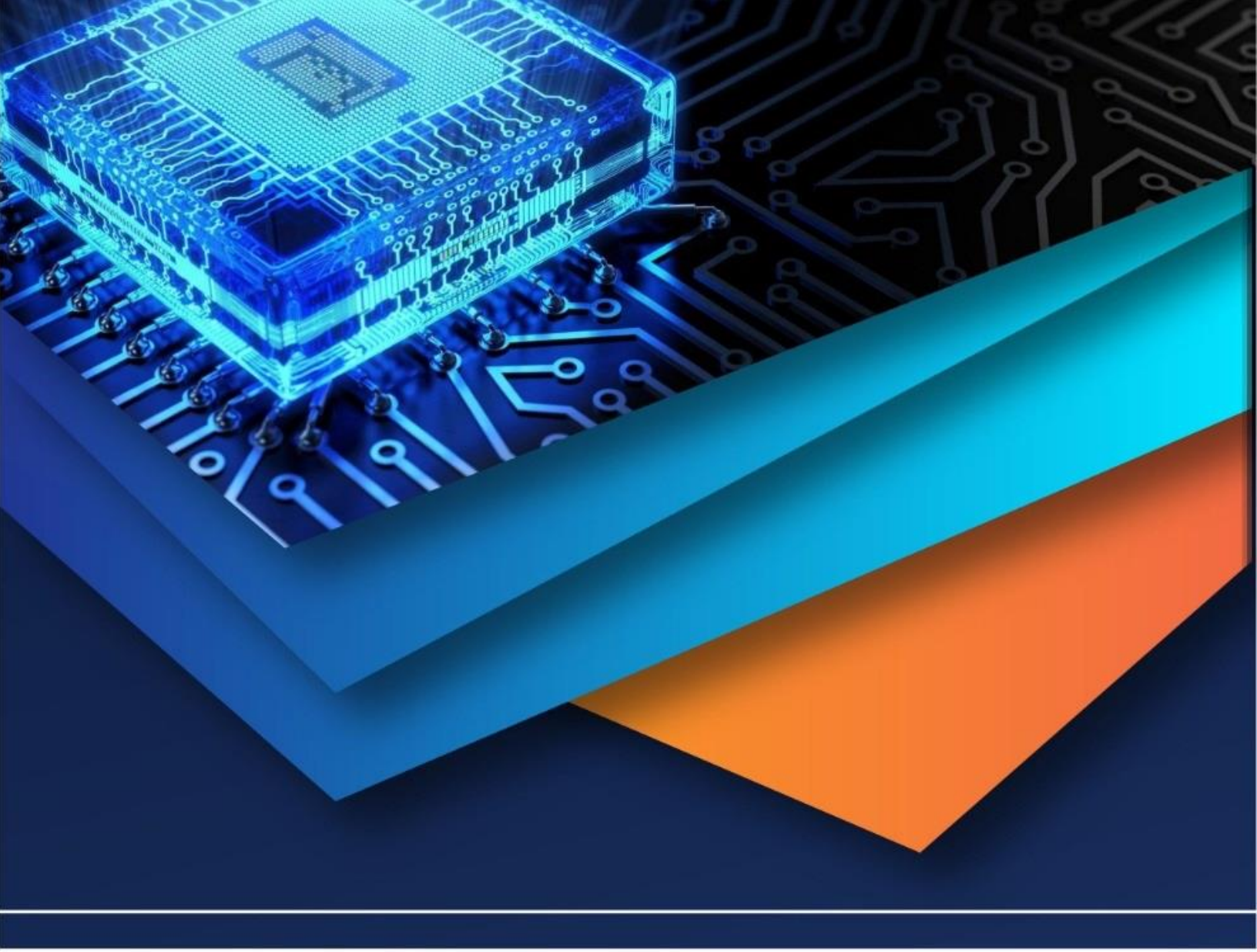
B. *Future Work*

Systems would probably result in much better accuracy and object detection speed with a human level of understanding rapidly with the advent in computing hardware and novel techniques. Autonomous driving models are based mostly on environment's variability and its impact on features [23]. Creating a computer vision system for autonomous vehicles will bring secure autonomous vehicles to roadways with a wide range of problems. The real-time object detection algorithm currently appears to be sufficient for high-precision tasks such as robotic surgery and autonomous vehicles. However, In future, it is possible to develop a deep learning algorithm which can beat human in terms of speed and accuracy, thus in return can execute more complex tasks. Also we can make use of this YOLO model to detect licence plates and perform OCR to carry out licence plate recognition using. At present this project made up of only software and webcam. The CCTV based system can be added to take real time video feed and take images in real time to correct surveillance monitoring at every interval of time.

REFERENCES

[1] Zhao, Z. Q., Zheng, P., Xu, S. T., & Wu, X. (2019). Object detection with deep learning: A review. *IEEE transactions on neural networks and learning systems*, *30*(11), 3212-3232.
[2] R. Girshick, Jeff Donahue, Trevor Darrell and Jitendra Malik “Fast R-CNN” arXiv preprint arXiv: 1504.08083, 2015.
[3] Joseph Redmon, Santosh Divvala, Ross Girshick and Ali Farhadi "You Only Look Once: Unified, Real-Time Object Detection" University of Washington, Allen Institute for AI, Facebook AI Research, arXiv: 1506.02640v5 [cs.CV] 9 May 2016.
[4] Paul Viola and Michael Jones "Rapid Object Detection using a Boosted Cascade of Simple Features" in 2001.
[5] Dalal and B. Triggs. “Histograms of oriented gradients for human detection” In CVPR, 2005.
[6] Alex Krizhevsky, Ilya Sutskever ad Geoffrey E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks" January 2012.
[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik. “Rich feature hierarchies for accurate object detection and semantic segmentation” In CVPR, 2014.
[8] Uçar, Ayşegül, et al. “Object Recognition and Detection with Deep Learning for Autonomous Driving Applications.” SIMULATION, vol. 93, no. 9, Sept. 2017, pp. 759–769, doi: 10.1177/0037549717709932.

[9] Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1, pp. I-I). IEEE.
[10] Pulkit Sharma "Yolo Framework: Object Detection Using Yolo." *Analytics Vidhya*, 24 May 2020, www.analyticsvidhya.com/blog/2018/12/practical-guide-object-detection-yolo-framewor-python/.
[11] Panwar, Himanshu. "Convolutional Neural Network: Basic Concepts." *Medium*, Analytics Vidhya, 3 Apr. 2020, medium.com/analytics-vidhya/convolutional-neural-network-basis-concepts-e059a76d9161.
[12] Singh, Garima. *Gradient Descent Algorithm*. 10 Aug. 2020, medium.com/swlh/gradient-descent-algorithm-3d3ba3823fd4.
[13] Hui, Jonathan. "Real-Time Object Detection with YOLO, YOLOv2 and Now YOLOv3." *Medium*, Medium, 27 Aug. 2019, medium.com/@jonathan_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088.
[14] "YOLO Deep Learning: Don't Think Twice." *MissingLink.ai*, missinglink.ai/guides/computer-vision/yolo-deep-learning-dont-think-twice/.
[15] P. F. Felzenszwalb, R. B. Girshick, D. Mcallester, and D. Ramanan, "Object detection with discriminatively trained part-based models," IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 9, p. 1627, 2010.
[16] Sumit Das and Aritra Dey "Applications of Artificial Intelligence in Machine Learning: Review and Prospect" International Journal of Computer Applications (0975 – 8887) Volume 115 – No. 9, April 2015.
[17] *Towards Data Science*, towardsdatascience.com/.
[18] Convolutional Neural Networks, accessed 11 May 2007, <https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53>.
[19] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bulletin of Mathematical Biology, vol. 5, no. 4, pp. 115–133, 1943.
[20] Subhashini Venugopalan, Marcus Rohrbach and others "Sequence to Sequence – Video to Text" arXiv: 1505.00487v3 [cs.CV] 19 Oct 2015.
[21] Shaoqing Ren and R. Girshick "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks" arXiv: 1506.01497 [cs.CV].
[22] P. Dollár, C. Wojek, B. Schiele and P. Perona "Pedestrian Detection: An Evaluation of the State of the Art" PAMI, 2012.
[23] Paul Viola and Michael J Jones. "Robust real-time face detection" IJCV, 57(2):137–154, 2004.
[24] C. Chen, A. Seff, A. Kornhauser and J. Xiao "Deep Driving: Learning Affordance for Direct Perception in Autonomous Driving" 15th IEEE International Conference on Computer Vision ICCV, 2015.
[25] J. E. Hoo, and K. C. Lim. "Accuracy and Error Study of Horizontal and Vertical Measurements with Single View Metrology for Road Surveying" ARPN Journal of Engineering and Applied Sciences. 11(12):7872-6, 2016.
[26] Lee CS, Tyring AJ, Deruyter NP, Wu Y, Rokem A, Lee AY. "Deep-learning based automated segmentation of macular edema in optical coherence tomography" Biomed Opt Express. 2017; 8(7):3440-3448. Published 2017 Jun 23. doi:10.1364/BOE.8.003440.
[27] Marra, Francesco & Poggi, Giovanni & Sansone, Carlo & Verdoliva, Luisa (2017),"A Deep Learning Approach for Iris Sensor Model Identification. Pattern Recognition" Letters. 113. 10.1016/j.patrec.2017.04.010".
[28] J. Long, E. Shelhamer, and T. Darrell "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15), pp. 3431–3440, IEEE, Boston, Mass, USA, June 2015.
[29] S. Ren, K. He, R. Girshick, and J. Sun. "Faster R-CNN: Towards real-time object detection with region proposal networks" arXiv: 1506.01497, 2015.
[30] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," Bulletin of Mathematical Biology, vol. 5, no. 4, pp. 115–133, 1943.
[31] W. Ouyang, X. Zeng, X. Wang et al., "Deep ID-Net: Object Detection with Deformable Part Based -amazing-examples-of-computer-and-machine-vision-in-practice/#4da402eb1018.
[32] N. Doulamis, "Adaptable deep learning structures for object labelling/tracking under dynamic visual environments," Multimedia Tools and Applications, pp. 1–39, 2017.
[33] S. Cao and R. Nevatia, "Exploring deep learning based solutions in fine grained activity recognition in the wild," in Proceedings of the 2016 23rd International Conference on Pattern Recognition (ICPR), pp. 384–389, Cancun, December 2016.
[34] H.Noh, S.Hong, and B.Han, "Learning de convolution network for semantic segmentation," in Proceedings of the 15th IEEE International Conference on Computer Vision, ICCV 2015, pp. 1520–1528, Santiago, Chile, December 2015.
[35] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR'15), pp. 3431–3440, IEEE, Boston, Mass, USA, June 2015.
[36] P. Arbel´aez, B. Hariharan, C. Gu, S. Gupta, L. Bourdev, and J. Malik. "Semantic segmentation using regions and parts" In CVPR, 2012.
[37] H. A. Rowley, S. Baluja and T. Kanade. "Neural network based face detection" TPAMI, 1998.
[38] P. Sermanet, K. Kavukcuoglu, S. Chintala, and Y. LeCun. "Pedestrian detection with unsupervised multi-stage feature learning". In CVPR, 2013.
[39] K. Simonyan and A. Zisserman."Very Deep Convolutional Networks for Large-Scale Image Recognition". arXiv preprint, arXiv:1409.1556, 2014.
[40] R. Girshick, P. Felzenszwalb, and D. McAllester. "Discriminatively trained deformable part models".
[41] Mr. Shismohammad Mulla, Dr. Mahesh Chavan."*Application of Machine Learning in Computer Vision: A Brief Review*", Volume 8, Issue VII, International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 508-516, ISSN : 2321-9653, www.ijraset.com
[42] Tai Sing Lee and David Mumford "Hierarchical Bayesian inference in the visual cortex" 1434 J. Opt. Soc. Am. A/Vol. 20, No. 7/July 2003.
[43] Machine Learning "Machine Learning." *Wikipedia*, Wikimedia Foundation, 20 Oct. 2020, en.wikipedia.org/wiki/Machine_learning.
[44] Karpathy, Andrej, and Li Fei-Fei. "Deep visual-semantic alignments for generating image descriptions." Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.
[45] Sharma, Nikita. "Introduction to Basic Object Detection Algorithms." *Medium*, Heartbeat, 2 Mar. 2020, heartbeat.fritz.ai/introduction-to-basic-object-detection-algorithms-b77295a95a63.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY