ELSEVIER Scopus doi

Journal of Vibration Engineering

ISSN:1004-4523

Registered

- SCOPUS
- GOOGLE SCHOLAR
- DIGITAL OBJECT IDENTIFIER (DOI)
- IMPACT FACTOR 6.1

UnderwaterTargetDetectionUsingYOLOv4Basedon Machine Learning

Prof.UmaKaranje
CSEDepartment
MarathwadaMitramandal'sInstituteofTechnology
pune,india
uma.karanje@mmit.edu.in

Pratik Salunkhe
CSEDepartment
MarathwadaMitramandal'sInstituteofTechnology
pune,india
pratik.salunkhe@mmit.edu.in

SujitKavitake
CSEDepartment
MarathwadaMitramandal'sInstituteofTechnology
pune,india
sujit.kavitake@mmit.edu.in

UdaySonaje
CSEDepartment
MarathwadaMitramandal'sInstituteofTechnology
pune,india
uday.sonaje@mmit.edu.in

SwapnilGawali
CSEDepartment
MarathwadaMitramandal'sInstituteofTechnology
pune,india
swapnil.gawali@mmit.edu.in

Abstract **- Tasks involving the discovery of targets visible in underwater photographs are referred to as underwater target detection tasks. In contrast to conventional target identification tasks, the effect of target detection will be impacted by underwater targets due to issues like illumination, camera motion, complicated background interference, and the diversification of target kinds. The identification of underwater targets is a crucial component of national military security. Deep learning-based technology for underwater target detection has advanced significantly in recent years, but there are still certain issues, such as features and contours that are not immediately evident. We provide a target detection technique based on deep learning and picture enhancement. The YOLO v4 algorithm first improves the contrast of the picture data to acquire a better target detection performance, and then it employs a deeplearning approach to distinguish the target from the background. The algorithm can produce superiordetection performance, according to experimental data.**

Keyword – Underwater Target Detection, complicated background interference, Convolution Neural Network, YOLO v4

I. INTRODUCTION

Underwater target detection aims to identify things in the underwater environment. This research may improve underwater target recognition efficiency and have practical implications for navigation and aquaculture in a range of use cases. Due to the underwater environment, significant color distortion,andsmallunderwaterobjects,itisdifficulttolocate targets underwater. As the target detection technique based on machine learning is used to underwater target recognition and has attained a certain levelofsuccess, more and more relevant target detection methods have been introduced and haveshown encouraging performance in real-world application scenarios. Due to the small size of the underwater target detection dataset, the underwater target is modest and only weakly supported by the data.

The existing deep learning target detection algorithms have difficulty identifying small-volume targets. Due to the poor underwater imaging effect, the image has a problem with blur and color distortion, which makes the underwater target features less obvious. In order to extract valuable target information and increase target detection accuracy, a more powerfulCNN isneeded.The study'schosen underweartarget identification technique, YOLO v4, successfully resolves the aforementioned problems. In YOLO v4, the Cross Stage Partial (CSP) structure is used to improve CNN's ability to learn, boost the recognition algorithm's recognition accuracy, and cut down on the amount of parameter calculation.

Whatismachinelearning?

With the aid of machine learning (ML), which is a form of artificial intelligence (AI), software programs can predict outcomesmoreaccuratelywithouthavingtobeexplicitly instructed to do so. In order to forecast new output values, machine learning algorithms use historical data as input.

Machine learning is significant because it aids in the development of new goods and provides businesses with a picture of trends in consumer behavior and operational business patterns. A significant portion of the operations of many of today's top businesses, like Facebook, Google, and Uber, revolve around machine learning. For many businesses, machine learning has emerged as a key competitive differentiation.

HowMachineLearningWorks?

1. A Decision-Making Process: Typically, machine learning algorithms are employed to classify or predict data. Your algorithmwillgenerateanestimateaboutapatternintheinput data based on some input data, which can be labelled or unlabeled.

2. An Error Function: An error function is used to assess how well the model predicts. If there are known examples, an error function can compare them to gauge the model's correctness.

3. A model optimization process: Weights are adjusted to lessen the difference between the known example and the model estimate if the model can fit the data points in the training set more accurately. Until an accuracy threshold is reached, the algorithm will iteratively evaluate and optimize, updating weights on its own each time.

A. OBJECTIVESOFSYSTEM

- TodetectunderwatertargetusingmethodofYOLO v4.
- To Use YOLO v4 it improves CNN's learning capabilities, increase the recognition algorithm's recognition accuracy, and decrease the quantity of parameter calculation

B. PROBLEMSTATEMENT

Small-volume targets are challenging for the current deep learning and machine learning target detection methods tofind. The imagehas an issue with blur and colour distortion as a result of the poor underwater imaging effect, making the underwater target features less evident. As a result, a more potent CNN is required to extract useful target features and improve target detection precision.

II. LITERATURESURVEY

1] Wen Zhang, Yanqun Wu, Yonggang Lin, Lina Ma, Kaifeng Han, Yu Chen, Chen Liu . In this study, machine learning or deep learning was used to successfully implement each of the four key components of underwater target detection (poor channel detection, classification, tracking, and localization).

2] Yuechao Chen, Xiaonan Xu et.al. In this study, The results demonstrate that the deep learning algorithms may significantly enhance the underwater target recognition effect, withtherecognitionaccuracyofDBNandSDAEbeinghigher than that of the other three approaches in every case.

3] YunliangZheng1,MengxueYu2,Zi'aoMa2,Rong Liu3andYang Liu4 et.al. In this study, a deep network and image enhancement-based target detection system used. The approach uses a deep learning algorithm after first enhancing theimagedatatoimprovecontrast.Thealgorithmcanproduce superior detection performance, according to experimental data.

4]XinhuaWang,YungangZhu,DayuLi&GuangZhanget.al. Inthisstudy,Inordertodeterminethetransitionprobabilityof each pixel, a variable radius sensing approach is first presented, and the reinforcement learning concept is then included into the motions of artificial ants. These techniques are meant to prevent some pixels in image borders from being missed or detected incorrectly. Second, a double-population ant colony search approach is suggested, which combines global and local search capabilities. The results of the experiments demonstrate that the algorithm can successfully extract the contour data of underwater targets, maintain the image texture, and have the best anti-interferenceperformance.

5] Hao Yue, Lilun Zhang, Dezhi Wang , Yongxian Wang and Zengquan Lu. In this study, Convolution neural networks (CNN) and deep short networks (DBN) are used in this study, and they can achieve accuracy levels of up to 94.75% and 96.96%, respectively, in both supervised and unsupervised ways. We also employ Wndchrm, which was originally a tool used for biological image analysis, and Support Vector Machine (SVM) to perform the same task in order to compare the results with those of conventional machine learning techniques.

6] Suraj Kamal et al. In this study, used DBN to perform passive target recognition tasks. The findings also demonstrated that generative DBN is superior than systems relying on expert knowledge of underwater acoustic signal processingintermsofitsabilitytoextractmoreconsistentand expressive features from the target. Additionally, using a dataset with 40 different types of underwater targets, it had an accuracy score of 90.23%.

III. SYSTEMARCHITECTURE

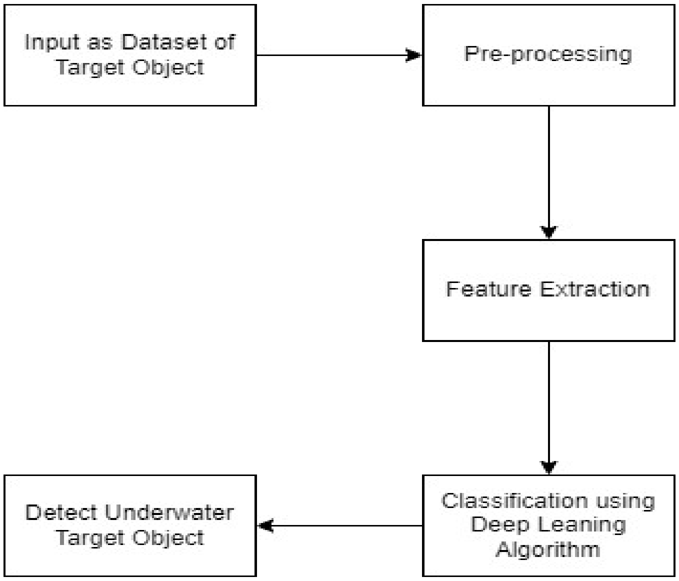

Fig.1SystemArchitecture

Dataset:UnderwaterImagesdataset.
Preprocessing is necessary because real-world data frequently has noise, missing values, and may even be in an unsuitable format that prevents it from being used directly by machine learning models. Cleaning up the data and preparing it for a CNNmodelarenecessarystepsthatalsoimprovetheaccuracy and productivity of a CNN model.
All humans can use feature extraction and feature selection procedures. However, feature extraction in CNN models that extract image pixel values, edges, height, breadth, and color presents a challenge for learning algorithms.
Classification is the process of grouping a set of data into categories. It can be done with both structured and unstructured data. The first step in the procedure is classifying the provided data points. The terms target, label, and classes arefrequentlyusedtodescribetheclasses.Toapproximatethe mapping function from input is the task of classification predictive modeling.

DatasetDescription:
The data set has been used in this study to develop a machine learning algorithm based system for the Underwater object detection. Dataset is used to build an efficient model to classify and predict the underwater target with the help of YOLO v4.

IV. ALGORITHMSUSED

ConvolutionalNeuralNetwork:

CNNfocus on image and video recognition applications.CNN isprimarily utilized forimage analysisapplicationssuch segmentation, object detection, and picture recognition.

Convolutional neural networks have four different kinds of layers:

1) Convolutional layer: Each input neuron in a conventional neural network is connected to the following hidden layer. Only a small portion of the input layer neurons in cnn are connected to the hidden layer of neurons.

2) Pooling layer:Thepooling layer isused to makethefeature map less dimensional. Inside the cnn's hidden layer, there will be numerous activation and pooling layers.

3) Flatten layer: Flattening is the process of reducing data to a 1-dimensional array so that it may be entered into the following layer. We flatten the convolutional layer output to produce a solitary, lengthy feature vector.

4) Fully connected layer: Fully connected tiers make up the network's final few layers. The output from the last pooling or convolutional layer is passed into the fully connected layer, where it is flattened before being applied.

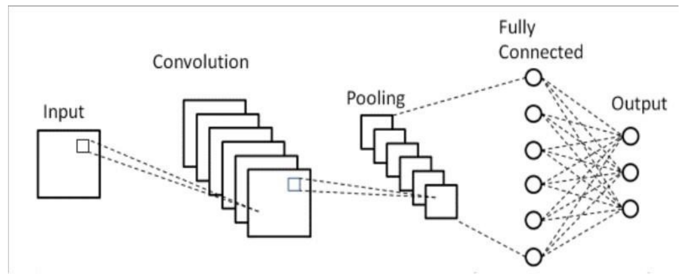

Fig.2CNN'sArchitecture

YOLO v4:

YOLO is a You Only Look Once and v4 is a forth version ofit.WiththereleaseofYOLOv4,thereal-timeobjectdetection market stays competitive and advances constantly. YOLO v4 performs significantly better than other object detectionmodels in terms of inference speed.

The fourth model in the You Only Look Once family is called YOLO v4. Real-time detection is prioritized in YOLO v4 and training is carried out on a single GPU. It is a one step target detection model. Thegoalof the entity isto makeitsimplefor visual engineers and programmers to use their YOLO v4 framework in unique areas.

YOLOmodelsallfunctionasobjectdetectionmodels.Models for object detection are trained to scan a picture for a certain subset of object types. When discovered, these object classes are placed inside the bounding boxes and given a class designation. The COCO dataset, which includes an extensive collection of 80 object classes, is commonly used to train and evaluate object detection models. From there, it is assumedthat if object detection models are exposed to fresh training data, they will generalize to fresh object detection tasks.

V. ADVANTAGES,DISADVANTAGESAND APPLICATIONS

A. ADVANTAGES

•High precision, simplicity, ease of installation and operation, and a respectable number of frames per second

• Doesn'tcallforextrahardware

• EasierComputerControl

• It will make it easier for amputees to use computers, which will lead to an increase in users.

• Simplercomputeroperation.

•Reducedworkinfeatureengineeringduetoautomaticfeature extraction from raw data.

•Duetotheirgreataccuracyandquickprocessingtimes, underwater robots are widely used.

•AnotherbenefitofYOLOv4technologyisagoodgeneralization.

B. DISADVANTAGES

•After the transfer to the target domain, the model is fixed and is greatly affected by the existing target domain.

•Themodelistoounstabletocontrol

•TheModelparametersaredifficulttoobtain.Themodelis less flexible and sensitive to abnormal samples.

C. APPLICATION

- Underwater target detection technology useful for a underwater environment studies. It is Essential forthe inspection and understanding of underwater scenarios.
- Underwater target detection technology useful the military sector.
- Underwater target detection technology useful the fishery sector.
- Underwater target detection technology useful for a rescue missions.

VI. CONCLUSION

Recent developments in deep learning algorithms haveallowed for some breakthroughs in underwater target recognition. But there are still some challenges. Along with poor underwater image effects, there is color distortion and noise. Underwater creatures are frequently tiny and obscure.YOLO v4 improve CNN ability to recognise small objects, and improve its learning capabilities whilemaintaining accuracy.

VII. FUTURESCOPE

The development of the underwater target recognition method will prioritise real-time performance, intelligence, autonomy, high precision removal rate, and robustness. Its broader influence will be advantageous to both the military and the civilian sectors.

VIII. REFERENCES

[1] AlexeyBochkovskiy*,Chien-YaoWang*,Hong-Yuan Mark Liao. Yolov4:Optimal speed and accuracy ofobject detection. arXiv preprint arXiv:2004.10934, 2020

[2] Chien-Yao Wang, Hong-Yuan Mark Liao, Yueh-Hua Wu, Ping-Yang Chen, Jun-Wei Hsieh, and I-Hau Yeh. CSPNet: A new backbone that can enhance learning capabilityofcnn.ProceedingsoftheIEEEConferenceon Computer Vision and Pattern Recognition Workshop (CVPR Workshop), 2020.

[3] JosephRedmonandAliFarhadi.YOLOv3:An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.

[4] KaimingHe,XiangyuZhang,ShaoqingRen,andJian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), 37(9):1904–1916, 2015.

[5] L. Shamir, N. Orlov, D. M. Eckley, T. Macura, J. Johnston,I.G.Goldberg,"Wndchrm–anopensourceutility for biological image analysis," Source Code for Biology and Medicine, 2008, pp. 3-13.

[6] C. C. Chang, C. J. Lin, "LIBSVM: A library for support vectormachines,"AcmTransactionsonIntelligentSystems& Technology, 2011, vol. 2, no. 3, pp.27-32.

[7] L. Shamir, C. Yerby, R. Simpson, A. M. von Benda-Beckmann,P.Tyack,F.Samarra,etal,"Classificationoflarge acoustic datasets using machine learning and crowdsourcing: Application to whale calls, Journal of the Acoustical Societyof America, 2014, vol. 135 No. 2, pp. 953-962.

[8] S.Kamal,S.K.Mohammed,P.R.S.Pillai,M.H.Supriya, "Deep learning architectures for underwater target recognition,".Sympol2013,pp.48-54.