Underwater Mines Detection using Neural Network

Shantanu
Dept. of Computer Science and Engineering
RV college of engineering
Bangalore, India

Aman Saraf
Dept. of Computer Science and Engineering
RV college of engineering
Bangalore, India

Atharv Tiwari
Dept. of Computer Science and Engineering
RV college of engineering
Bangalore, India

Mukesh Kumar
Dept. of Computer Science and Engineering
RV college of engineering
Bangalore, India

Prof. Manonmani S
Assistant Professor
Dept. of Computer Science and Engineering
RV college of engineering
Bangalore, India

Abstract—The usage of underwater mines in warfare poses a great security and safety threat to the naval defense system. With the evolving technological advancements, the use of underwater mines is an ever-growing concern. Previous studies have relied upon the use of side-scan sonar imagery where the accuracy of the model is concerning. It is susceptible to false alarm in rigorous geographic locations. The purpose of this study is to investigate a system that can provide the naval forces with accurate data in shortest time possible. In this paper Mask RCNN model has been used for mines detection. ResNet-50 architecture has been used to implement Mask RCNN. Image pre-processing has been done which is followed by Mask RCNN using FPN for feature extraction. On successfully implementing the system it was found that mines were detected with satisfactory accuracy. This study can be extended to detect other marine objects using faster RCNN.

Keywords—Mask RCNN; Computer Vision; Neural Network; Image Processing; ResNet

I.   INTRODUCTION

Underwater mines also known as naval mines are explosives used in warfare. The mines are used to destroy enemies naval attack. It is also used in defense sector where the countries oceanic region is guarded by mines acting like a border. These mines prevent the enemy ship from entering unmarked dominion. The opponent need to sweep the entire area for mines. Underwater mines force the opponent to attack in unmined location where the defense are prepared for a battle. The modern mines are detonated by push of a button unlike the older mines where [1].

Detection of underwater mines is highly essential to make sure civilians are not harmed in any way. The mines help in assuring the security of high level defense bases and avoid the leak of valuable information. A reliable and cost effective system will help battle group to determine the exact location of mines and avoid casualties.

The neural network can be compared to the working of a human brain. It is used to represent the relationship between data throughout the computer system. Machine Learning is heavily based on this artificial network. The neural network performs its task by learning on the data provided. The more the network learns the better the result. It is made up of number of cells interlinked by neurons. Each of these cell work alone on only a small objective. These cells communicate among each other using the neurons to form a larger system.

Mask RCNN is a deep neural network which is used to solve segmentation problem along with the object detection in an image or a video. Mask RCNN generates a proposal about the region in the image where object might be present and later generates bounding boxes and mask in pixel level and predicts the class of the object. Mask RCNN uses FPN as backbone for generating feature vector form raw images. RPN for searching object in regions. For clubbing the feature vector with the location in raw image anchor boxes are used which can be used for comparing with ground truth while detection using the concept of IoU value [9].

II.   LITERATURE REVIEW

Underwater image processing has been significantly used in detection of mines. This is implemented by using Autonomous Underwater Vehicles (AUV). These AUV are deployed in the detection region to gather the location of the mines. The collected data is then send to the base where necessary action is taken. These Autonomous vehicles use underwater camera sensors. These sensors offer major drawback. The underwater image obtained from the AUV are

not accurate due to scattering and noise. The primary cause of this phenomena is the light transportation characteristic in water along with sea floor's biological activity [2].

Use of Underwater Optical Imaging is a bothersome. There are many obstacles that come into play when compared to normal photography. Underwater Images get blurred due to scattering effect. The color is reduced due to absorption of wavelength. Noise along with water residuals out-turn the captured image. Underwater imagery does not use natural lightning, thus it results in deprecation of the image. The flicker cannot be avoided in day time [2].

A study on Side-Scan Sonar was conducted. Here the sonar signatures are fed to the network. This technology maps large sea floor area for surveying. The signal that is fed to the network preprocesses the data and then applies the training algorithm. Training is followed by segmentation where the sonar image is segmented into sub frames. Feature Extraction then defines the object property and portions the data set [3].

Side–Scan Sonar Imagery is a challenge in mines detection. The environment of the deployed system may vary. This causes the accuracy of the image to drop. The target mines are of various shapes and the side scan sonar finds it difficult to deal with the variations. Underwater habitation like coral and reef can be wrongly detected. These can even hide the mine like objects. The sonar signal is also unreliable due to its time taken to detect. The side scan sonar sends a sonar wave to the target object and receives the data by mapping the object. The time taken is considerable high and in extreme oceanic conditions the side scan sonar is highly inaccurate [3].

Underwater object can also be detected using transformable template matching approach. In this method feature extraction takes place by constructing templet from sonar video sequences. This is done by analysis of acoustic shadows along with highlighting regions. Fast saliency detection is a technique used to identify the target region. Normalized gradient feature is extract in the next step followed by calculation of similarity which is done between target and template [4].

Upon studying the Threshold based model it was found that is has been commonly used in the past. This model is not the best for the mines detection purpose. The model has high computation and the contour model which is active affects the initial contour. The model's sonar image is of poor quality. Noise and decoy targets do not lead to accurate recognition [4].

Forward Looking Sonar imagery uses the integral-image representation. This competes features in a small amount of time. The computational load is reduced significantly. The work is done on small portions of the image. This algorithm works fine for real time object detection. The high demand for real time processing of the signal make the detection a challenge. The algorithm lacks density filtering. Here the algorithm needs to ignore the fact that many mines will lie together. Shadow casted by mine like object also gets detected [5].

Multi beam sonar image processing is another method that has been used for underwater image detection. This method uses the BluView (BV) Sonar. The data collected is real-time and this data is transformed to an image and then preprocessed. Contour Detection algorithm separates the foreground from background. Objet tracking is done in this algorithm. This is achieved by particle filter tracking method. The tracking method implements the adaptive fusion tracking method. Sonar image have high noise and low contrast making the method unreliable [6].

Adaptive Fuzzy Neural Network is another interesting way to study the underwater objects. Here feed forward and pattern recognition is used. MATLAB is an important tool used in its making. Texture features are computed. Objects are classified using features like autocorrelation, sum variance, sum average, and sum entropy. The texture feature is trained and then the model is tested for classification [7].

Monocular Vision sensors are used in underwater object detection. Light Transmission is used in this technique. Region of Interest (ROI) is identified using global contrast feature. Monocular Camera is used to create dataset. It produces varying images and advances the testing model. It can remove noise and increase accuracy of the system. There are drawbacks to the method. The camera used goes through intensity degradation. Distortion of color and haze effect add to its disadvantage [8].

III. PROPOSED SYSTEM

In the proposed system, we are working on identifying underwater mines using neural network. The system takes an image or a group of image as an input. The input can be in any format. This input image is supplied to the pre trained model where image processing takes place. The whole process is split into several necessary stages, starting with gathering of images to detection using deep neural networks all of which have been mentioned in the subsections below. Figure 1 presents the overall flowchart of module.

Dataset creation

For a machine learning algorithm to work the primary step is to collect data. In this paper data is an image. The images have been downloaded from multiple websites and sources consisting of different types of underwater mines. These images are further divided into two parts for training and validation of the neural network module.

Image pre-processing and labelling

The collected images may have varying quality and resolution. To resolve this issue the images initially need to be pre-processed. Pre-processing of image is done by clipping operation and adjusting the aspect ratio by resizing and reshaping. Since some images are blurred they have been adjusted for contract and sharpness to make it clear. For training any neural network two thing are required dataset and its annotation. Here training set is labelled manually using annotation tool and for labelling of images MS COCO format is used.

Augmentation Process

To increase the accuracy of a neural network model the size of dataset matters. One of the common techniques to increase the size of the dataset is image augmentation. The augmentation technique uses random rotation, horizontal flip, vertical flip and shifts operations on the pre-processed images. Performing augmentation also leads to rotational invariance to neural network.

Prepare Training and Test Image Sets

Split the whole dataset into training and validation data. Pick 70% of images from the set as the training data and the remaining 30% for the validation data. Randomization operation is performed during the split to avoid biasing the results.

Fig-1 *The overall flowchart of the system*

Transfer learning

To optimize a predictive model and overcome dataset constrains, knowledge of any base network is repurposed to a second network that is to be trained on a target dataset by using learned feature. The type of transfer learning used in this paper is inductive learning. The overall goal is to infer mapping from a set of training example. The pre-trained model used is ResNet-50 architecture. ResNet-50 introduced the technique of skipping convolutional layer from stack of convolutional layers which allow the model to learn an identification function which allows the higher layer to perform as good as lower and not worse [10].

Detection

For identifying the target object form the image, a deep learning algorithm named Mask RCNN is used, which is the extension of faster RCNN. Mask RCNN is the new form for detection of object along with instance segmentation. The backbone used for generating feature vector from the given image is a CNN model. In this paper the backbone is ResNet-50. After generating feature map, concept of anchor box is used to determine the target object. In this paper the size of anchor box varies from 8 to 128. For generating mask and defining class of the object use a neural network called RPN. The training of whole model is divided into two parts. In first part only heads of model are trained by taking help of coco pre-trained weights. The second part is training of bottleneck layer which is just before the output layer. For training both layer different number and epoch is used and learning rate that is 6 and 4 epochs and 0.001 and 0.0001 as learning rate. The activation function used for output layer is a softmax function which gives the probability of the object it belongs to the class [9].
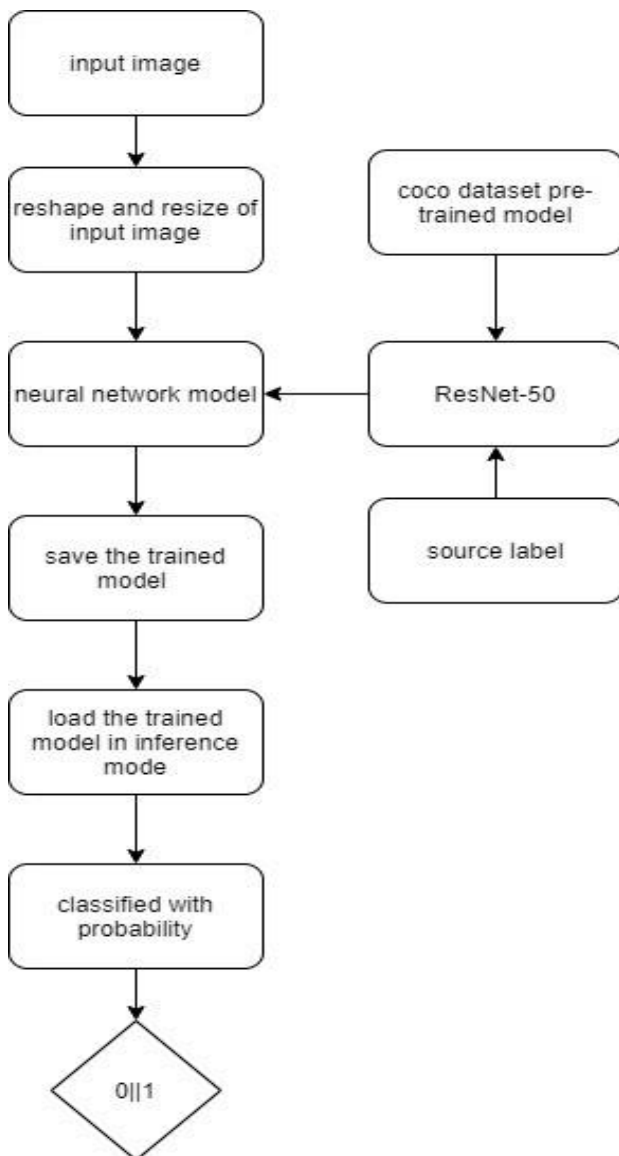
IV. RESULT

For improving the result of the overall system attributes of the model like learning rate, no of epoch and steps per epoch are the deterministic factor. For evaluation of system the metric used here is mAP (mean average precision). After training the model for 6 and 4 epochs with learning rate of 0.001 and 0.0001 for head and all layers respectively, the value of mAP is 0.740 by running the model on training and validation dataset.

```python
from mrcnn.model import load_image_gt
from mrcnn.model import mold_image
from numpy import expand_dims
from numpy import mean
from mrcnn.utils import compute_ap
from numpy import zeros
from numpy import asarray


# evaluate model on training dataset
train_mAP = evaluate_model(dataset_train,model,inference_config)
print("Train mAP: %.3f" % train_mAP)
# evaluate model on test dataset
test_mAP = evaluate_model(dataset_val,model,inference_config)
print("Test mAP: %.3f" % test_mAP)

Train mAP: 0.740
```
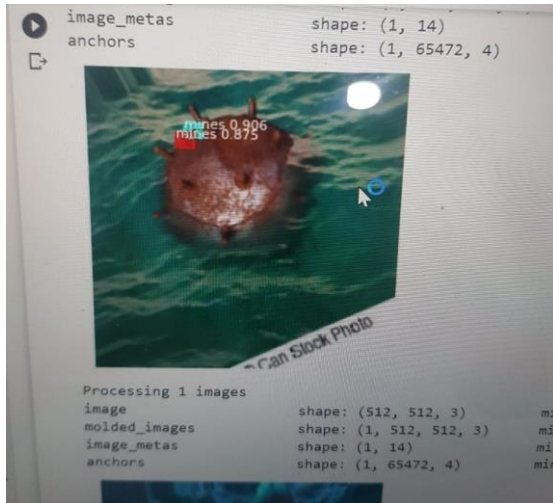
Fig-2 showing mAP value of model

Fig 3 showing the output of system

V. FUTURE SCOPE

This prototype system is just a base of a simple implementation. This field offers a lot of scope for work and research in the coming future. There is room for improving the overall system so that it can work and adapt in various underwater environment. The module can be further extended to accept video as an input and detect different marine objects like fishes, plants, and minerals.

VI. CONCLUSION

In this paper a neural network was used to detect target object. Self-created image dataset is used to train the neural network model. Augmentation and pre-processing of image has been carried out before passing it through the neural network. Since the model uses Mask RCNN which is extended version of faster RCNN, it can be used for generating mask along with object detection. Thus it can be used for wide number of application like autonomous underwater vehicle (AUV) by attaching a camera module and training on different type of marine object dataset.

ACKNOWLEDGMENT

REFERENCES

[1] S. N. Geethalakshmi, P. Subashini, S. Ramya, "A Study on Detection and Classification of Underwater Mines Using Neural Networks", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, ISSN: 2231-2307, Volume-1, Issue-5, November 2011.

[2] Huimin Lu1,*, Yujie Li2,*, Yudong Zhang3, Min Chen4, Seiichi Serikawa1, Hyoungseop Kim1, "Underwater Optical Image Processing: A Comprehensive Review".

[3] Chinmay Rao Kushal Mukherjee Shalabh Gupta Asok Ray Shashi Phoha, "Underwater Mine Detection using Symbolic Pattern Analysis of Sidescan Sonar Images", 2009 American Control Conference, Hyatt Regency Riverfront, St Louis, MO, USA, June 10-12,2009.

[4] Jianjiang Zhu,1 Siquan Yu,2,3 Zhi Han ,2 Yandong Tang,2 and Chengdong Wu3, "Underwater Object Recognition Using Transformable Template Matching Based on Prior Knowledge", Hindawi Mathematical Problems in Engineering Volume 2019, Article ID 2892975, 11 pages https://doi.org/10.1155/2019/2892975.

[5] Enric Galceran, Vladimir Djapic, Marc Carreras, David P. Williams," A Real-time Underwater Object Detection Algorithm for Multi-beam Forward Looking Sonar", DPI2011-27977-C03-02) and the TRIDENT EU FP7-Project under the Grant agreement No: ICT-248497.

[6] Min Li, Member, IEEE, Houwei Ji, Xiangcun Wang, Liyuan Weng and Zhenbang Gong, "Underwater Object Detection and Tracking Based on Multi-Beam Sonar Image Processing", International Conference on Robotics and Biomimetics (ROBIO) Shenzhen, China, December 2013.

[7] U. Anitha and S. Malarkkan, "International Conference on Robotics and Biomimetics (ROBIO) Shenzhen, China, December 2013", Indian Journal of Geo Marine Science Vol. 47 (43), March 2018, pp 665-673.

[8] Zhe Chen,1,2 Zhen Zhang,1 Fengzhao Dai,3 Yang Bu,3 and Huibin Wang1, "Monocular Vision-Based Underwater Object Detection", NCBI PMID: 28771194.

[9] Kaiming He ; Georgia Gkioxari ; Piotr Dollár ; Ross Girshick, "Mask R-CNN", 2017 IEEE International Conference on Computer Vision (ICCV) ISSN: 2380-7504

[10] Kaiming He ; Xiangyu Zhang ; Shaoqing Ren ; Jian Sun, "Deep Residual Learning for Image Recognition", 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) ISSN: 1063-6919