iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** V **Month of publication:** May 2023

DOI: https://doi.org/10.22214/ijraset.2023.50919

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 11 Issue V May 2023- Available at www.ijraset.com

Object Detection in Underwater Using Deep Learning Techniques

Mr. R. Trinadh[1], M. Chaitanya Deepika[2], M. Manojna[3], K. Sindhu Lavanya[4], H. Pranay Deep[5], K. Ramya Sri[6]
[1]*Assistant professor,* [2, 3, 4, 5, 6]*UG Schlors, Department of Electronics and Communication Engineering, Sir C. R. Reddy College of Engineering, Eluru-534166, India*

Abstract: ***Based on the architecture of convolutional neural networks, a model is suggested. The model was created using underwater photography. YOLO is used in this method to locate items underwater. An autonomous underwater item-detecting system is necessary to reduce the cost of underwater inspection. An autonomous underwater item detection system is necessary to reduce the cost of underwater inspection. The main goal of this project is to create a model that can identify and recognize objects. This can be done using deep learning techniques. Object detection has two parts. One is object classification and the other is object localization. Classifying objects into predefined classes classifies objects by location under object localization. Our goal is to test the input images after the system is trained by matching the objects in the training dataset to the training dataset. I suggested using the YOLO model to find objects in images. YOLO is a method that enables real-time object recognition using neural networks. You Only Look Once is known by the acronym YOLO. The precision and speed of this algorithm are what make it so popular. MATLAB will be used to implement YOLO. With MATLAB, there is a deep learning toolset.***
Keywords: Object Detection, YOLO V2, ResNet, Bounding Boxes.

I. INTRODUCTION

Detecting objects in photographs or films taken underwater is a very difficult task. This research aims to develop a model that can detect and locate underwater objects using deep learning techniques. These techniques help computers classify untagged data as human. Underwater object detection has several functions Impairments such as amblyopia, colour reduction, and uneven lighting make objects look less shiny and truer to their true colours. Accurate and reliable detection allows proper identification of different targets in different experimental settings submerged object. The best and most accurate results come from training a big dataset, whereas insufficient training and worse accuracy come from training a smaller dataset. As there aren't many sizable datasets relevant to the undersea environment, it's critical to have a range of performance enhancement techniques that make the most of the sparse data sets used to train our models. We employ a pre-trained ResNet50 model that is based on transfer learning and uses ImageNet weights for picture categorization. Object localization and object classification are the two components of object detection. Classifying an item into predetermined classifications is known as object classification while identifying an object by its position is known as object localization. The input image will be used by the object detection system, which will then pre-process it. At this point, the image will be scaled. The system collects the characteristics from the image and then classifies the item using Deep Learning algorithms. Bounding boxes are used to locate the item using a finite number of classes.

II. LITERATURE SURVEY

Real-time Detection of Underwater Objects Using Multiscale ResNet, a Programmed Framework for Identifying Underwater Objects According to Huang-Chu Huang2, Tien-Szu Pan1, Jen-Chun, Lee2, and Chung-Hsien Chen [2] reduce costs Underwater test. So they use a multi-scale ResNet to detect underwater objects. The trained underwater video Frame proposes a convolutional neural network architecture, a modified architecture of ResNet called Multiscale ResNet(M-ResNet), which improves efficiency and further develops by using multi-scale tasks with precise locations of various types of objects, and sizes, especially for small objects.
Lightweight Deep Neural Network for Concurrent Learning of Colour rConversion and Underwater Item Detection, To recognize things underwater, Chia-Hung Yeh, Chih-Hsiang Huang, Chu-Han Lin, Li-Wei Kang, Min-Hui Lin, Chuan-Yu Chang, and Chua-Chin Wang [3] created a network called a Lightweight Deep Neural Network. The main components of the proposed model of a lightweight deep neural network for submerged object recognition are an image colour conversion network and an underwater item detection network. A colour conversion network module is created in order to alter the colour information in underwater photographs and improve object recognition, as underwater images frequently suffer from colour distortion. They have focused on marine divers, fish, and detritus.

©IJRASET: All Rights are Reserved

III. RESEARCH METHODOLOGY

A. YOLO V2 Architecture

The YOLO-V2 architecture has two sub-networks. Recognition network after feature extraction network. In this example, ResNet-50 is used for feature extraction. The input image size is 224 x 244 x 3. There are 22 convolution layers and 5 layers for maximum pooling. The model proposed in this article is based on a unique neural network architecture. This technique uses YOLO to find underwater objects. YOLOv2 is sometimes referred to as YOLO9000. YOLO9000 is a real-time object detection system can detect more than 9000 kinds of objects. By using fully connected layers, YOLO directly predicts the coordinates of bounding boxes on the segment of the convolution function. The main objective of this study is to develop a model to recognize and locate underwater objects using deep learning. YOLO is an efficient algorithm for real-time object recognition ("You Only Look Once").

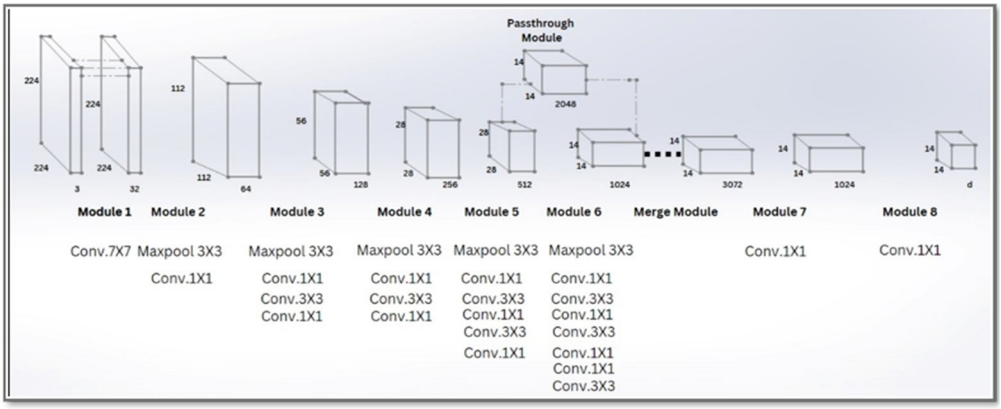


Fig 1: YOLO V2 Architecture

Image classification is the process of classifying an image into one of the various categories.
Object localization helps us find the required object in the image is known as object localization.

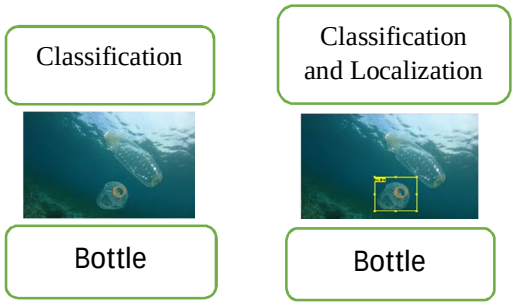


Fig 2: Object Detection for A Single Object

Every object in a picture may be recognized using object recognition tools, which also enable the generation of bounding boxes around those objects. Bounding boxes are used by a neural network to detect and identify the items in a picture. In the context of bounding boxes, there are various situations where we need to establish the exact limits of our objects. Instance segmentation is the phrase used to describe this process. In YOLO, our objective is to determine an object's class and to use bounding boxes to indicate the object's position.

Four descriptors can be used to characterize each bounding box:

- bounding box's center (bx, by)
- width (bw)
- height (bh)
- Value corresponds to an object's class (e.g., fish, covers, bottles, etc.)

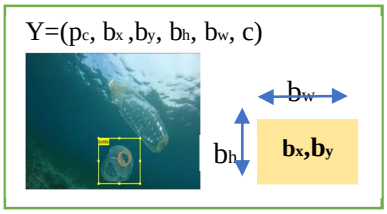


Fig 3: Descriptors of Bounding Box

It is necessary to predict the pc value, which indicates the likelihood that an object is present within the bounding box. Each cell in a picture is given the task of guessing five bounding boxes. With one image, several bounding boxes came. The majority of these cells and bounding boxes won't contain any items. A method called non-max suppression is used to exclude boxes with low object probability and bounding boxes with the highest common area.

B. Activation Function

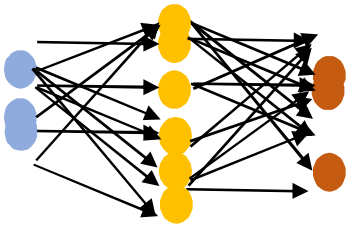
We most likely employed two different types of deep learning activation functions in our model, one for hidden layers and one for output layers. At buried layers, the ReLU activation function is used. ReLU stands for Rectified Linear Unit. It's a non-linear activation function in deep neural networks or neural networks with several layers. The equation for Relu is f(x)=max (0, x). We utilized the soft Max activation function in the output layers since it frequently resolves multiclass classification problems.

C. Pooling Layers

"Pooling layers" is a method for down sampling feature maps that adds up the number of features across different feature mapping regions. Average pooling and max pooling are two well-liked pooling algorithms that, respectively, characterize a feature's average presence and its most active existence.

D. Fully Connected Layer

Combining linear functions (y = Wx + b) with nonlinear functions (Sigmoid and ReLu) yields fully connected or hidden layers. The fully connected layer first receives the input from the flattening layer, and then the fully connected layer's neurons linearly modify the input vector using the weight matrix. The output vector is generated by computing the probability distribution of all classes in the final set after a nonlinear transformation of the product using a nonlinear function.



Input Layer Hidden Layer Output Layer

Fig 4: Fully Connected Layer

E. Intersection Over Union

Intersection over Union (IoU) is used to evaluate the performance of object detection by comparing the ground truth bounding box with the predicted bounding box, providing a ratio of overlap to total area. The prediction is positive if the IoU number is greater than 0.5. If it is below that, the forecast is negative.

$$\mathbf{IOU}=\frac{\boldsymbol{Area\ of\ Overlap}}{\boldsymbol{Area\ of\ Union}}$$

F. Confusion Matrix

Confusion matrix is a representation of the classifier model's predicted value and real value in a table-like structure, where all rows correspond to predicted value and all columns to real value. All values in the table correspond to the number of inputs that were given to the model for classification.

We must first become familiar with the following terminologies in order to completely comprehend the confusion matrix for this binary class classification problem:

1) *True Positive (TP):* A sample that is accurately identified as being in a positive class is said to be true positive (TP).
2) *True Negative (TN):* A sample that was accurately identified as being in the negative class is known as a true negative (TN).
3) *False Positive (FP):* A samples are those that belong to the negative category but are mistakenly identified as being in the positive category.
4) *False Negative (FN):* A samples are those that belong to the positive class but are mistakenly categorised as negative samples.

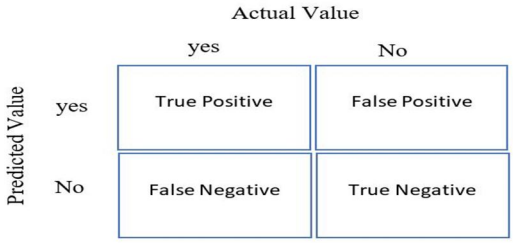


Fig 5: Confusion matrix

- Precision = $\frac{TP}{TP+FP}$
- Recall = $\frac{TP}{TP+FN}$
- Accuracy = $\frac{TP+TN}{TP+FP+TN+FN}$
- F1 Score = $\frac{2*(precision*recall)}{precision+recall}$

G. Residual Network (ResNet-50)

The 50-layer neural network residual deep learning model is called ResNet50. The architecture of ResNet50 includes 4 levels. ResNet is a typical neural network that acts as a framework. For computer vision applications. ResNet has successfully trained over 150 very deep neural networks. Skip Connection was first invented by ResNet. Use jump links or shortcuts to jump some layers (Highway Net can also learn the jump weights itself via the gate's additional weight matrix). Common ResNet models are implemented with two or three layers of hops with nonlinearity (ReLU) and batch normalization in between.

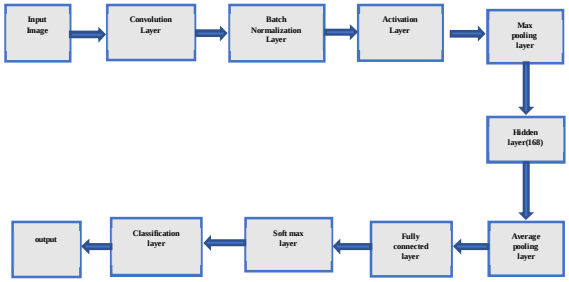


Fig 6: Block Diagram of Resnet-50

IV. RESULT



Fig 7: Images with object detection using YOLO-V2

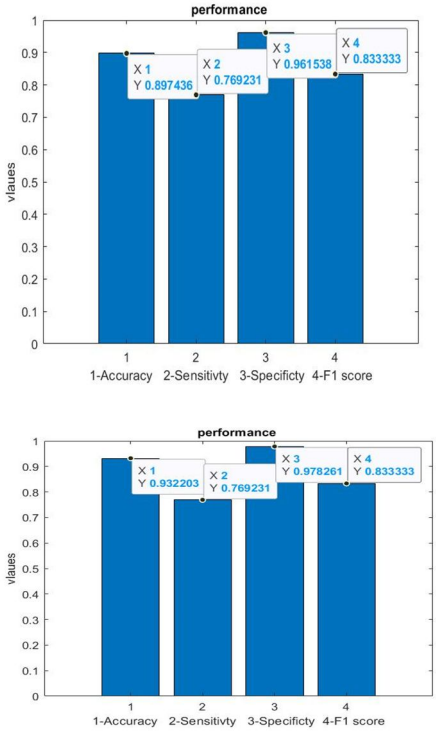


Fig 8: Performance with learning rate 0.001 using YOLO-V2

V. CONCLUSION

This model is proposed to recognize underwater objects. The model can be used to identify divers, marine life, plastic, and several other objects**.** Because of its excellent accuracy, fast execution, and efficient performance, the YOLOV2 model is used. Classifying an object into specified classes and isolating it from other objects depending on their locations are both aspects of object localization and object classification. We used non-max suppression to eliminate boxes with the lowest item probabilities and bounding boxes with the highest common area. The intersection over union metric is used in object identification models to evaluate localization accuracy and calculate localization errors.

In MATLAB, a deep learning toolkit is offered. Using algorithms and learning models, Deep Learning Toolbox is a tool for constructing and deploying deep neural networks. We have trained our model with different learning rates and epochs values among them our models stand at the best accuracy of 95.9%. In the Future we want to do with Large Dataset and with an advanced version of YOLO to enhance Accuracy and other Metrics**.**

REFERENCES

[1] Tien-Szu Pan, Huang-Chu Huang, Jen-Chun Lee, Chung-Hsien Chen, "Multi-scale ResNet for realtime underwater object detection." Signal, Image and Video Processing, 15(5), 941–949. doi:10.1007/s11760-020-01818-w Received: 16 July 2020 / Revised: 2 November 2020 / Accepted: 6 November 2020 © Springer-Verlag London Ltd., part of Springer Nature 2020]

[2] Chia-Hung Yeh, Chu-Han Lin, Li-Wei Kang, Chu-Han Lin, Li-Wei Kang, "Lightweight Deep Neural Network for Joint Learning of Underwater Object Detection and Color Conversion. ", IEEE Transactions on Neural Networks and Learning Systems, 1–15. doi:10.1109/tnnls.2021.3072414

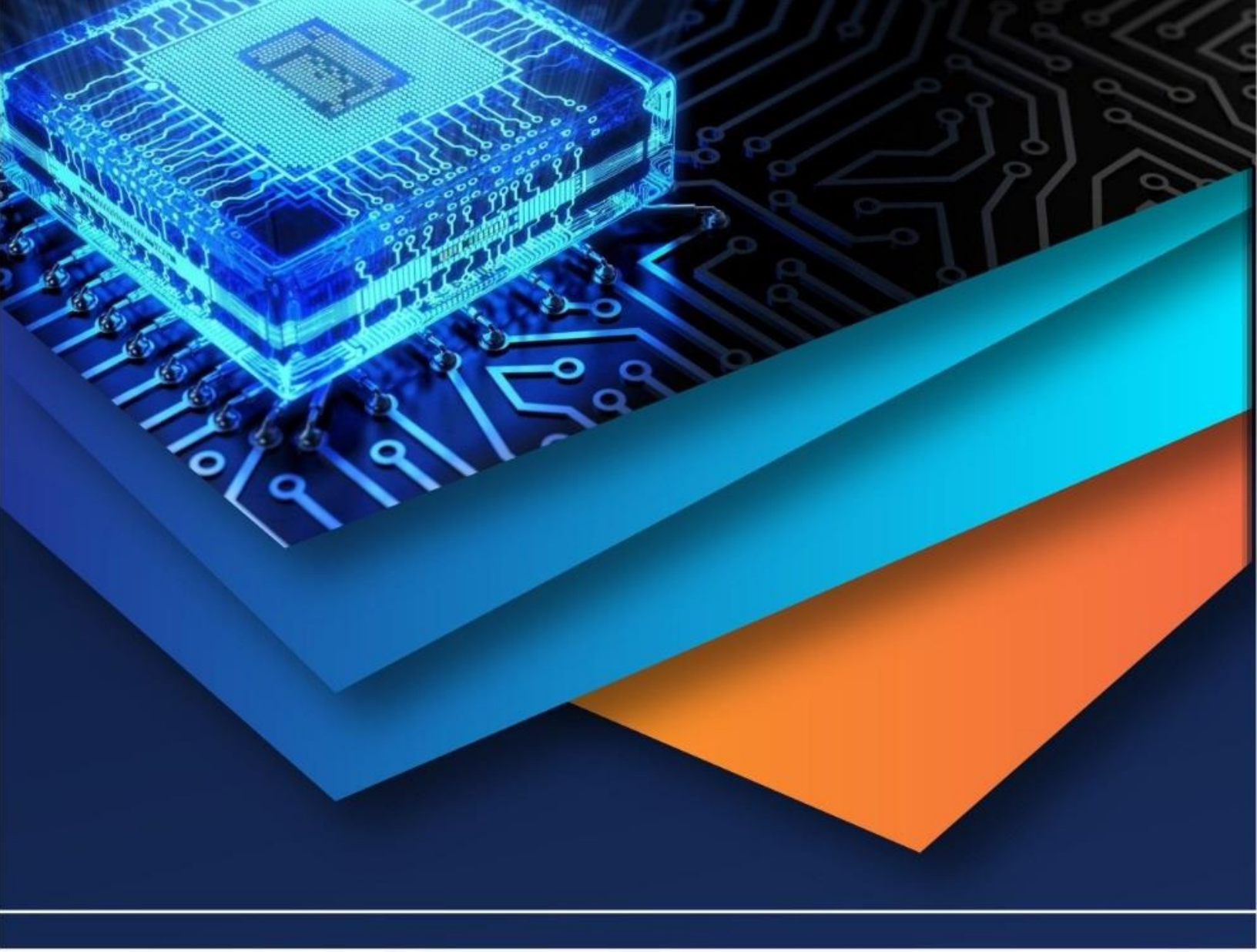
[3] Gautam Tataa, Jay Lowea, Olivier Poirionb, and Sarah-Jeanne Royerc," A Robotic Approach towards Quantifying Epipelagic Bound Plastic Using Deep Visual Models" arXiv:2105.01882v4 cs.CV19 Oct 2021

[4] Priyadharsini, R., & Sharmila, T. S. (2019). Object detection in underwater acoustic images using edge-based segmentation method. Procedia Computer Science, 165, 759-765.

[5] Zhang, L., Li, C., & Sun, H. (2022). Object detection/tracking toward underwater photographs by remotely operated vehicles (ROVs). Future Generation Computer Systems, 126, 163-168.

[6] Yeh, C. H., Lin, C. H., Kang, L. W., Huang, C. H., Lin, M. H., Chang, C. Y., & Wang, C. C. (2021). Lightweight deep neural network for joint learning of underwater object detection and color conversion. IEEE Transactions on Neural Networks and Learning Systems.

[7] Karimanzira, D., Pfützenreuter, T., & Renkewitz, H. (2021). Deep learning for long and short-range object detection in the underwater environment.



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)