iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 10 **Issue:** VI **Month of publication:** June 2022

DOI: https://doi.org/10.22214/ijraset.2022.44069

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: 7.538
Volume 10 Issue VI June 2022- Available at www.ijraset.com

3d Image Processing Robot

Ridhiman Kumar[1], Pulkit Singhal[2], Sanchit Goel[3]
[1, 2, 3]*Electronics And Communication, Maharaja Agrasen Institute Of Technology*

Abstract: ***For an 3D IMAGE PROCESSING robot to truly succeed, there must be a variety of ways for robot to interpret its surrounding environment, both quickly and accurately. This report focuses on investigating improvements for a sub- problem of environmental perception, namely 3D image processing, referring to localizing and classifying objects of interests in a specified environment. The objects of interests in a process security scan of baggage or to analyse scans of materials to understand their structure. Recently python3 & machine learning programming approaches have made significant progress in the field of 3d object detection. Thus, several techniques have been developed for the real time tasks such as the autonomous robot. However, as images lack some information about depth which is essential for environmental perception for autonomous robot, they struggle a lot to preserve distance between object and robot. So, in this project in order to avoid such delays we have used high quality camera OF 1080p. This Report investigates the performance of both the minicomputers and as well as the cameras being mounted on the robot for 3d analysis which uses the state-of-the-3d image processing models. The end-to-end models of python3 and OpenCV are ultimately performed and experimented various times to perform 3D Detections. The models were trained and evaluated on the recently released data set and the most promising model was able to detect the 3d image of the object.***
Keywords: Python3, Machine Learning, OpenCV, 3d image processing, 3d Object Detection

I. INTRODUCTION

3D image processing robot represents one of the most notable challenges present in modern computer science and electronics. The possibilities that AD may provide in peoples` everyday life are vast [7]. Partly in terms of saving lives as most driving accidents have been linked to the human error, but also in terms of lifestyle as time spent commuting could, at least in theory, be spent more productively. One key technology needed for succeeding in the field of 3D image processing robot is environmental perception, specifying a way for robot to interpret their surroundings with 3D image processing [4] .

It is incredibly crucial for safe and reliable 3D image processing Robot that this technology is both accurate and fast. Firstly, accuracy is vital as the interpreted environment should reflect the surroundings as correctly as possible [3]. A system with accurate environmental perception can better perform tasks such as route planning, avoiding objects, how is the distance between Robot and Object. Secondly, the speed at which the perception is applied is also vital as 3D image processing robot should be utilized in real-time. If a system is not fast enough the driving situation may have changed from the last environment interpretation, making the processed information irrelevant [2].

A subproblem to environmental perception is 3d Object Detection with image processing. The proposed master's thesis focuses on investigating the improvement of 3D object recognition by image processing in everyday situations. Simply put, 3d object Recognition with image processing means identifying and classifying specific objects in a specific environment and measuring the distance between the robot and the objects. Status parking for city cars can be unusual or complex. This is a problem for pedestrians and other objects with different appearances that are commonly and reliably detected. One way to solve this problem is to run it in OpenCV using the python3 image processing method [1].

II. SENSOR INPUTS

A. Camera

When acquiring data from the actual world, one of the options is to use a camera. Images are made by projecting a real-world image onto a two-dimensional plane and collecting the light intensity and frequency recorded at each projected location. The values of the projected sites are saved as RGB pixel values [10].

As there is no depth in conventional 2D photos, this simply means that information about how far away objects are lost. Because one of the most important concerns in environmental perception is object distance, 3D image processing robot models do not only rely on cameras as sensory input [12]. Furthermore, cameras, like input sensors, are sensitive to changes in light and weather. Because information is saved in the form of intensity.

©IJRASET: All Rights are Reserved

III. OBJECT DETECTION USING IMAGE PROCESSING AND ITS ASPECTS.

Object detection is a middle thing of many robotic belief pipelines, and is a difficult computer imaginative and prescient challenge. The item detection challenge includes subtasks, which can be the localization challenge and the classication challenge. More specically, given an picture I, the approach must generate 2D bounding boxes B1:::Bn, and classication ratings C1:::Cn, wherein n is the variety of detected objects . Bi is a 2D bounding container which may be represented via way of means of its centre role and width and top dimensions, Bi = bu; bv;w; h, or may be represented with their pinnacle left and backside proper nook positions Bi = (u1; v1; u2; v2). The classication rating represents an expected possibility that the detected item is of a sure class. The 3-D item detection challenge extends the 2D challenge right into a 3-D estimation. In addition to the classication rating, a way must produce the 6-DOF pose estimation of every item represented via way of means of a 3-D centroid, T = (tx; ty; tz), bounding container dimensions D = (dx; dy; dz), and orientation $\Psi = (\Theta, \Phi, \psi)$.

A. Feature Extraction

Letting a computer to understand content of an image is just a set of all values of a matrix, it is a very challenging task. Especially as the image data vary depending upon the conditions like lightening conditions. The general solution is to learn general features from the actual data to try and to find some representation of the objects from that separates subject of interest from the object and the other subjects. These kinds of networks can assign significance withinside the shape of weights to particular homes in patches, developing greater summary characteristic representations. The networks are skilled via way of means of being proven examples of inputs and their meant outputs. Finding which homes withinside the enter are vital to make the favoured decisions. Given sufficient examples, ideally, the community could be capable of commonly differentiate one elegance from some other via the assist of the discovered features.

B. Region Proposals

One issue regarding 3D IMAGE PROCESSING is the no of the objects i..e.. the size of the objects. There are the several ways to solve this problem. One way can be the sliding window method, that may as it suggests consists of sliding the window across input image and extracting the smaller sections of that image, which are being also called as the region proposals. This technique can be used to handle the objects of different sizes by scaling the input image and by rerunning the sliding window.

C. Classification

Each of those aforementioned vicinity proposals (bounding boxes) are problem to class, to discover whether or not the bounding field consists of an item or not. The class is executed both thru a separate class gadget or as an integrated a part of the version producing the vicinity proposals. The class typically yields a chance for every magnificence to be gift withinside the proposed bounding boxes. Typically, a further magnificence relating to heritage is brought to discard bounding boxes with out gadgets in them.

D. Regression

As the generated vicinity proposals aren't anticipated to surround the gadgets as closely as possible, extra regression is generally applied. The regression is similarly because the class both achieved one by one or as a greater included a part of the vicinity proposals. The aim for the regression is to in the long run tighten (or loosen) the proposed bounding boxes, through offset values, to higher incorporate the gadget.

E. Pruning

Each of the produced bounding bins from the version is usually situation to pruning. The maximum broadly used technique to do that is referred to as Non-Maximum Suppression (NMS). NMS includes first putting off all bounding bins wherein the category opportunity output is under a hard and fast threshold, putting off the predictions that the version do now no longer agree with are objects. Generated bounding bins that discuss with the identical object also are removed. This is performed through calculating the overlap among bins, and if the overlap is above a hard and fast threshold, then handiest the bounding field with the highest category actuality is kept.

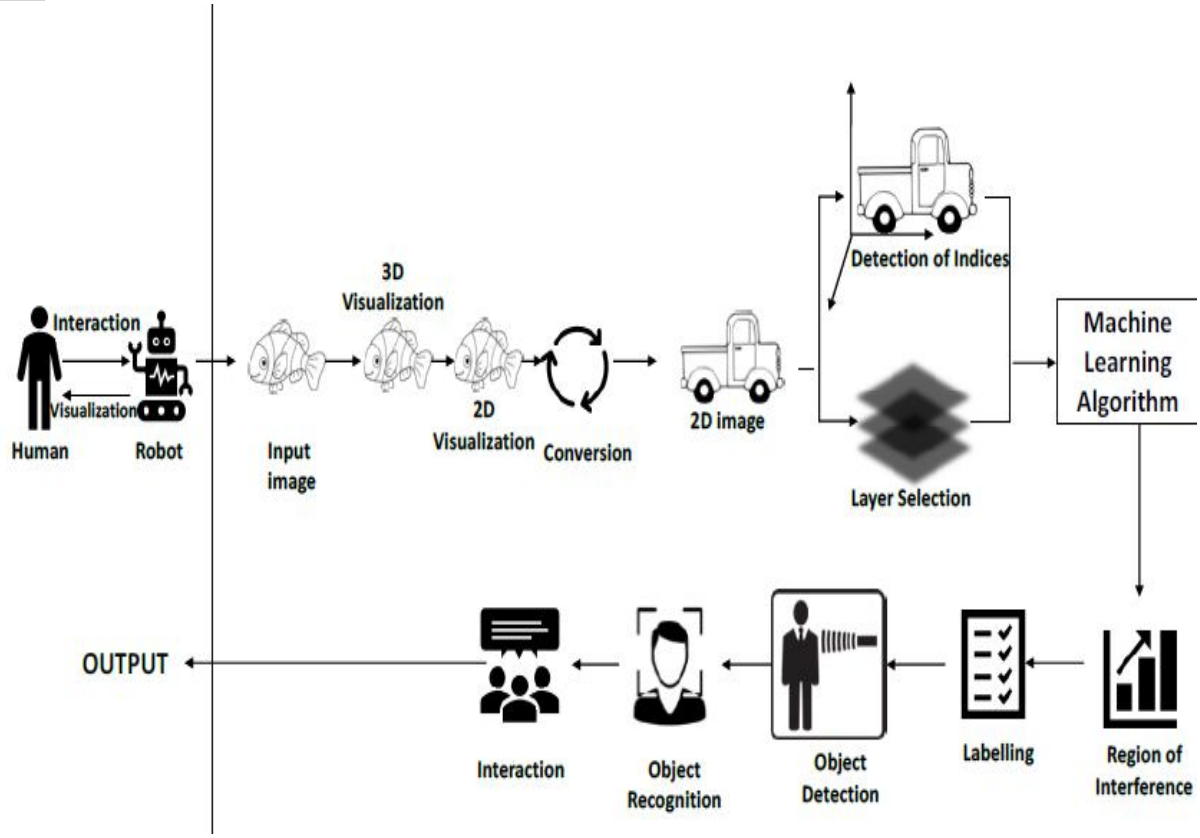


Fig 1 Proposed model working using python3 and OpenCV for 3d image processing

IV. RELATED WORK

Input processing that relies on ML algorithms and projects is the only factor in image recognition based on a particular entity and prevents misidentification[15]. Iterative analysis is performed using a common index based on the discovered components. Timeline analysis is for identification based on previous analysis[12]. The time interval is defined by an algorithm that clicks on an image of an object from different locations at a particular time interval, processes the image in 3D, and displays it as output on a computer screen. You can use plane calculations to find the index and position of an object. In this case, the factors may vary depending on the 3D to 2D type of the captured object[14]. There are two ways to perform this optimization. The plane in which the object resides and its position are the two main constraints based on the object being selected from the plane and producing output[16].

V. RESULT OBTAINED

3D images from the repository are used to evaluate the overall performance of the proposed system. The system trains the ruleset primarily based on a fully populated processing model. Thousands of sample images from numerous instructions are considered a 3D model. Correlation and match validation is performed using checkouts and school datasets containing 500x2 images. The OpenCV software program is used to study these images. 50 images were trained for each matching instance. Input correlation and verification of indices is executed for each instance. Error, processing complexity, processing time and reputation ratio are analysed for estimating the performance of the proposed method. In all tests, the affiliation aspect and variety of objects are verified. Existing schemes like multi-scale convolutional neural network, Context-Assisted 3D, Conventional voxel-primarily based totally occupancy grid, radio-frequency identification, support vector regression and Dynamic Statistical Parametric Mapping are compared in TABLE I [11]. The camera provides a series of objects on which data indices are performed. The object features are defined using labels. Reliable matching is performed on detection of object despite variation in indices. The data indices increase with the increase in the plane and matching as shown in TABLE II [11].

TABLE I as example of patent in [11]

OBJECTS	2 INDICES	4 INDICES	8 INDICES	PLANE CORRELATION
2	0.7	0.83	0.64	0.6
4	0.74	0.84	0.66	0.7
6	0.7	0.78	0.65	0.8
8	0.66	0.75	0.98	0.9
10	0.73	0.73	0.87	1
12	0.67	0.74	0.9	1.1
14	0.7	0.74	0.99	1.2
16	0.73	0.78	1	1.3

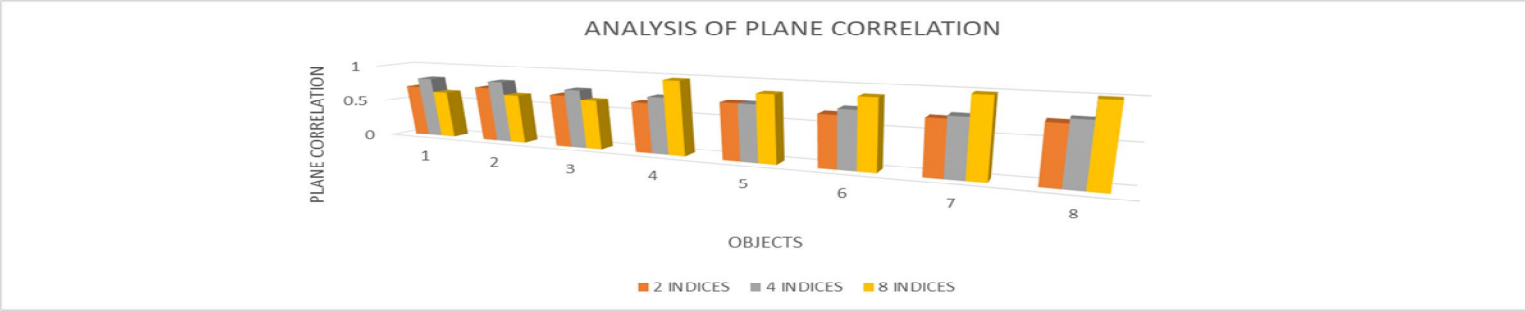


Fig 2 Analysis of plane correlation as an example in patent [11]

TABLE II as an example in patent [11]

XCORRELATION WITH INDICES	SINGLE PLANE	DUAL PLANE	THREE PLANE
0	0.84	0.88	0.91
2	0.84	0.881	0.93
4	0.856	0.882	0.93
6	0.856	0.883	0.93
8	0.856	0.884	0.933
10	0.856	0.887	0.934
12	0.856	0.8872	0.935
14	0.84	0.888	0.939
16	0.858	0.889	0.939
18	0.858	0.9	0.9395
20	0.86	0.889	0.939
22	0.86	0.889	0.94
24	0.863	0.91	0.94
26	0.87	0.91	0.941
28	0.873	0.92	0.945
30	0.871	0.92	0.95
32	0.873	0.93	0.96
34	0.872	0.92	0.97
36	0.873	0.919	0.969
38	0.88	0.921	0.979
40	0.879	0.94	0.98

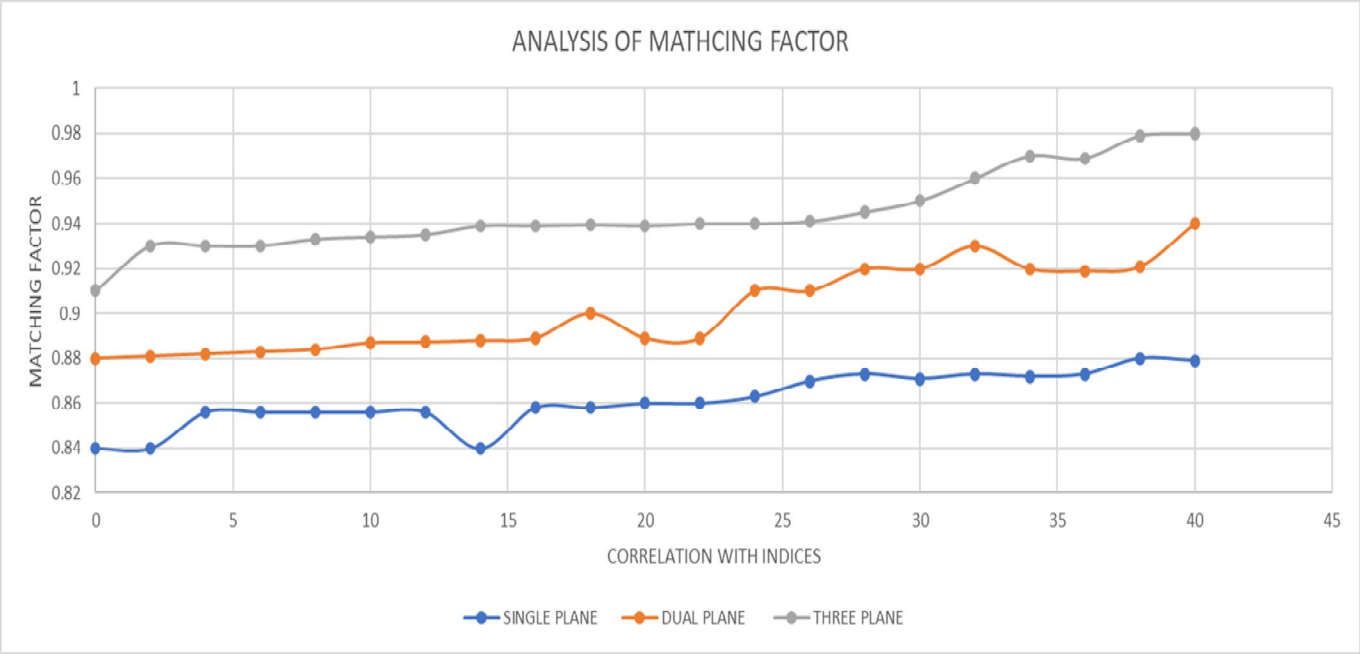


Fig 3 Analysis of MATCHING FACTOR as an example in patent [11]

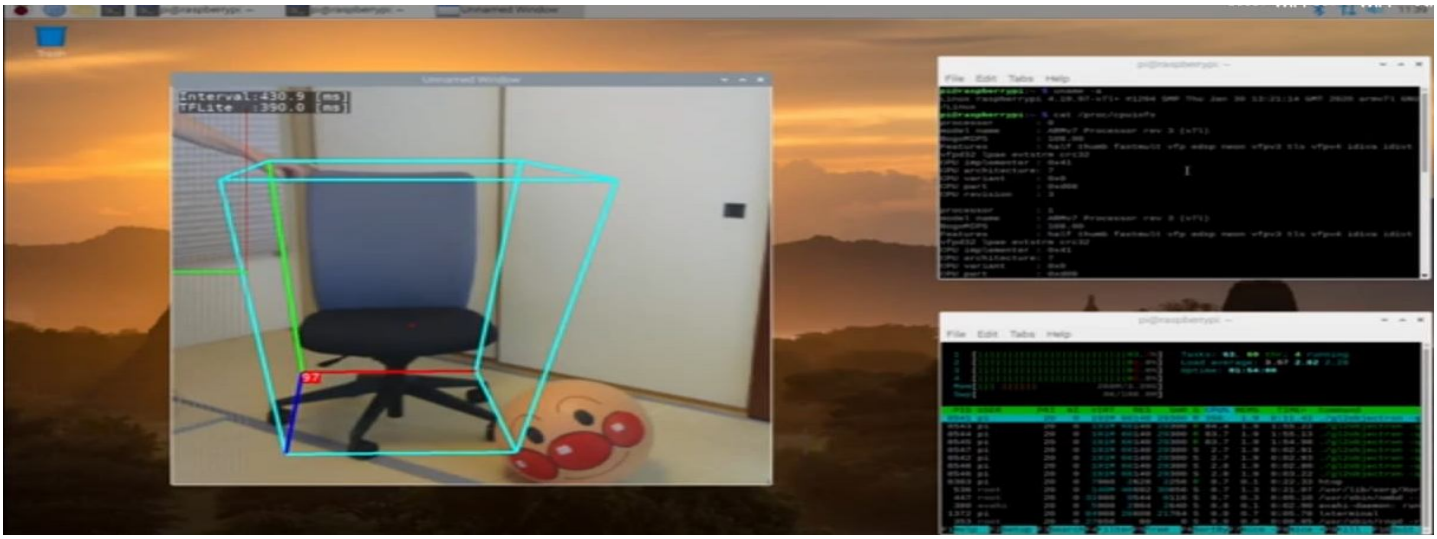


Fig 4 Output obtained on screen

VI. CONCLUSIONS

The robot was finally working but we have experienced some glitches while robot was working. So in order to figure it out we have installed the new firmware in the mini computer due to which we have experienced the minimal glitches in the robot. We have learned a lot of the new things and facts about the microprocessors and microcontrollers and deep and clear knowledge of the ports and the architecture which were being involved in it. After that when we have gone to the part of the designing in the algorithm we have gone up through the each and every step for the design in the algorithm. We have to check each and every part of the algorithm till we have reached the conclusion. Because algorithm design was a crucial task for us. We have to first check whether the firmware can be installed or not. Then during our algorithm design we have done firstly the camera calibration that is we have to test the camera about it's range and as well as it's capability and then we have done the object segmentation. After designing the algorithm, we have seen that our output was coming accurately, there might be 0.01 percent of the human error in the output and after that as you have seen in the above graph we have done our analysis and also seen that analysis in a very care full manner so that it is easy for everyone to understand that analysis. Rest is the output which we have seen in the above topics in this report.

©IJRASET: All Rights are Reserved

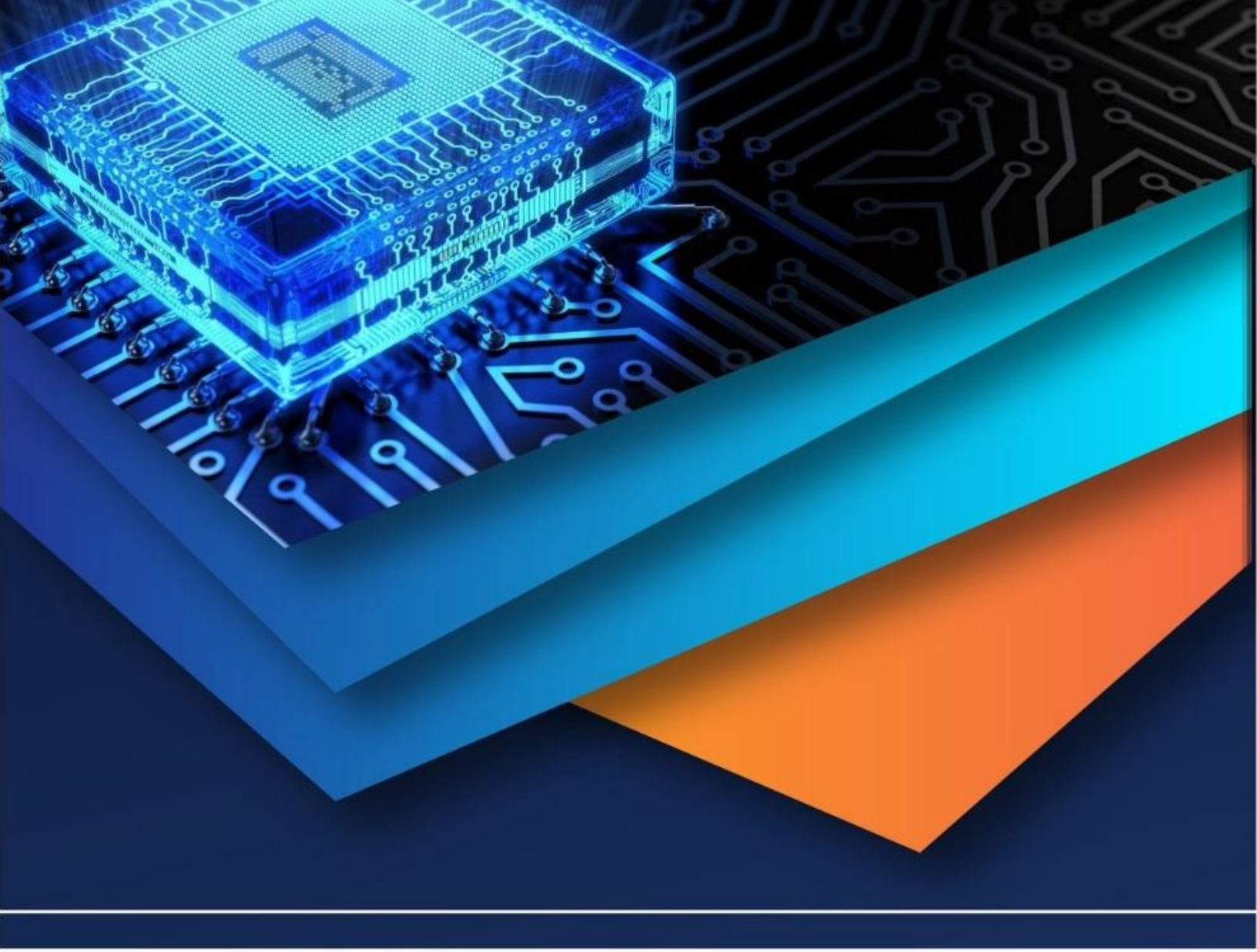
VII. ACKNOWLEDGMENT

First and foremost, We wish to express my profound gratitude to Dr. SUNIL KUMAR MATHUR SIR, for allowing me to carry out my project at MAIT. We find great pleasure to express our unfeigned thanks to our group head Dr. ROHIT RANA SIR, Assistant Professor in ECE, for his invaluable guidance, support and useful suggestions at every stage of this project work. No words can express my deep sense of gratitude to MR. BINAY KUMAR SINGH, without whom this project would not have turned up this way. My heartfelt thanks to him for his immense help and support, useful discussions and valuable recommendations throughout my project work. I wish to thank my respected faculty and my lab mates for their support. Last but not least we thank the almighty for enlightening me with his blessings.

REFERENCES

[1] Christopher Ingraham, The astonishing human potential wasted on commutes 2016. https://www.washingtonpost.com/. Accessed: 2019-01-25.
[2] Global status report on road safety. World Health Organization, 2015.
[3] Katie Pyzyk, Gridlock Woes: Traffic congestion by the numbers, 2018. https://www.smartcitiesdive.com/. Accessed: 2019-01-25.
[4] National Highway Traffic Safety Administration. National Motor Vehicle Crash Causation Survey. 2008.
[5] Shivang Agarwal, Jean Ogier Du Terrail, and Frédéric Jurie. Recent advance, in object detection in the age of deep convolutional neural networks. *CoRR*s, abs/1809.03193, 2018.
[6] Léon Bottou. Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*, pages 177–186. Springer, 2010.
[7] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. nuscenes: A multimodal dataset for autonomous driving. *arXiv preprint arXiv:1903.11027*, 2019.
[8] Xiaozhi Chen, Huimin Ma, Ji Wan, Bo Li, and Tian Xia. Multi-view 3d object detection network for autonomous driving. *CoRR*, abs/1611.07759, 2016.
[9] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In *international Conference on computer vision & Pattern Recognition (CVPR'05)*, volume 1, pages 886–893. IEEE Computer Society, 2005.
[10] John Duchi, Elad Hazan, and Yoram Singer. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning, Research*, 12(Jul):2121–2159, 2011.
[11] Akay Sungheeta, Rajesh Sharma for 3D image processing using machine learning based input processing. Journal of Innovative Image Processing(2021)
[12] Vincent Dumoulin and Francesco Visin. A guide to convolution arithmetic for deep learning. *arXiv preprint arXiv:1603.07285*, 2016.
[13] Mark Everingham, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman. The pascal visual object classes (voc) challenge. *International journal of computer vision*, 88(2):303–338, 2010.
[14] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2012.
[15] Ross Girshick. Fast r-cnn. In *The IEEE International Conference on Computer Vision (ICCV)*, December 2015.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)