iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** IV **Month of publication:** April 2024

DOI: https://doi.org/10.22214/ijraset.2024.60228

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 12 Issue IV Apr 2024- Available at www.ijraset.com

Live Object Recognition Using YOLO

Prathamesh Sonawane[1], Rupa Gudur[2], Vedant Gaikwad[3], Harshad Jadhav[4]
[1,2,3,4]*Undergraduate Student (IT), Vasantdada Patil Pratishthan's College of Engineering and Visual Arts, Mumbai, India*

Abstract: ***Live object recognition refers to the real-time process of identifying and categorizing objects within a given visual input, such as images. This technology utilizes computer vision techniques and advanced algorithms to detect objects, determine their dimension, area and weight and often classify them into defined categories. Our system proposesR-CNN and YOLO to determine the dimensions of the objects in real time. YOLO takes a different approach by treating object detection as a single regression problem. A single neural network is trained to directly predict bounding boxes and class probabilities for multiple objects in an image. The input image is divided into a grid, and each grid cell is responsible for predicting the objects whose center fall within that cell. Live object recognition finds applications in various fields, including autonomous vehicles, surveillance systems, robotics,augmented reality, and more. By providing instantaneous and accurate insights into the surrounding environment, this technology contributes to enhanced decision-making,interaction, and automation across numerous domains. The objects which can be recognized are solid objects which we use daily such as electronic items, stationery items, culinary items and many more. Our system "Live Object Recognition using YOLO" is aimed to detect and determine the objects and dimensions in real time.***

Keywords: Live Object Recognition, YOLO, Accuracy, Dimension measure, CNN, Deep Learning, Convolution Neural Network, Object Detection, Machine Learning.

I. INTRODUCTION

The capacity for comprehension and interaction with the visual world has become increasingly important in the current era of rapid technological innovation. "Live Object Dimension Recognition" is one of this domain's outstanding accomplishments. Object recognition is the process of identifying items in pictures and videos. The autonomous vehicles can recognize and classify items in real time thanks to this computer vision approach. An autonomous vehicle is a car that can sense and respond to its surroundings in order to navigate on its own without assistance from a human. Since object detection and recognition enable the vehicle to detect impediments and determine its future trajectory, they are regarded as among the most crucial duties. This breakthrough is extremely significant since it has applications in many other fields, including manufacturing, logistics, augmented reality, and more.

Through the use of cutting-edge object recognition algorithms such as YOLO (You Only Look Once) or related technologies, this combination aims to give machines the ability to recognize, classify, and comprehend the spatial properties of things in real time. But this is not an easy task; in order to deal with perspective distortions and other real-world difficulties, sophisticated computer vision algorithms must be incorporated. Furthermore, this endeavor's development of user-friendly interfaces, calibration capabilities, and smooth interaction with current systems are crucial components that,in the end, promise to completely transform how we communicate with and comprehend our physical environment.

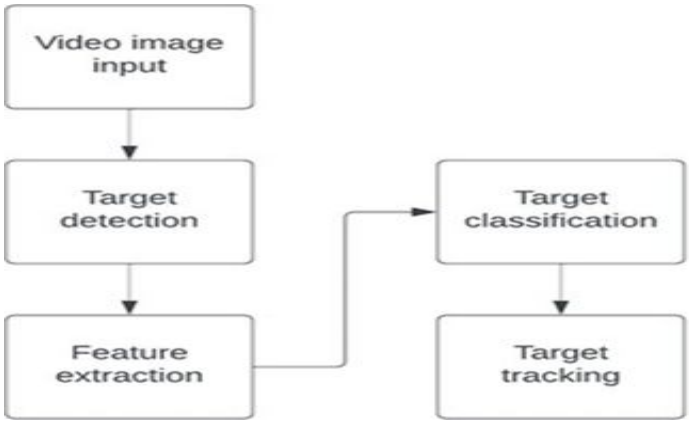


Fig. 1. A general object detection system

©IJRASET: All Rights are Reserved

II. LITERATURE REVIEW

The method described for object tracking in video sequences utilizes morphological characteristics for intelligent tracking. It involves three key procedures: initial background estimation, distinctive geometry analysis, and object registration. By subtracting subsequent frames, moving items are identified, and morphological traits are used for object registration. The process is repeated for each frame, and optical flow is computed to track objects efficiently. While computationally efficient compared to other methods, it still faces challenges such as high processing overhead and complexity. However, it excels in monitoring multiple objects and resolving object merging issues during tracking [1].Feature Pyramid Networks (FPN) address scale invariance in object identification systems by constructing feature pyramids from image pyramids. Unlike previous methods, FPN combines high-resolution and semantically weak features with low-resolution and semantically strong ones through top-down and bottom-up pathways and lateral connections. This allows for rich semantics extraction at all levels and enables end-to-end training without sacrificing speed or memory efficiency. FPN is versatile and can be applied to various computer vision tasks like instance segmentation and region proposal creation [2].The process of object detection involves eliminating irrelevant noise from
input data, such as images or videos, to train models effectively. Convolutional neural networks (CNNs) serve as the backbone of object detection models, handling each significant component intricately. After collecting and extracting data, the model is trained to recognize objects, refined for precision, and deployed for real-world applications. Efforts are underway to develop object detection models compatible with smartphones and small devices, aiming to broaden accessibility beyond traditional large computers and Macs, primarily focusing on the Android system[3].].The process of image analysis and categorization begins with noise elimination from the input image. This technique is similarly employed in video processing for real-time object detection, where resolution influences tracking accuracy. Object detection entails both classification and localization, resulting in bounding boxes around objects in the image and identifying them. Convolutional Neural Networks (CNNs) are integral to this process, as they excel in pattern recognition and directly learn from image data to facilitate object detection. The Caffe framework, initially developed at UC Berkeley, offers a comprehensive toolkit for training, testing, fine-tuning, and deploying CNN models for object detection [4]

The paper introduces a method for object identification, dimension measurement, and recognition using video or image input from a computer's webcam or external camera. The system, developed with Python and OpenCV libraries, comprises three modules: setting up the camera and environment, detecting and identifying objects using techniques like Gaussian blurring and thresholding, and determining object dimensions using contour data and mathematical calculations. The proposed approach demonstrates high accuracy in object detection and dimension measurement, often achieving above 95% accuracy [5]. The paper introduces a novel multi-scaled deformable convolutional object detection network based on YOLO v3. It combines techniques from FPN and deformable convolutional networks to improve object detection accuracy. By modifying convolution operations and incorporating multi-scaled feature fusion, the network enhances its capability to detect objects of various sizes and shapes while maintaining computational efficiency. Experimental results demonstrate significant improvements in both speed and accuracy compared to alternative object detection methods such as R-CNN and SSD. Future research will focus on further optimizing the network structure and exploring its application in real-time object detectionin videos[6].

III. PROPOSED METHODOLOGY

In this paper, we propose an approach that uses the video or image of the surroundings captured by the computer's webcam or external camera as an input to identify items,measure their dimensions, and identify them.With a stand, we are able to keep the camera at a specific distance while also adjusting the camera's height, width, and depth. With the use of this system, we are able to identify many objects at once, provide their dimensional measurements, and obtain other information such as the object's area of occupancy. The system is created using a python program which makes use of python libraries for queue, math, numpy, and computer vision (opencv-cv2). This system is composed of three modules(a,b and c) that act in accordance with various system-related tasks and are discussed below sequentially.

Module a - the first module is used for obtaining the video frames input, in this module we configure our environment by setting the camera width, height and also we set the frame rate of the video input.

Module b - the second module is used to detect object by capturing the objects boundaries. There are multiple functions performed in this module. Initially the input frame to gray scale to get better understanding of the details in input. The function's final parameter, which is obtained using OpenCV, supports multiple types of thresholding.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

Module c - the third module is used to find the object dimensions by using the contour data we find the various lengths of the object from the math library we use hypot method which is used to calculate euclidean distance, there by allowing us to calculate length,breadth, and also also area of the object.

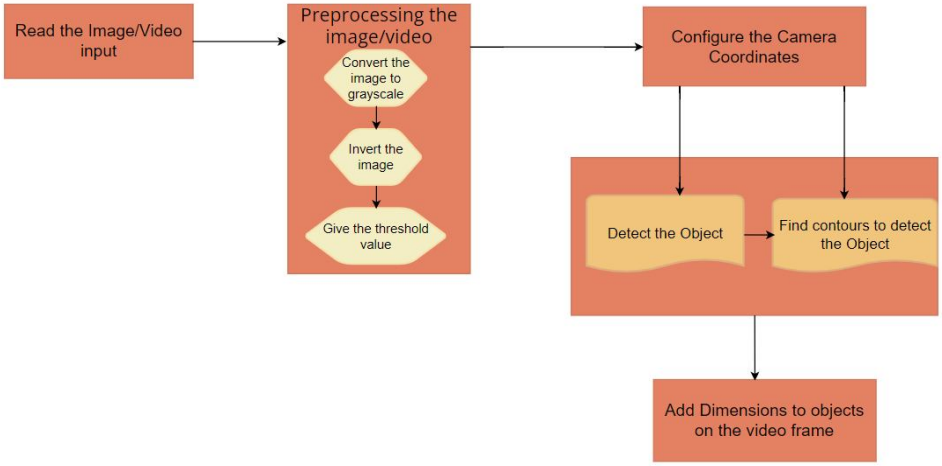


Fig 2.PROPOSED ARCHITECTURE

IV. RESULTS



FIG 3. SYSTEM USER INTERFACE

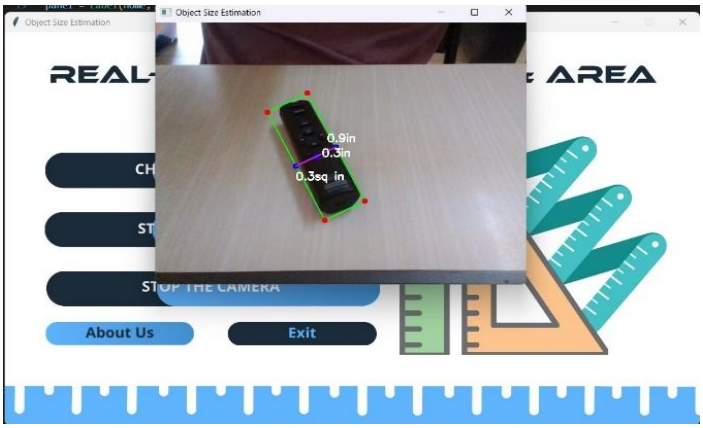


FIG 4. OUTPUT 1

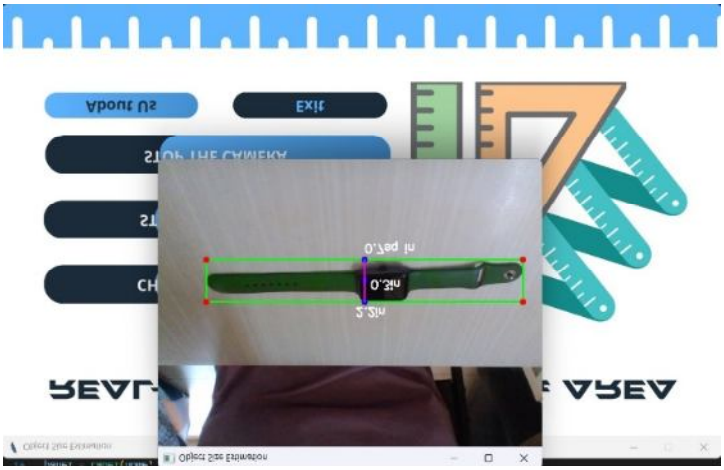


FIG 5.OUTPUT 2



FIG 6. OUTPUT 3

TABLE 1.COMPARISON TABLE

object	Actual length (inch)	Actual breadth (inch)	Actual area (inch)	Observed length (inch)	observed breadth (inch)	observed area (inch)	Accuracy (%)
mobile	6.10	2.83	493.37	6.04	2.77	424.20	96.54
pen	4.72	0.51	61.41	4.69	0.52	61.38	99.95
book	10.82	6.45	1775.59	10.86	6.43	1776.67	99.94
box	11.88	12.00	3626.78	11.59	12.20	3595.45	99.14
wallet	3.58	2.87	261.53	3.65	2.94	272.70	95.72
Pencil box	5.51	1.77	248.03	5.55	1.77	249.98	99.21

V. CONCLUSION

The importance of computer vision in contemporary technological applications—particularly in the processing, analysis, and comprehension of digital videos—seems to be emphasized in this work. Your suggested study shows remarkable accuracy in both item detectionand size measurement tasks by utilizing computer vision algorithms.The focus on surpassing 95% accuracy in dimension measuring is significant since it highlights the dependability and accuracy of your methodology. This degree of precision is essential for many real-world uses, such figuring out how much space an object takes up.

It can be found in daily life, in industrial settings where precise measurements are necessary for manufacturing processes, in e-commerce platforms where accurate sizing information is critical for online shoppers, or even in the development of AI-driven cars where object detection and measurement help with navigation and safety systems, there are a lot of possible uses for what you do. Accurate measurement of dimensions creates opportunities for increasing productivity, optimizing user experience, and developing technology capabilities in several domains. Your suggested work has the potential to significantly impact numerous sectors and areas by offering a reliable solution for computer vision-based object detection and dimension measurement.

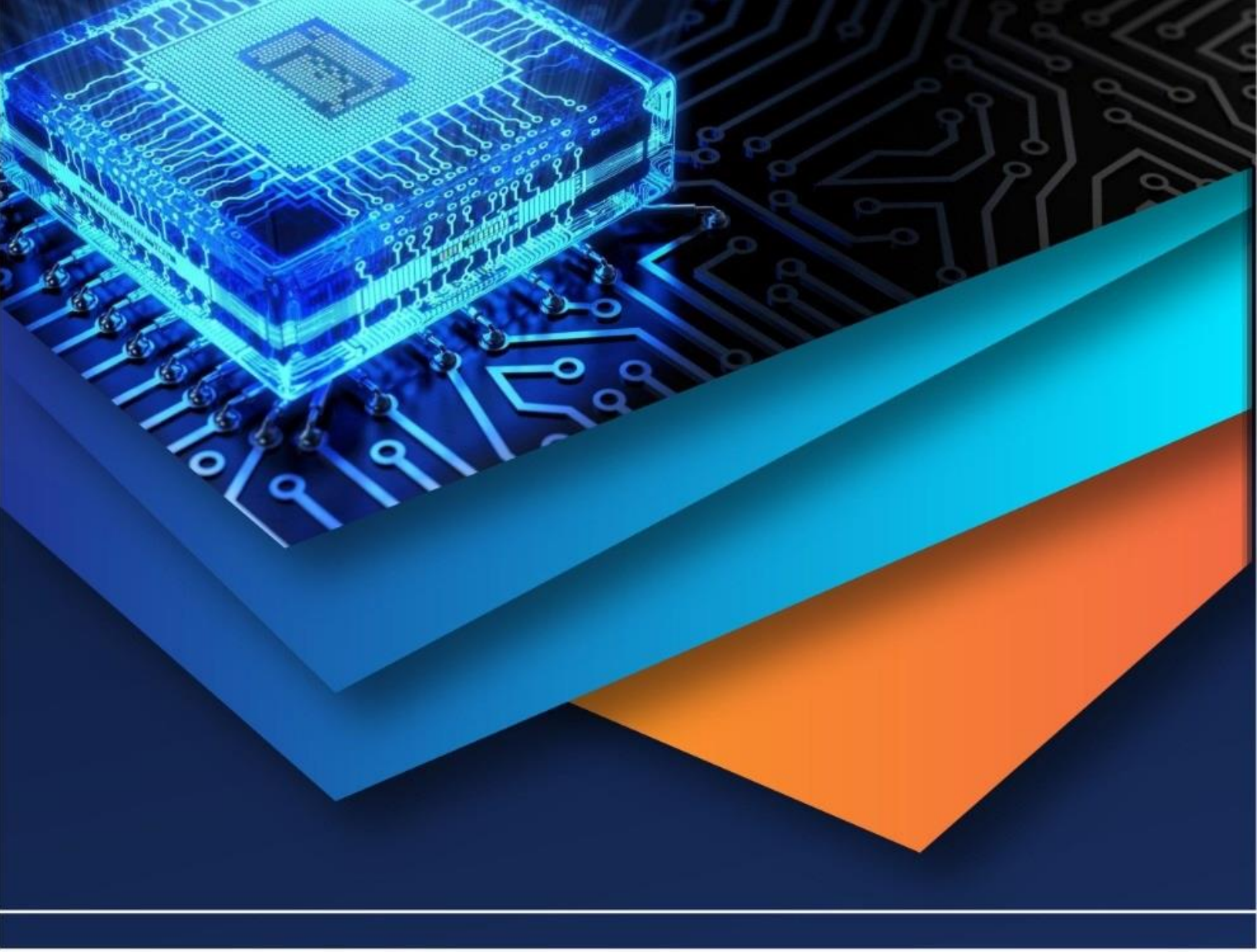
VI. FUTURE WORK

Object recognition will become more accurate and efficient as AI models grow more complex and capable of tackling difficult jobs. Developing object recognition is crucial to the advancement of autonomous vehicles. In order to improve the safety and dependability of autonomous vehicles, object recognition technology will be increasingly integrated in the future. The creation of smart cities will heavily rely on object recognition. It can be applied to trash management, traffic control, security, and other municipal service optimization, all of which will enhance the standard of living in cities.In order to ensure responsible and secure deployment, continued advancements in privacy-preserving methods and ethical considerations will also be crucial in determining the direction of this subject.It is anticipated that object recognition technology will be more deeply integrated into the creation of smart cities in the future. Urban applications for object recognition are numerous and include trash management, traffic control, security monitoring, and service optimization. The potential for this thorough integration of AI-powered item detection might completely transform urban living. In the end, it can improve urban dwellers' overall quality of life by resulting in more efficient city planning, safer, better traffic flow, and lower energy use.

VII. REFERENCES

[1] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu, “Object Detection with Deep Learning”, IEEE Transactions on Neural Networks and Learning Systems (Volume: 30, Issue: 11, November 2020), pp. 3212- 3232, doi: 10.1109/TNNLS.2018.2876865
[2] Hammad Naeem, Jawad Ahmad_and Muhammad Tayyab, “Real-Time Object Detection and Tracking”,HITEC University Taxila Cantt, Pakistan,Conference: 202016th International Multi Topic Conference (INMIC), doi:10.1109/INMIC.2020.6731341
[3] Anjali Nema, Anshul Khurana "Real Time Object Identification Using Neural Network with Caffe Model." International Journal of Computer Sciences and Engineering 7.5 (2021), p.175-182.
[4] Abhishek Kamble, Abhijit G. Kavathankar, Prathamesh P. Manjrekar, Priyanka Bandagale, “Object Detecting Artificial Intelligence (ODAI)”;International Conference on Smart Data Intelligence, DOI:10.2139/ssrn.3851990
[5] Madhavi Karanam1, Varun Kumar Kamani1,*, Vikas Kuchana1, Gopal Krishna Reddy Koppula1 , and Gautham Gongada1:Object and it’s dimension detection in real time;E3S Web of Conferences 391, 01016 (2023) ICMEDICMPC2023;https://doi.org/10.1051/e3sconf/202339101016
[6] Danyang Cao1,2* , Zhixin Chen1 and Lei Gao1 : An improved object detection algorithm based on multi-scaled and deformable convolutional neural networks Cao et al.Hum. Cent. Comput. Inf. Sci. (2020) 10:14;https://doi.org/10.1186/s13673-020-00219-9
[7] K. Shreyamsh, UAV: Application of object detection and tracking techniques for unmanned aerial vehicles, Texas A & M University, (2015).
[8] M. Pietikinen, T. Ojala, T. Maenpaa, Multi-resolution gray scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume 24, Issue 7, 971 - 987,(2002).
[9] R. Girshick, Fast R-CNN, In Proceedings of the IEEE international conference on computer vision pp. 1440-1448,(2015)

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



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)