DETECTION AND EXTRACTION OF SEA MINE FEATURES USING CNN ARCHITECTURE

Sheethal S, Soundarya S, Vidhisha V, Kiran G, Dr A B Rajendra
Department of Information Science and Engineering
VidyaVardhaka College of Engineering, Mysuru

Abstract

The Conventional worry of navel asset is naval mines; these mines are stationary and were planted during war times and now they have been working as a threat to naval ships, and submarines. Detection of those naval mines has been one of the foremost risk-taking tasks, with modern technology various techniques are wont to detect these mines Using Ultrasonic signals, Symbolic pattern analysis of side-scan sonar images but detection through image processing has been one of the most challenging and efficient ones since it can solve the real-time problem with less error, the image classification model like uses FRCNN(Fast Region Convolutional Neural Network) algorithm to classify the objects as mine or not. The cloud platform is employed to watch the mine and as soon as the changes are observed the Android application will reflect the changes.

Keywords: FRCNN, Neural network, Image processing, Deep Learning, ResNet, TensorFlow, Python

1. INTRODUCTION

Aquatic mines also known as nonmilitary mines are used in warfare. The mines are used to destroy naval assets during war. It's also used in the defence sector where the country's oceanic region is guarded by mines acting like a border. These mines help the adversary marine assets from entering unmarked dominion. The opponent needs to sweep the entire area for mines. Aquatic mines force the opponent to attack in the unmined position where the defence is prepared for a battle. The ultramodern mines are exploded by the drive of a button, unlike the aged mines. The Discovery of aquatic mines is largely essential to make sure civilians aren't harmed in any way. The mines help in assuring the security of high position defence bases and avoiding the leak of precious information. A dependable and cost-effective system will help battle groups to determine the exact position of mines and avoid casualties. The neural network can be compared to the working of a mortal brain. It's used to represent the relationship between data throughout the computer system. Machine Literacy is heavily grounded on this artificial network. The neural network performs its task by learning from the data handed. The more the network machine learns the better the result. It's made up of several cells connected by neurons. Each of these cells works alone on only a small ideal. These single cells communicate among each other using the neurons to form a larger system. Mask RCNN is a deep neural network used to break segmentation problems and object discovery in an image or a videotape. Mask RCNN generates an offer about the region in the image where the object might be present and latterly generates bounding boxes and mask at pixel position and predicts the class of the object. Mask RCNN uses FPN as the backbone for generating point vectors from raw images. RPN for searching objects in regions. For allying the point vector with the position in raw image anchor boxes are used which can be used for comparing with ground verity while discovery using the conception of IoU value.

2. LITERATURE REVIEW

[1] The paper mentions the simplest way of implementing various deep learning techniques, hence the modifications need to be done to the techniques this could be the major challenge of the paper. In this paper

modules and sub-modules used are **CNN, Autoencoders, Deep Belief Networks, and GAN.** This paper provides an overview of the simplest ways to implement target recognition and hence not very efficient.

[2] In this paper, the author Huu-Thu Nguyen, Eon-Ho Lee 1, and Sejin Lee specified **sonar sensor** needs and challenges to auto-detect submerged **human bodies** underwater. Sonar images need to be tested at different levels of polarization and intensities, the target background must be considered as there will be scatterings and noises in the sonar image. The same model needs to be retrained on sonar images of various polarization and intensity.

[3] Self-Supervised Learning of Pretext-Invariant Representations is to construct image representations that are semantically meaningful via PIRL (Pretext Invariant Representation Learning) that do not require semantic annotations for a large training set of images. To achieve the highest single crop top-1 accuracy of all self-supervised learners that use a single **ResNet-50 model.**

[4] In this paper, the author has specified more on various techniques used for detecting sea mines. The major challenge of this paper was images need to be manually annotated and the overhead of labelling is very large when dealing with huge datasets. In this paper, they have used the **Mask RCNN model**
which Region-Based Convolutional Neural Network.

[5] in this paper, the methodology involves a Masked RCNN module that comprises numerous convolutional layers. In this referral paper, the dataset is a set of images downloaded from the web hence a lot of pre-processing, labeling, and augmentation is required.

[6] In this referral paper, **Gabor Filter and K-means clustering** algorithm is used. Gabor Filter is used for feature extraction, and the K-means Clustering algorithm is used for segmentation. The accuracy and efficiency of K-Means clustering are not accurate

[7] In this referral paper mainly two techniques are used, **Homomorphic, CLAHE, and Wavelet Filtering Techniques**

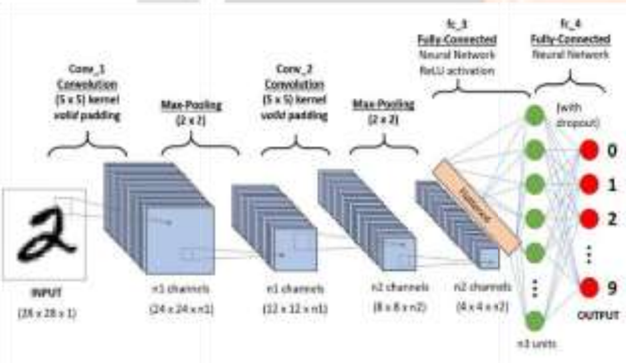
These techniques help to enhance the image quality and removing the unwanted noise enhances the image quality.
This survey paper also mentioned various other techniques for image enhancement such as a **median filter** for better quality and **RGB** for color level stretching

[8] In this referral paper they have specified image denoising techniques which include **Spatial Domain Filter, Frequency Domain Filter, Mean Filter, Median Filter, and Adaptive Filter.** The paper does not consider any hybrid filters which are more efficient in de-noising images and for a given dataset, the right kind of filter cannot be decided beforehand, we'll have to implement each of the filters on the dataset.

Title of Paper	**Authors of the paper**	**Model/Sub Model used**
A Review on Deep Learning-Based Approaches for Automatic Sonar Target Recognition	Dhiraj Neupane and Jongwon Seok	CNN, Autoencoders, Deep Belief Networks, GAN
Study on the Classification Performance of Underwater Sonar Image Classification Based on Convolutional Neural Networks for Detecting a Submerged	Huu-Thu Nguyen, Eon-Ho Lee 1 and Sejin Lee	AlexNet, GoogleNet

Human Body		
Self-Supervised Learning of Pretext-Invariant Representations	Ishan Misra, Laurens van der Maaten	PIRL (Pretext Invariant Representation Learning)

Underwater Fish Detection	Aditya Agarwal,Manonmani S,Gaurav Rawal,Tushar Malani,Navjeet Anand	Masked RCNN
Image Segmentation Using Gabor Filter and K-Means Clustering Method	Agyztia Premana, Akhmad Pandhu Wijaya, Moch Arief Soeleman	Gabor Filter and K-means clustering algorithm.



3. BACKGROUND

3.1 DEEP NEURAL NETWORK

A deep neural network is an artificial neural network (ANN) with several layers between the input and output layers (DNN) as shown in Fig.1. Neurons, synapses, weights, biases, and functions are all basic components of neural networks, which come in a range of forms and sizes.

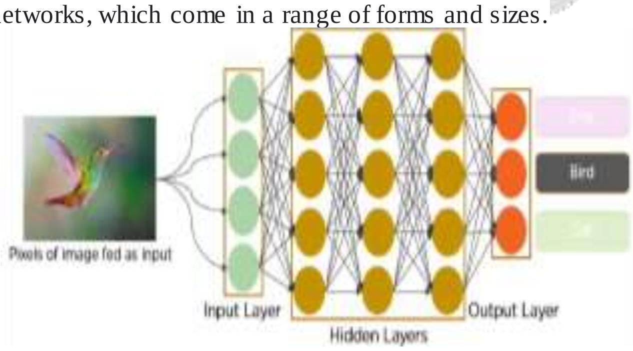


Fig.1 Deep Neural Network Architecture

3.2 CONVOLUTIONAL NEURAL NETWORK

The CNN is a sort of neural network as shown in Fig.2 that is quite similar to human vision and thought. Various computer vision applications have become an important part of this over time. CNN's were constructed for the first time in the 1980s. At the time, this neural network was the best at recognizing manual digits. The code reader has mostly been used or built to read zip codes, pin codes, and other codes of a similar kind.

Fig.2 A traditional CNN Architecture

3.3 GENERATIVE ADVERSARIAL NETWORK

A generative adversarial network (GAN) is a machine learning (ML) model in which two neural networks compete to improve the accuracy of their predictions as shown in Fig.1. GANs are frequently unsupervised and learn by playing a cooperative zero-sum game

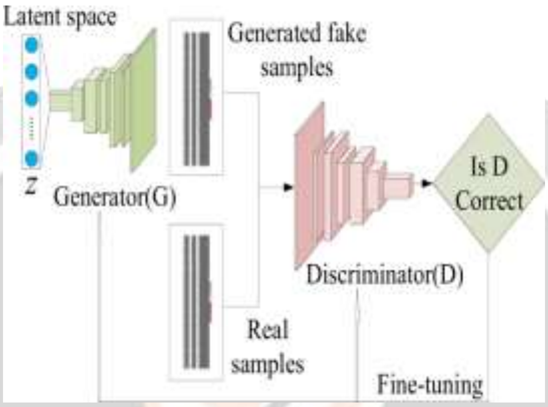


Fig.3 Generative Adversarial Network

4. IMPLEMENTATION OF DETECTION AND EXTRACTION OF NAVAL MINE FEATURE SYSTEM.

The implementation phase is an important step in the development of a project because the design needs to be carried out in order to solve the problem. During this step, low level/detailed designs will be transformed into language-specific programs to comply with the SRS document standards. The actual execution of concepts proposed in the design and analysis phase is part of this stage. The software implementation methodologies and processes must foster reusability, facilitate maintenance, and be generally understood. It is vital that the velocity and discipline of coding are maintained.

4.1 ARCHITECTURAL STRATEGIES

In this Section, the overall organization of the system along with its sub modules and detailed design is provided. It also provides a key insight into the system architecture in order to develop the application without any loopholes. The application should fit all the requirements as mentioned.

4.1.1 PROGRAMMING LANGUAGE SELECTION

Python 3 is chosen to implement the Machine Learning Model development pipeline and to develop the Backend of the Web Based Sea Mine Detection System. The main reasons for using Python 3 are listed below.

- Python is a free and open-source programming language.
- The Language is simple to use and consistent.
- Python has a vast set of libraries specifically and extensively developed for Machine Learning and associated disciplines. Popular examples include Keras, Tensorflow, Pandas, Scikit, OpenCV etc.

Python's extensive set web application frameworks can be extremely useful and timesaving, such as Flask, which has very less to no dependencies to external libraries, therefore light, flexible and easy to learn and implement.

Platform Independent :

HTML, CSS, and Bootstrap have been used to develop the Frontend. Bootstrap has been specifically chosen because there are fewer cross-browser bugs, the customizability, and the responsiveness of the structures and styles.

4.1.2 USER INTERFACE DESIGN

The user interface of the flask-based web application mainly consists of two pages, namely

1. **User Input Page**

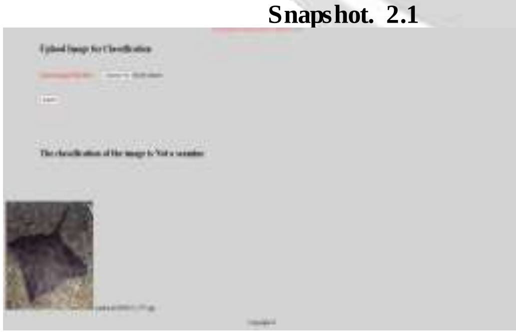
   The user interface of the user input page consists of two buttons, namely the select image button to select the image and then load button to upload the selected image. Once the user clicks on the Upload button, after selecting the image from local disk, the selected image is predicted based on the pretrained models which have been loaded into the flask app. The results of this prediction can be seen on the redirected page which is the final output page.

Snapshot. 1.1

2. **Output Page**

   This is the output page that shows us the prediction of each of the four different algorithms. It also consists of a classification report for each of the four CNN architectures. This classification report helps us understand if the predicted accuracy is reliable or the degree to which we can rely on the results of a particular CNN architecture.

Snapshot. 2.1



Snapshot. 2.2

4.1.3 TRAIN-TEST SPLIT

The dataset used for our study is a self-developed dataset by manually collecting images from various sources across the web. Although the original size of the dataset is very small, suitable augmentation techniques have been applied to increase the size of the dataset which also helps in better predicting the images

4.1.4 DATASTORAGE MANAGEMENT

The images and the models are the only data that have been used for the purpose of this study. The images which are provided as the input by the user are stored in the user's local directory. Images are also provided as the input for denoising/pre-processing techniques as well as the augmentation process. The output images are also stored in the local directory. The CNN architectures are trained, the models are saved first onto the Google drive and then loaded into the flask app. These loaded models are used to predict the input image uploaded by the user tothe web portal.

4.2 SYSTEM ARCHITECTURE

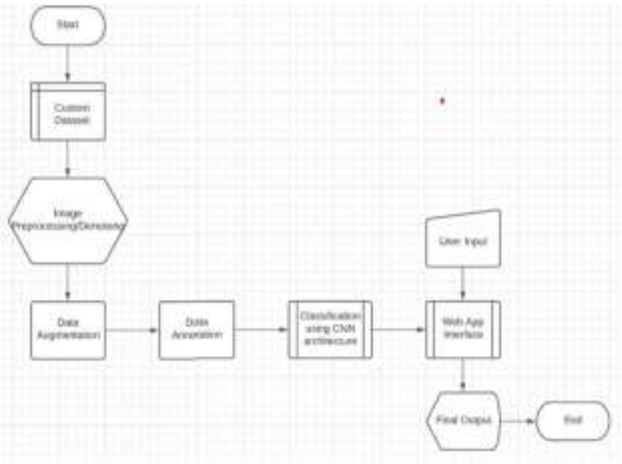


Fig.4 Naval Mines Detection System Architecture

First things first, we need to collect or prepare a good dataset that can feed to the machine learning models. But for this area of study, there are very few available images or data that can be directly taken as a dataset, so we have collected various images of the sea mines and have performed image pre-processing/denoising, and then have performed data augmentation by flipping the images vertically and horizontally. Thereby, preparing our own custom dataset. Now, this dataset undergoes automatic labelling using unsupervised techniques such as K-Means, Agglomerative, DBSCAN, and Birch. Now that we have the labelled data, it's time to feed the data to the training models i.e., performing the classification using CNN architecture using algorithms such as Resnet-50, VGG-16, Inception, and Xception and finallywe have a web application built for a comfortable user interface, where the user can inputthe image and the model gives out or basically identifies the image as sea mine or non-sea mineobjects as the final output.

4.3 Data Flow Diagrams

The Data Flow Diagram (DFD) shows the data flow through the whole system in a graphical way. Data Flow Models focus on the data flow between different sub modules through some processing steps. Data Flow Diagrams are composed of four primary components i.e. Processes, Data Flows, External Entities, and Internal Entities. Pictorial representation of the flow gives the user a clear picture of the working of the whole system. DFD's provide end- users with an idea about the processing of data and the effect it will have on the whole system when some data is entered as input to the system and ultimately the output of the system is obtained.

4.3.1 Data Flow Diagram - Level 0

The Level 0 DFD describes the overall operation of the system. It represents the system and user and the inputs and the outputs between the user and the system. The Level 0 DFD of the system is shown below

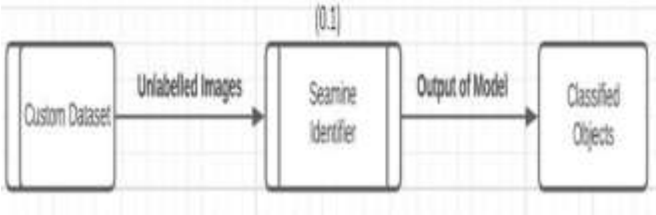


Fig. 5 Level 0 DFD for Naval mine Detection

4.3.2 Data Flow Diagram - Level 1

The Level 1 Data Flow Diagram breaks down the system into sub-modules specifically. The Level 1 DFD of the system is shown below.

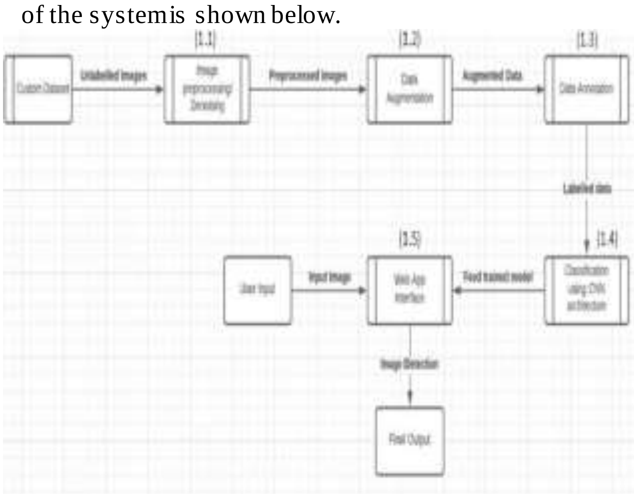


Fig.6 Level 1 DFD for Underwater Naval Mine Detection System

4.3.3 Data Flow Diagram – Level

The processes in Level 1 DFD are expanded in Level 2 DFD. Level 2 DFD gives a detailed description of the system as shown in the figures below, the figure below expands on the process of Image Pre-processing specifically Denoising.

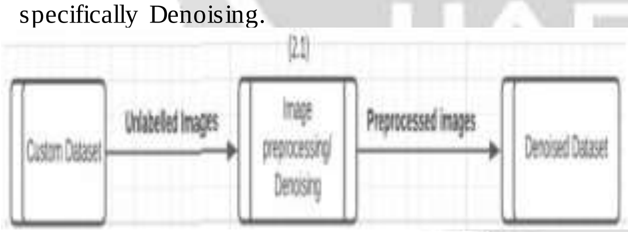


Fig.7 Level 2 DFD for Underwater Naval Mine Detection System

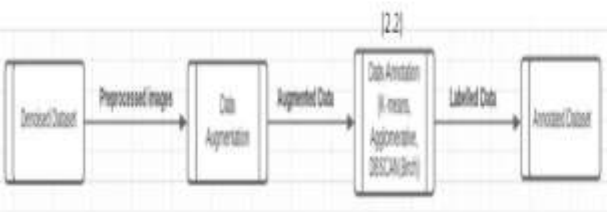


Fig.8 Level 2 DFD for Underwater Naval Mine Detection System

The figure below expands on the process of training CNN models and these pretrained models on the webapp detect sea mines from user input images.

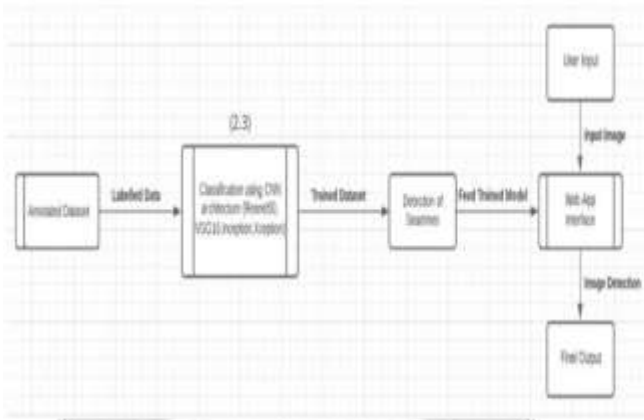


Fig.8 Level 2 DFD for Underwater Naval Mine Detection System

5. Functional Description of Modules

Fig. 9 Structure Chart for Sea Mine Detection

This section describes the internal functioning of each module. Alsocovered are software components and subcomponents.

5.1 Image Denoising Module

This module performs the function of denoising input images.

- **Purpose:** This module is intended just to denote the photos in the selected dataset. This is the first stage to identify marine mines since it is essential to remove the noise from photos.
- **Input:** The input provided to this module will be the Custom Underwater Image dataset.
- **Output**: The output of this module will be the set of corresponding denoised images. It will be stored in a specific destination folder which is internally provided.
- **Functionality:** The functionality of this module is to denoise the input image dataset fed to it and produce corresponding denoised images dataset that is used in the later stages of the system.

5.2 Data Augmentation Modul

This module is responsible for increasing the size of the image dataset.

Purpose: In order to address the shortfall by increasing the number of images in this module.

- **Input:** The input provided to this module will be the denoised image dataset.
- **Output**: The output of this module will be the expanded image dataset achieved throughaugmentation.

- **Functionality:** The functionality of this module is to apply image augmentation on each input image to generate augmented equivalents to expand the existing image dataset. The resultant expanded dataset is stored in a specified directory.

5.3 Data Labelling Module

This module is responsible for labeling images in the image dataset as Sea Mine or No Sea Mine.

Purpose: This module is intended to label the picture dataset.

- **Input:** The input provided to this module will be the Augmented dataset of images.
- **Output**: The output of this module will be the images labeled and categorized.

6. CONCLUSION

The detection of the sea mines can be done in real-time or in real-time in order to be more effective in its usage to the naval forces of the country. A more extensive dataset that involves actual SONAR-generated images and contains a wide selection of mines of all shapes, sizes, and specifications, images of different visibilities, etc can be used to train a more comprehensive detection model that detects accurately across all these variations. Models may be trained using different machine-learning techniques like YONO v3 and a few more recent approaches which can lead to improved performance and outcomes.

7. REFERENCES

[1] Dhiraj Neupane and Jongwon Seok, A Review on Deep Learning-Based Approaches for Automatic Sonar Target Recognition, MDPI Electronics, 2020.

[2] Huu-Thu Nguyen, Eon-Ho Lee 1 and Sejin Lee, Study on the Classification Performance of Underwater Sonar Image Classification Based on Convolutional Neural Networks for Detecting a Submerged Human Body, MDPI Sensors, 2019.

[3] Ishan Misra, Laurens van der Maaten, Self-Supervised Learning of Pretext-Invariant Representations, arXiV, 2019.

[4] N Abhishek, Arjun, Bharathesh, Kavitha KS, Prof. Manonmani S, Dr. Shanta Rangaswamy, “Underwater Mine Detection using Image Processing”. 2020, International Research Journal of Engineering and Technology (IRJET).

[5] Aditya Agarwal, Manonmani S, Gaurav Rawal, Tushar Malani, Navjeet Anand, “Underwater Fish Detection”, 2020, International Journal of Engineering Research & Technology (IJERT).

[6] Agyztia Premana, Akhmad Pandu Wijaya, Moch Arief Soeleman, “Image Segmentation Using Gabor Filter Clustering Method”, 2017 International Seminar on Application for Technology of Information and Communication.

[7] Manonmani S, Dr. Shanta Rangaswamy,” Comparative analysis of combining various enhancement filtering techniques for images”, 2018, IEEE International Conference on Control, Power, Communication and Computing Technologies.

[8] Manonmani, Lalitha V., Dr. Shanta Rangaswamy. “Survey On Image Denoising Techniques”, 2016, International Journal of Science, Engineering and Technology Research (IJSETR).