As per UGC guidelines an electronic bar code is provided to secure your paper

International Journal for Modern Trends in Science and Technology
Volume 11, Issue 07, pages 42-47
ISSN: 2455-3778 online
Available online at: http://www.ijmtst.com/vol11issue07.html
DOI: https://doi.org/10.5281/zenodo.15770262

Intelligent Object Detection System

Dr.S.Mallikharjuna Rao, N.V.Krishna Prasad, E.Sai Sudheer, G.Prem Tharun, Dyaswanth

Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering & Technology, Vijayawada, India.

To Cite this Article
Dr.S.Mallikharjuna Rao, N.V.Krishna Prasad, E.Sai Sudheer, G.Prem Tharun & Dyaswanth (2025). Intelligent Object Detection System. International Journal for Modern Trends in Science and Technology, 11(07), 42-47. https://doi.org/10.5281/zenodo.15770262

Article Info
Received: 06 June 2025; Accepted: 26 June 2025.; Published: 29 June 2025.

KEYWORDS

YOLO, SSD,Facial Recognition , OpenCV , Convolutional Neural Network,Aspect Ratio

ABSTRACT

As technology improved, object detection, which is connected to video and image analysis, caught researchers' interest. Earlier object recognition techniques are based on hand-crafted features and imprecise architectures and trainable algorithms. One of the main issues with many object detection systems is that they rely on other computer vision methods to support their deep learning-based methodology, which leads to slow and subpar performance. In this article, we present an end-to-end solution to the object detection problem using a deep learning based method. The single shot detector (SSD) technique is the quickest method for object detection from an image using a single layer of a convolution network. Our research's primary goal is to enhance accuracy of SSD method

1. INTRODUCTION

Image classification, which is defined as figuring out the class of the image, was one of the essential issues. The Challenge of image localization, when one item is present in the picture and the system must predictits class and position within the image, is rather challenging (a bounding box around the object). The fact that objects discovery includes both identification and localisation makes it a more challenging challenge. In this instance,an image will be used as the system's input, and the output will be a bounding box that corresponds to every object in the image and specifies the type of object in each box. We built a solution that uses less processing power than the existing techniques while operating at enhanced FPS and fast object detection [1,2]. The SSD mobile net method is used by our object discovery model to identify and celebrate the item in the image. The algorithm in our model analyses appearance existing in an image to pinpoint a specific object. Object detection is a computer vision technique that helps identify and locate objects in images and movies. With this form of identifying and localizing, detection of objects may be used to count the items in a scenario, locate and identify them precisely, and name them.Have you ever noticed how adeptly

Face book can recognise your pals in your photos? In order to tag friends in photographs on Face book, you used to have to click on the friend's profile and enter

their names [3-5]. These days, Face book automatically tags everyone in your photos as soon as you upload them.This method is known as face-recognition.Facebook's algorithms may recognise your friends' faces after just a few times of being tagged. Face book has a facial detection accurac of 98%, which is comparable to human performance. Faces in picture and video streams on social media and mobile devices may be used to recognise people [6-8].

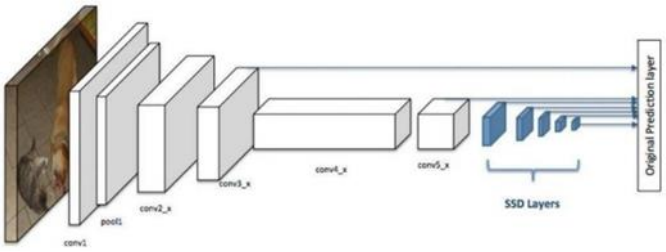


Fig 1 .1

1.2 Objective of the Project:

The objective of this project is to develop a lightweight object detection algorithm for unmanned surface vehicles (USVs) that balances high detection speed and accuracy in complex maritime environments. The proposed algorithm, based on YOLOv5, is designed to enhance the USVs' ability to perceive their surroundings, providing critical data for tasks like path planning and collision prevention during reconnaissance missions. The improvements, including the integration of the Ghost module and Transformer, focus on efficient feature extraction, reduced model size, and increased detection precision and speed, making it robust even under challenging marine conditions like sea fog

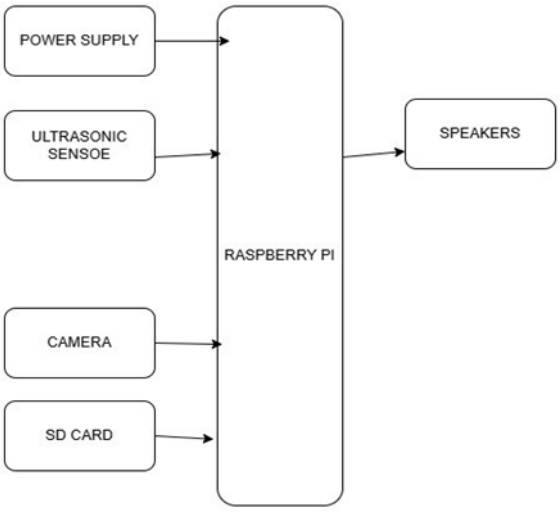


FIG:2

1.3 Overview of the Paper:

The YOLO Algorithm: YOLO (You Only Look Once) is a groundbreaking real-time object detection framework. Unlike traditional methods that run object detection algorithms on multiple regions of an image, YOLO processes the entire image in a single pass, making it remarkably fast.

SSD Methodology: Single Shot Detector (SSD) enhances real-time object detection by eliminating the need for region proposal networks, resulting in higher speeds. SSD achieves accuracy comparable to Faster R-

CNN while being faster, especially on low-resolution images.

Face Recognition Accuracy: Facebook's facial recognition technology boasts an impressive accuracy rate of 98%, rivaling human-level performance. This showcases the power of deep learning algorithms applied to real-world applications.

Applications in Attendance Systems: Object detection and facial recognition have made their way into innovative applications like automated attendance systems, particularly in the education sector, streamlining processes and reducing human error.

Historical Leap in Image Recognition: The journey of image recognition technology began in the 1980s and has since evolved to power applications in security, social media, healthcare, and even industrial automation.

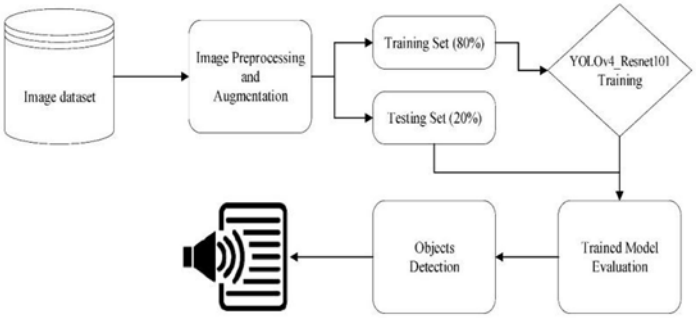


FIG:3

2. RELATED WORK

A library or programming package called OpenCV was created primarily to assist programmers in learningabout computer vision. OpenCV is an abbreviation for free computer vision software, and the package was developed by Intel Corporation and made accessible to the public between 1999 and 2000. (Library).

The most popular, well-known,and well documented library for computer vision. As the programme is open-source,there is no licencing required to use it. As is probably previously known, the bulk of machine learning algorithms require numerical orquantitative

inputs. A libraryor programming package called OpenCV was created primarily to assist programmers in learning about computer vision. OpenCV is an abbreviation for free computer vision software, and the package was developed by Intel Corporation and made accessible to the public between 1999 and 2000. (Library). The most popular, well-known, and well documented libraryfor computer vision. As the programme is open-source, there is no licencing required to use it. As is probably previously known, the bulk of machine learning algorithms require numerical or quantitative inputs. Despite the fact that OpenCV makes it possible for us to apply machine learning techniques to pictures , the raw images are usually need to be processed in order to transform them into features(columns of data).They benefit our machine learning algorithms and are utilised by them

NumPy is a Python package . The name "Numerical Python " 0refers to a collection of procedures for working with multidimensional array objects and arrays.Pandas are a rapid, powerful, adaptable, and user-friendly open-source programme for data analysis and manipulation. Using the Python programming language as a foundation

The Python Imaging Library gives the Python interpreter the capacity to process pictures . This library provides a broad variety of file format compatibility, a helpful internal representation, and rather powerful image processing tools.

3. PROPOSED APPROACH

The paper introduces an upgraded methodology for object detection that utilizes an enhanced Single Shot Detector (SSD) algorithm. The authors aim to improve the accuracy and speed of real-time object detection by integrating depth-wise and spatial separable convolutions within the convolutional layers of the SSD. These enhancements address challenges like small object detection and optimize the feature extraction process with a resolution multiplier. The architecture combines a multi-layer convolutional neural network with carefully designed filters to apply the most suitable aspect ratios for object detection. By eliminating the region proposal network commonly used in Faster R-CNN, the approach accelerates the detection process while maintaining comparable accuracy. The SSD's enhancements, such as default boxes and multi-scale functionality, ensure high accuracy even for low-resolution images. Additionally, the paper highlights the algorithm's ability to generate bounding boxes, segment the image, and iteratively refine predictions through back-propagation, enabling efficient identification of multiple objects. This methodology demonstrates significant improvements in both precision and speed, making it ideal for real-time applications..

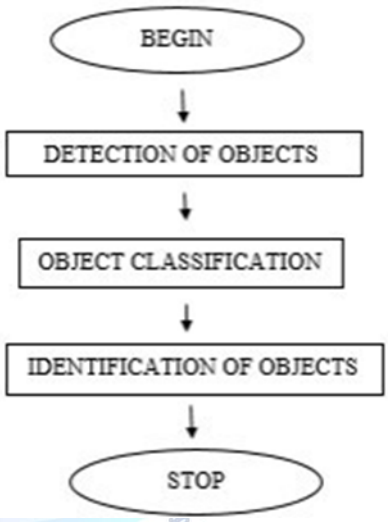


FIG:4

4. SHIP DETECTION BASED ON YOLO

Since the YOLO algorithm was published, it has been widely studied because of its good computational efficiency and detection accuracy. Lee et al. (2018) applied the YOLOv2 algorithm to ship detection and classification. In comparison with other machine learning algorithms, their model has better robustness and scalability. Li and Qiao (2021) proposed a shipdetection and tracking algorithm based on YOLO ,YOLOv5 has high performance in terms of detection speed and accuracy. According to the depth and width of the network, it is divided into four versions: YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv5x. The basic network of the four versions is similar.

The structure of YOLOv5 is mainly composed of the input, backbone, Neck, and Prediction. At the input, we perform data augmentation operations, such as Mixup

and Mosaic, which can enrich the ship dataset and improve the detection efficiency of small objects. Feature maps of different scales are extracted at the backbone network. The FPN and path aggregation network (PANet) at the Neck strengthen the feature fusion ability. The FPN transfers high-level semantic features in a top-down manner, and the PANet transfers low-level strong localization features in a bottom-up manner after the FPN. The final output is the prediction of the network, and the prediction uses the nonmaximum suppression (NMS) algorithm to filter the object boxes. Then, we make predictions on the image features, generate bounding boxes and predict classes.

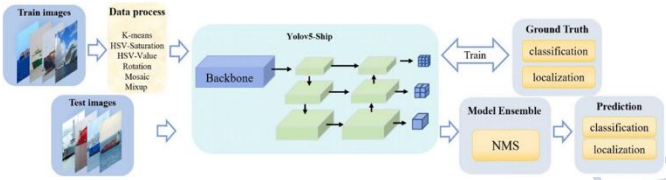


FIG:5

5. DATASET DESCRIPTION

Among the 300 photos in our collection are depictions of a boat, a bicycle, a cow, a human, a bottle, etc. Our

Technique is examined using a real-time web camera that records the Items . Following pre-processing , the figure below displays a few examples of pictures.

Telegram

The detected objects from the object detection system were shared with users through a Telegram bot, utilizing Telegram's API for automated and efficient communication. After the SSD algorithm processes images or video streams and identifies objects, the results—including object labels, bounding box coordinates, and confidence scores—are packaged into a structured message format. The Telegram bot then transmits these messages to designated recipients or groups, allowing users to receive real-time updates on detected objects directly on their devices. This integration enhances the practical application of the object detection system by offering seamless accessibility and instant notifications, making it a convenient solution for real-time monitoring and communication. Let me know if you'd like more details on how the bot operates or its implementation process!s.

PYTHON:

Python is a general-purpose, high-level, dynamically typed, and interpreted programming language that supports object-oriented programming (OOP) for application development. Known for its simplicity and readability, Python provides a rich set of high-level data structures, making it easy to learn yet powerful enough for complex software development. Its dynamic typing and interpreted nature make it an ideal language for scripting and rapid application development.

Python supports multiple programming paradigms, including object-oriented, imperative, functional, and procedural programming styles. Unlike specialized languages, Python is multipurpose, allowing developers to build web applications, enterprise solutions, 3D CAD software, AI, and automation tools. Its dynamically typed feature eliminates the need to specify variable types explicitly, enabling faster development with statements like a = 10 for assigning integer values.

One of Python's major advantages is its fast development and debugging process. Since it does not require compilation, developers can edit, test, and debug code efficiently. This makes Python particularly popular for prototyping, data science, machine learning, and automation. With a vast ecosystem of libraries and frameworks, Python continues to be a preferred language for beginners and experienced developers alike.

FIG:7

Result

The following steps are involved in our proposed system.

Step-1 .It uses the use's camera to capture the picture as input.

Step-2. It transforms the picture.

Step-3.It takes all the required features out of the picture.

Step4.To recognise more objects in the image , it divides it into smaller bits.

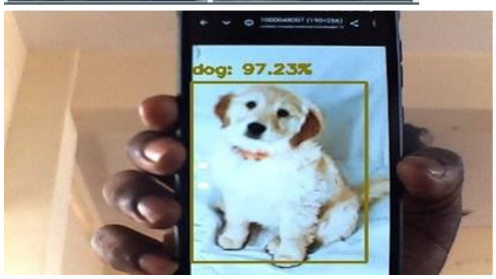
Step-5. Try to categorise and identify the objects after segmenting them.

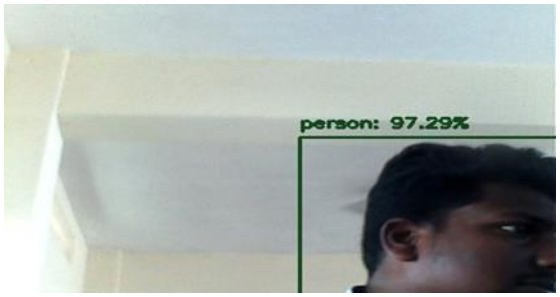
Step-6.Then the process of finding things in the image be gins.

Step- 7.It shows the output to the use

OUTPUT:

Telegram Output:





5. CONCLUSION AND FUTURE SCOPE

This study creates a deep learning-based item recognizer for identifying objects in images. The study uses an upgraded SSD technique and a multilayer convolution network to recognise things quickly and accurately. Both still images and moving images are handled well byour technology. More than 80% of the predictions made by the proposed model are correct. After removing feature data from the image, convolution neural networks employ feature mapping to get the class label. The major objective of our solution is to improve SSD's object detection process by selecting default boxes with the best feasible aspect ratios.

Object identification technology has the ability to relieve people from regular jobs that robots can carry outmore quickly and efficiently, much like the first Industrial Revolution did. This technologyis now being tested.

Conflict of interest statement

Authors declare that they do not have any conflict of interest.

REFERENCES

[1] Bai, Y., Lei, S., and Liu, L. (2021). "The ship object detection based on Sea-Sky Line," in 2021 6th International Conference on Automation, Control and Robotics Engineering(CACRE), IEEE.456–460. doi: 10.1109/CACRE52464.2021.9501

[2] Bochkovskiy, A., Wang, C. Y., and Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934. doi: 10.48550/ arXiv.2004.10934

[3] Chen, X., Ling, J., Wang, S., Yang, Y., Luo, L., and Yan, Y. (2021). Ship detection from coastal surveillance videos via an ensemble canny-gaussian- morphology framework. J. Navigation 74, 1252–1266. doi: 10.1017/S037346332100

[4] Chen, T., Wang, N., Wang, R., Zhao, H., and Zhang, G. (2021). One-stage CNN detector-based benthonic organisms detection with limited training dataset. Neural Networks 144, 247–259. doi: 10.1016/j.neunet.2021.08.014

[5] Dzvonkovskaya, A., and Rohling, H. (2010). "Cargo ship RCS estimation base on HF radar measurements," in 11-th International Radar Symposium. IEEE. 1–4.

[6] Electronic Quality Shipping Information System (2020) The world merchant fleet in 2020. Available at: https://www.equasis.org (Accessed May 17, 2000).

[7] Girshick, R. (2015). "Fast r-cnn," in Proceedings of the IEEE international conference on computer vision. IEEE. 1440–1448.

[8] Gupta, V., Gupta, M., and Singla, P. (2021). Ship detection from highly cluttered images using convolutional neural network. Wireless Pers. Commun. 121, 287–305. doi: 10.1007/s11277-021-08635-5

[9] Han, K., Wang, Y., Tian, Q., Guo, J., and Xu, C. (2020). Ghostnet: More features from cheap operations. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. IEEE. 1580–1589.

[10] Jain, A. (2021) Game of deep learning: Ship datasets. Available at: https://www.kaggle.com/datasets/arpitjain007/game-of-deep-learning-ship- datasets (AccessedJune 24, 2021).

[11] Jie, Y., Leonidas, L., Mumtaz, F., and Ali, M. (2021). Ship detection and tracking in inland waterways using improved YOLOv3 and deep SORT. Symmetry 13(2), 308. doi: 10.3390/sym130203

[12] Lee, S.-J., Roh, M.-I., Lee, H.-W., Ha, J.-S., and Woo, I.-G. (2018). "Image-based ship detection and classification for unmanned surface vehicle using real-time object detection neural networks," in The 28th International Ocean and Polar Engineering Conference (OnePetro). 726–730.

[13] Li, G., and Qiao, Y. (2021). A ship object detection and tracking algorithm based on graph matching. In Journal of Physics: Conference Series (IOP Publishing), vol. 1873 (1), 012056. doi: 10.1088/1742-6596/1873/1/012056

[14] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., et al. (2016). "Ssd: Single shot multibox detector," in European Conference on computer vision (Springer), 21–37. doi: 10.1007/978-3-319-46448-0_2

[15] Liu, C., and Li, J. (2021). Self-correction ship tracking and counting with variable time window based on YOLOv3. Complexity 2021, 1–9. doi: 10.1155/2021/7428927

[16] Liu, T., Pang, B., Zhang, L., Yang, W., and Sun, X. (2021). Sea Surface object detection algorithm based on YOLO v4 fused with reverse depthwise separable convolution (RDSC) for USV. J. Mar. Sci. Eng. 9, 753. doi: 10.3390/jmse9070

[17] Liu, Z., Zhou, F., Bai, X., and Yu, X. (2013). Automatic detection of ship object and motion direction in visual images. Int. J. Electron. 100, 94–111. doi: 10.1080/00207217.2012.687188