Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 6 (2023)

Hybrid Deep Learning Model for Detecting Underwater Naval Mines

Manonmani Srinivasan[1*], Shanta Rangaswamy[2]
[1, 2] *Rashtreeya Vidyalaya College of Engineering, Bengaluru, Karnataka, India*
Visvesvaraya Technological University, Belagavi, Karnataka, India

Abstract:- In the realm of maritime warfare, naval mines stand as formidable threats, strategically positioned explosives lurking beneath the water's surface, poised to inflict damage upon unsuspecting ships or submarines. This proposed model harnesses the power of deep learning methodologies to discern and identify these submerged hazards. Leveraging contemporary advancements in deep learning technology, our aim is to construct robust and economical models capable of reliably detecting naval mines. Through this study, an array of deep learning models is employed to gauge their efficacy, utilizing accuracy as a primary metric for comparison and evaluation.

Among the models employed were CNN (Convolutional Neural Network), YOLOv5, VGG-19, and a hybrid fusion of CNN-VGG-19 and CNN-MobileNet. Remarkably, CNN showcased outstanding performance, achieving an impressive accuracy rate of 98%. Additionally, YOLOv5 demonstrated robust performance, closely trailing behind with an accuracy score of 97%.Surpassing them all, the hybrid models, specifically the CNN-VGG19 and CNN-MobileNet fusion, showcased the highest accuracy, reaching an outstanding 99%.

Keywords*:* YOLO, CNN, TensorFlow, Python, VGG-19, Mobilenet.

1. Introduction

Throughout history, conflicts among nations, groups, and regions have spurred the development of defense mechanisms to resolve disputes and safeguard territorial integrity. Ground warfare has witnessed the utilization of grenades and artillery, while torpedoes and air-launched missiles have been deployed by nations worldwide. Similarly, maritime arenas have seen the employment of submarines, seaplanes, and naval mines in defense strategies. Presently, these techniques are commonplace across nations for defense purposes. Naval mines, a type of defensive arsenal, have a long history dating back to the 14th century, evolving from early precursors into modern forms. Functioning as self-contained explosive devices, naval mines serve both offensive and defensive purposes, obstructing hostile vessel movements and safeguarding allied ships, thereby establishing safer maritime zones.

A naval mine serves as an underwater explosive device designed to target and damage ships or submarines. Functioning autonomously, it can be deployed for offensive or defensive purposes, safeguarding friendly vessels and delineating secure maritime zones. These mines are strategically placed in ocean waters, and upon contact with a passing vessel, such as a ship or submarine, they detonate, inflicting damage. The presence of sea mines compels opposing forces to navigate cautiously, as they must choose between undertaking resource-intensive minesweeping operations, risking casualties by challenging the minefield, or diverting to less fortified waters where enemy forces may concentrate. The process of minesweeping entails the removal of naval mines, usually conducted through specialized vessels or methodologies designed specifically for capturing mines.

In contrast to early gunpowder-based mines that relied on manual ignition, modern mines equipped with sophisticated electronic fuses and high explosives are considerably more potent. Deployment of sea mines can be executed via boats, aircraft, submarines, or even individual swimmers and boaters, offering a diverse range of tactics. While international regulations mandate the marking and declaration of mined areas, the precise locations of these mines are typically classified. Nonetheless, the presence of mines poses significant challenges to shipping and trade, persisting long after conflicts cease unless measures are taken to mitigate their longevity. Therefore, prompt detection and removal of naval mines post-war are imperative. A range of methodologies is utilized with the goal of safely removing these hazards from maritime environments.

"Deep Neural Networks" encompass Artificial Neural Networks (ANN) comprising multiple layers, enabling them to handle vast datasets. Widely embraced by Data Scientists across various domains, these networks owe their name to the mathematical operations, termed convolutions, conducted between matrices. CNN, a prominent variant, comprises distinct layers including convolutional, nonlinearity, pooling, and fully-connected layers. Compared to standard feed-forward neural networks with similar-sized layers, CNNs boast significantly fewer connections and parameters, rendering them faster and simpler to train. Leveraged in diverse applications like image classification datasets (e.g., ImageNet), computer vision, anomaly detection, and NLP, CNNs exhibit versatility. In this investigation, the current model is specifically trained for the detection of underwater naval mines, employing deep learning techniques to accurately identify and classify naval mine images.

2. Literature Review

The authors of this paper [1] conducted a review of the underwater image data acquisition process and operational procedures concerning naval mines. Furthermore, they extensively explored a variety of deep learning models for identifying, tracking, and detecting objects within sonar data.

In this paper [2], the authors employed the FRCNN (Fast Region Convolutional Neural Network) model to distinguish objects as either mines or non-mines. They utilized a cloud platform to oversee mine detection, ensuring real-time monitoring. Any alterations detected were promptly reflected in an Android application for immediate observation and response.

In their paper [3], the authors provided a comprehensive overview of techniques crucial for detecting and classifying underwater mines in side-scan sonar imagery. Their thorough examination encompassed more than 30 research papers, concentrating on image processing tools and methodologies aimed at object detection and classification.

In their paper [4], the authors addressed the issue of limited data availability in underwater image processing by emphasizing synthetic dataset generation. They employed deep learning methodologies to create a model capable of autonomously detecting and classifying underwater mine images. These images, generated by Synthetic Aperture Sonar (SAS), offer a solution to the demand for automated detection and classification of underwater mines.

In this paper [5], the authors conducted a thorough review of image enhancement techniques aimed at improving the quality of underwater images. They specifically explored methods such as Color Stretching, USM filters, and others. Their focus encompassed a comparative analysis of various image enhancement techniques as well as noise removal methods, all geared towards enhancing image quality to facilitate better prediction models.

In this research [6], the authors introduced a method capable of automatically labeling image datasets and determining the presence of naval mines within them. They conducted investigations using four distinct CNN architectures: ResNet50, VGG-16, InceptionV3, and Xception. The accuracy rates for these architectures were found to be 77%, 68%, 82%, and 86%, respectively. While Xception exhibited the highest accuracy during training, it displayed signs of overfitting on the test dataset. Consequently, the InceptionV3 algorithm emerged as the most effective for naval mine detection. ResNet50 also performed well, particularly following InceptionV3. Conversely, VGG-16 demonstrated poorer performance with a 68% accuracy rate.

In this study [7], Feature Extraction techniques like Histogram of Oriented Gradient (HOG) and Canny edge detection were utilized for underwater mine detection. Prior to applying the detection algorithms, the data underwent preprocessing to enhance results. The accuracy rates achieved were 95.83% for HOG and 94.44% for Canny edge detection.

In this paper [8], the authors provided insights into various naval mines utilized by different countries, including Brazil's MCF-100, Spain's Mila-6B Sea Mine, Iraq's SIGEEL/400 and MDM series, Germany's SM G2 mine, United States MK-52, and Italian Manta Mine. Additionally, the paper explored the significance of explosives in detonation, discussing types such as HBX-1, H-6, TNT, and HBX-3.

In this paper [9], the authors conducted a comparative analysis of various techniques aimed at enhancing the quality of images. They evaluated the efficacy of these techniques using metrics such as SNR, PSNR, MSE, and SSIM. The simulation results indicated that the CLAHE filter outperformed others in terms of SNR and PSNR, while Homomorphic and Wavelet techniques yielded similar results in terms of MSE and SSIM. The study also included graphical representations, generated using GNU plot, depicting the metrics SNR, PSNR, MSE, and SSIM for ten randomly selected images.

In their study [10], the authors presented an innovative model leveraging deep CNN-based feature extraction with reduced parameters, facilitating swift object identification in just 28 milliseconds. Their methodology focused on categorizing vehicle images into five distinct groups. By employing the AdaBoost algorithm alongside deep Convolutional Neural Networks (CNNs), the proposed approach achieved an impressive classification accuracy of 99.5% on the test dataset. Compared to traditional algorithms, this model showcased superior performance. Moreover, its minimized parameter count translates to significantly reduced storage requirements compared to other CNN models.

In the paper [11], a comprehensive examination of recent advancements in deep learning models for underwater image analysis is provided. The authors systematically categorized various techniques based on feature extraction methods, deep learning architectures, and object detection approaches. The analysis was structured around specific objects targeted for detection, highlighting the features and deep learning architectures utilized. The study concluded that substantial potential exists for automating the analysis of digital seabed images using deep convolutional neural networks, particularly in the detection and monitoring of seagrass.

In their paper [12], the authors conducted an analysis of the technical challenges encountered in underwater target recognition methods employing Autonomous Underwater Vehicles (AUVs). The article extensively discusses various deep learning approaches for analyzing underwater images and briefly outlines the fundamental principles underpinning different underwater target recognition methods.

In this paper [13], the authors investigated recent techniques employed for underwater object detection, conducting a thorough and comprehensive comparative analysis.

In their paper [14], the authors conducted a review of recent advancements in underwater marine object detection, elucidating the strengths and weaknesses of current solutions for each challenge. They meticulously analyzed the most commonly used benchmark datasets, offering critical insights. The paper also presented comparative studies with prior reviews, particularly focusing on artificial intelligence-based approaches, and discussed future trends in this dynamic field.

In this study [15], the authors trained and tested six different deep-learning CNN detectors for object detection. Five of these detectors were based on the You Only Look Once (YOLO) architectures (YOLOv4, YOLOv4-Tiny, CSP-YOLOv4, YOLOv4@Resnet, YOLOv4@DenseNet), while one utilized the Faster Region-based CNN (RCNN) architecture. Evaluation metrics such as detection accuracy, mean average precision (mAP), and processing speed were employed to assess the model's performance on a custom dataset comprising underwater pipeline images. The findings revealed that YOLOv4 outperformed other models in underwater pipeline object detection, achieving an mAP of 94.21% and demonstrating the capability to detect objects in real-time.

In their paper [16], the authors introduced the Underwater-YCC optimization algorithm, which is based on You Only Look Once (YOLO) v7, aiming to enhance the accuracy of detecting small targets underwater. Their proposed algorithm incorporates the Convolutional Block Attention Module (CBAM) to extract fine-grained semantic information, determined through multiple experiments to optimize positioning. Additionally, they integrated the Conv2Former as the Neck component to address underwater blurred images effectively. Finally, they implemented Wise-IoU, a technique proven to enhance detection accuracy by assigning varying weights between high- and low-quality images. Experimentation on the URPC2020 dataset revealed that the Underwater-YCC algorithm achieves a mean Average Precision (mAP) of up to 87.16% in complex underwater environments.

In this study [17], a novel model for small object detection using neural network architecture was introduced. Named Sample-Weighted Hyper Network (SWIPENet), the proposed method leverages the Invert Multi-Class Adaboost (IMA) algorithm for noise removal. Demonstrating superior performance compared to several state-of-the-art object detection approaches, SWIPENet presents a promising solution for accurate small object detection.

In this paper [18], a thorough review of deep-learning-based object recognition for both surface and underwater targets was conducted. To ensure a comprehensive overview, the authors initially summarized key concepts and typical architectures within a unified framework. They meticulously gathered popular benchmark datasets for marine object recognition and provided a comprehensive analysis of deep learning methodologies through extensive comparisons. Additionally, the paper delved into experimental results and discussed future trends in marine object recognition in depth.

In this paper [19], the authors conducted an in-depth survey on several deep learning models, including Fast RCNN, Faster RCNN, and the original YOLOV3, to ascertain their effectiveness in detecting underwater objects. Achieving a detection speed of approximately 50 FPS (Frames per Second) and a mean Average Precision (mAP) of about 90%, the study demonstrated promising results. The implemented program was integrated into an underwater robot, and real-time detection results indicated accurate and swift detection and classification capabilities, facilitating efficient underwater operations for the robot.

In this paper [20], a thorough examination of underwater object detection techniques was presented, encompassing current research challenges, future development trajectories, and potential applications. The study delved into the interconnection between underwater image enhancement and object detection, scrutinizing potential implementation strategies for integrating underwater image enhancement into object detection tasks to amplify their effectiveness.

3. The Proposed System Architecture

The system architecture is structured according to the order of execution requirements, as depicted in Fig-1. This design fosters a highly modular structure, comprising augmentation, annotation, model training, and evaluation metrics stages.

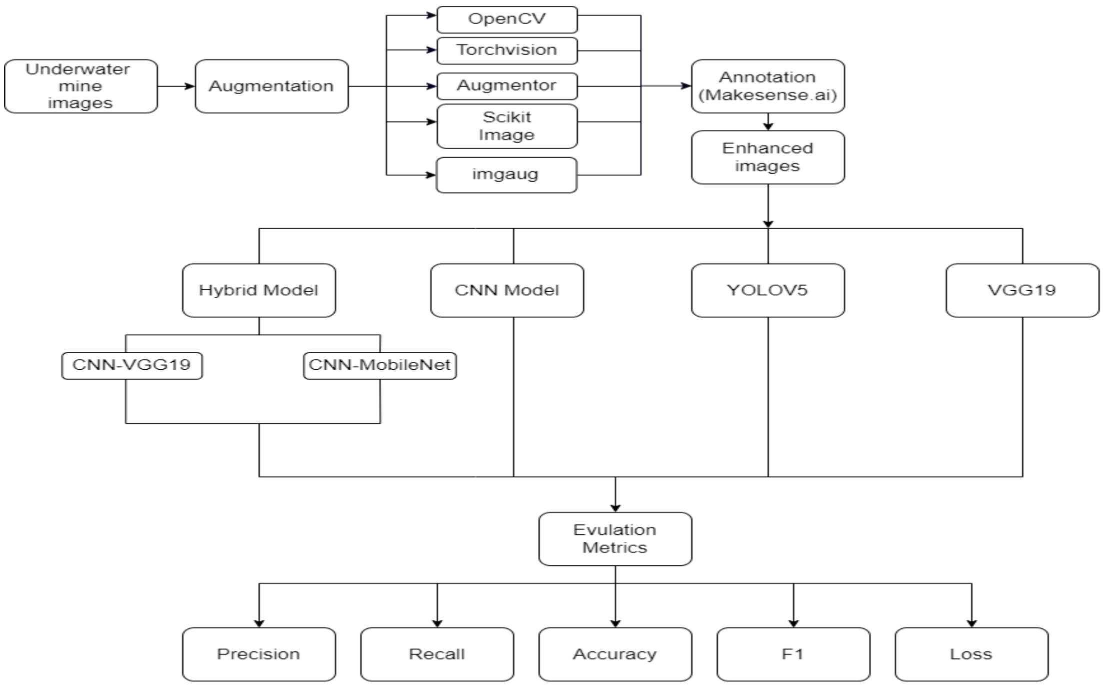


Fig-1 The Proposed System Architecture

As illustrated in Fig.1, the implementation process initiates by acquiring Underwater Naval mine images from various online sources, crucial for enhancing the accuracy of neural network models. Image augmentation, a

ISSN: 1001-4055
Vol. 44 No. 6 (2023)

widely-used technique, is employed to expand the dataset, utilizing augmentation libraries such as Augmentor, OpenCV, Scikit-image, and imgaug.

Next, the Annotation module utilizes the makessense.ai tool to annotate the image dataset with bounding boxes and export the annotations into YOLO format for YOLOv5 model training. Subsequently, a program is utilized to divide the dataset into training and validation sets with an 80:20 ratio. Further augmentation, scaling, and resizing of images are performed in the Enhanced Images module to ensure uniformity in size for model training.

For models like CNN, VGG-19, and hybrid models CNN-MobileNet and CNN-VGG19, the dataset does not undergo annotation, as the CNN model does not predict bounding boxes. However, for the YOLOv5 model, the mine dataset is both enhanced and annotated using the makesense.ai tool to enable bounding box detection and multiple classification.

The Evaluation Metrics module assesses various metrics such as Loss, F1-score, Precision, Accuracy, and Recall for each model to determine the best model for detecting and classifying mines. The proposed system architecture provides a detailed representation of these processes.

Additionally, sample input and output images for the Scikit-image, Augmentor, Imgaug, OpenCV, and Torchvision packages are depicted in figures 2, 3, 4, 5, and 6 respectively.

Sample input & output for Image Augmentation

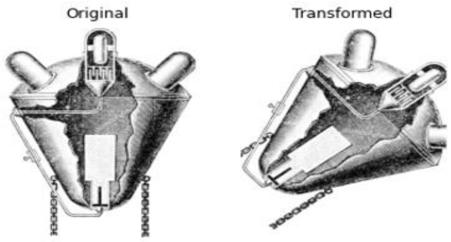


Fig-2.Sckit-image augmentation

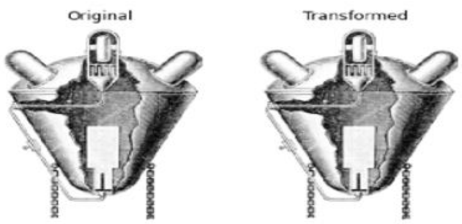


Fig-3 Augmentor augmentation

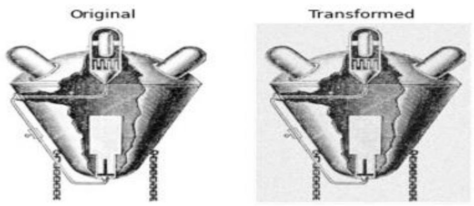


Fig-4 Imgaug augmentation

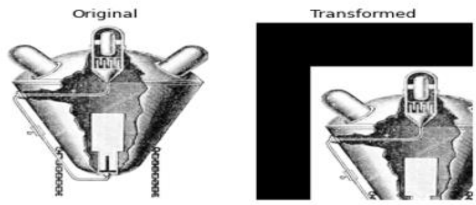


Fig-5OpenCV augmentation

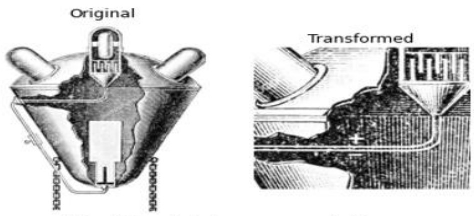


Fig-6 Torchvision augmentation

In this proposed model architecture, we conducted a comparative analysis of four different neural networks: CNN, VGG19, YOLOV5, and Hybrid models. Our objective was to evaluate the performance of these models based on various evaluation metrics, including accuracy and mean average precision (mAP).

The mAP is a crucial metric calculated by determining the Average Precision (AP) for each class and then averaging over all classes. It accounts for the trade-off between precision and recall while considering false positives (FP) and false negatives (FN). This characteristic makes mAP a suitable metric for a wide range of detection applications.

Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 44 No. 6 (2023)

To compare the training profiles of the test cases and analyze the metrics obtained, we calculated them from the generated confusion matrix using the following equations (1, 2, 3, 4, and 5):

$$Precision = \frac{TP}{TP + FP} \quad ----------------------------- Eq\ 1$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad ------------------------ Eq\ 2$$

$$Recall = \frac{TP}{TP + FN} \quad ----------------------------- Eq\ 3$$

$$Specificity = \frac{TN}{TN + FP} \quad ---------------------------- Eq\ 4$$

$$F1\ score = \frac{2 * Precision * Recall}{Precision + Recall} \quad ------------------------ Eq\ 5$$

Deep Learning Models

In this implementation, we utilized three distinct CNN models—CNN, VGG-19, and MobileNet. We gathered a comprehensive dataset comprising 13 different types of naval mine images sourced from various online platforms, including both authentic and synthetic underwater naval mine images generated through augmentation libraries. The dataset encompasses a total of 3441 photos, representing the diverse range of mine types. Of these, 2941 data points were earmarked for training the models, while the remaining 668 were set aside for testing purposes. The dataset was meticulously partitioned into an 80:20 test-train split, with an additional control dataset included for validation.

For the VGG-19 and MobileNet algorithms, we leveraged weights sourced from the ImageNet database, incorporating pre-trained models that have been extensively trained on millions of images. These pre-trained weights were seamlessly integrated into the models during the training process, enhancing their performance and robustness. A batch size of thirty-two was deemed suitable for training the dataset, with the training process spanning thirty epochs. The categorical cross-entropy loss function was employed, while the Adam optimizer was utilized to address noise issues. The rectified linear unit (relu) activation function was applied in the hidden layers. To prevent overfitting, dropout layers with a rate of 0.5 were added. Additionally, the outputs were normalized using the Softmax activation algorithm. Each model underwent training and was subsequently saved.

The CNN Model's layer configuration is depicted in Fig-7, while the overall CNN architecture is illustrated in Fig-8.

```
#CNN Model

def create_model():
    model = Sequential([
        Conv2D(filters=128, kernel_size=(5, 5), padding='valid', input_shape=(IMG_WIDTH, IMG_HEIGHT, 3)),
        Activation('relu'),
        MaxPooling2D(pool_size=(2, 2)),
        BatchNormalization(),

        Conv2D(filters=64, kernel_size=(3, 3), padding='valid', kernel_regularizer=l2(0.00005)),
        Activation('relu'),
        MaxPooling2D(pool_size=(2, 2)),
        BatchNormalization(),
         Conv2D(filters=32, kernel_size=(3, 3), padding='valid', kernel_regularizer=l2(0.00005)),
        Activation('relu'),
        MaxPooling2D(pool_size=(2, 2)),
        BatchNormalization(),

        Flatten(),

        Dense(units=256, activation='relu'),
        Dropout(0.5),
        Dense(units=13, activation='softmax')
    ])

    return model

cnn_model = create_model()
```

Fig-7 Proposed CNN Model summary

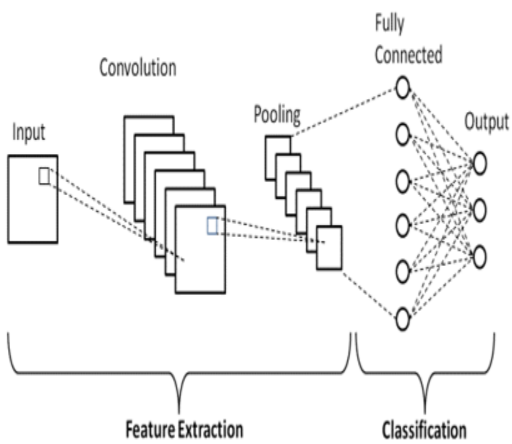


Fig-8 CNN Model Architecture

The CNN model is structured with three primary layers: convolutional layers, pooling layers, and fully-connected layers. Alongside these core layers, two additional parameters—dropout and activation functions—are employed to enhance the CNN's performance. Table 1 illustrates the functionality of each layer within the CNN architecture.

Table 1: CNN layers and its Functionality

Sl.No	CNN Layers	Functions
1	Convolutional Layers	Feature Extraction (mathematical operation of convolution is performed)
2	Pooling Layers	Decrease the size of the convolved feature map to reduce the computational cost
3	Fully-Connected Layers	Connect the neurons between two different layers.
4	Droupout	Dropping few neurons to avoid overfitting in the training dataset
5	Activation Functions	Decides which information should fire in forward direction and also adds non-linearity to the network.

VGG19

VGG19 stands as a deep convolutional neural network featuring 19 layers in its architecture. To adapt VGG19 to our dataset, the existing pretrained weights are initialized as false, enabling training with our specific data. Figure 9 illustrates the model architecture of VGG19, while Table 2 delineates the functionality of each layer within the network.

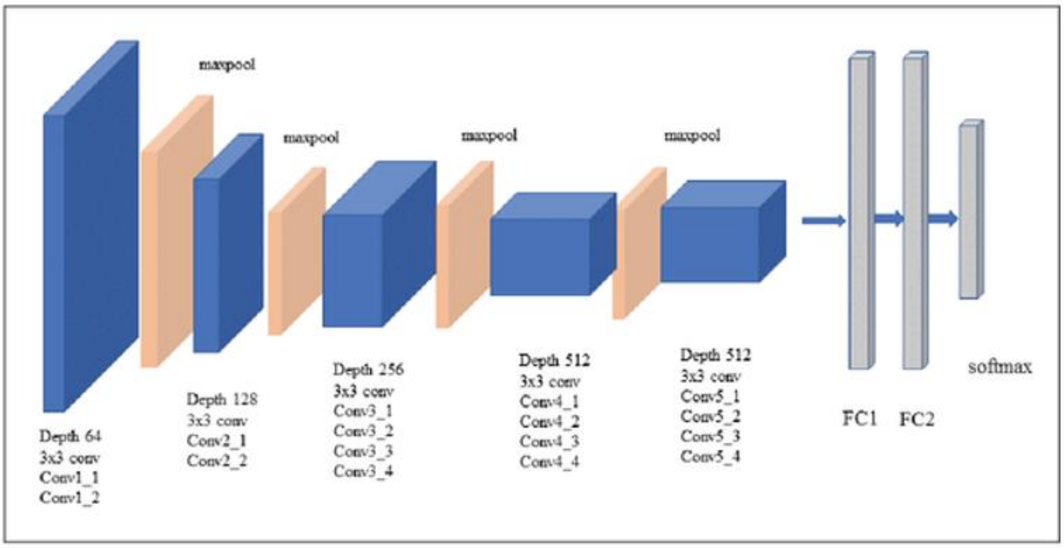


Fig-9 The Model Architecture of VGG19

Table 2: VGG19 layers and its Functionality

Sl.No	Functions	Descriptions
1	Input	Matrix of shape (224,224,3)-A fixed size of (224*224) RGB Image
2	Preprocessing	Subtracted the mean RGB value from each pixel, computed from the overall training set.
3	Kernels	3*3 size with stride size of 1 pixel
4	Padding	Spatial padding to preserve the spatial resolution
5	Pooling	Max pooling over a 2*2 pixels windows with stride 2
6	ReLu	To introduce non-linearity to make the model classify better and improve computational time as the previous model

Sl.No	Functions	Descriptions
7	Fully Connected layers	The first two layers of size 4096 and with 1000 channels for 1000-way ILSVRC classification
8	Final Layer	Softmax functions

CNN-VGG19:

The hybrid CNN and VGG-19 model were trained using the underwater naval mine dataset. Initially, the VGG-19 model underwent training with the dataset. Subsequently, the trained VGG-19 model was integrated as a base layer into the underlying CNN model. In Figure 10, the functional layer of Model 2 is depicted as added to the base CNN model.

MobileNet:

MobileNet employs depth-wise separable convolution as its core building block. Figure 11 illustrates the two layers of depth-wise separable convolution: depth-wise convolution and point convolution.

CNN-MobileNet:

Similarly, the hybrid CNN and MobileNet model were trained on the underwater naval mine dataset. Here, the MobileNet layers serve as additional layers to the base CNN model. Notably, MobileNet exhibited favorable performance in terms of accuracy and loss, achieving a test accuracy of 99%. The summary of the CNN-MobileNet hybrid model is presented in Figure 12.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 model_2 (Functional)        (None, 13)                20130893

 flatten_3 (Flatten)         (None, 13)                0

 dense_3 (Dense)             (None, 128)               1792

 dropout (Dropout)           (None, 128)               0

 dense_4 (Dense)             (None, 13)                1677

=================================================================
Total params: 20,134,362
Trainable params: 109,978
Non-trainable params: 20,024,384
```

Fig-10 CNN-VGG-19 Model Summary

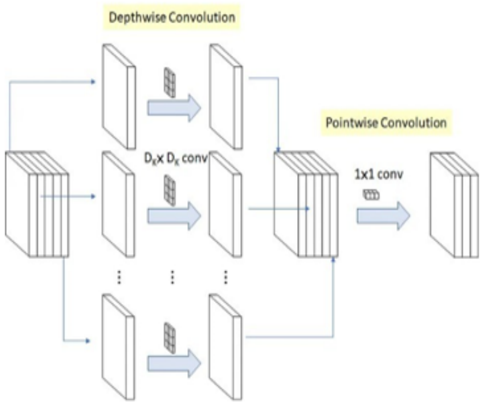


Fig-11 MobileNet Architecture

```
Model: "sequential_1"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 batch_normalization (BatchN  (None, 224, 224, 3)      12
 ormalization)

 mobilenet_1.00_224 (Functio  (None, 7, 7, 1024)       3228864
 nal)

 batch_normalization_1 (Batc  (None, 7, 7, 1024)       4096
 hNormalization)

 global_average_pooling2d_1   (None, 1024)             0
 (GlobalAveragePooling2D)

 dropout_3 (Dropout)         (None, 1024)              0

 dense_8 (Dense)             (None, 13)                13325

=================================================================
Total params: 3,246,297
Trainable params: 15,379
Non-trainable params: 3,230,918
_________________________________________________________________
```

Fig-12 CNN-MobileNet Model Summary

YOLOv5:

For YOLOv5, the underwater naval mine dataset underwent annotation using the makesense.ai tool, an open-source platform designed for image annotation. Subsequently, the yolov5 model was trained using 173 annotated images, with 143 allocated for training and 30 for validation purposes. The architectural diagram of YOLOv5 is illustrated in Figure 13.

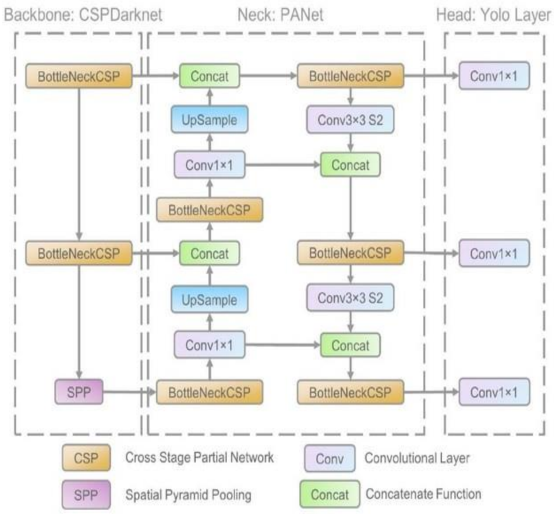


Fig-13 Yolov5 Model Architecture

The programming language utilized in this study is Python 3.0, along with a suite of tools and libraries including MakeSense.ai, Scikit-image, Labelimg, Keras, various augmentation libraries, Panda Library, TensorFlow 2.9, and Scikit-learn for evaluation metrics.

Dataset:

The dataset utilized in this study comprises a curated collection of images sourced from diverse online platforms. It encompasses a total of 3441 images depicting underwater naval mines, with 2941 images allocated for training and 668 images for testing. These images represent 13 distinct classes of naval mines.

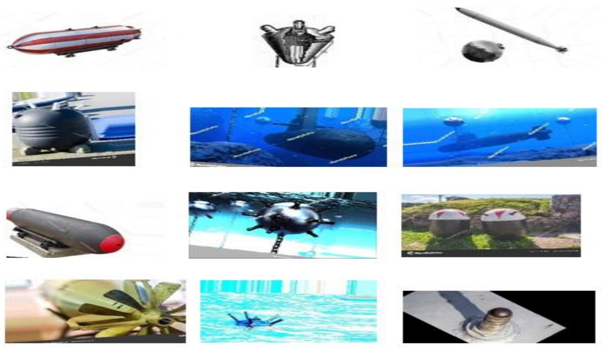


Fig-14 Naval Mines Dataset

4. Result and Analysis

CNN Binary Classification Model

This CNN model serves as a binary classifier, tasked with determining whether an input image contains a naval mine or not. With an achieved accuracy of 60%, it successfully categorizes images into the two classes. A comprehensive summary of the CNN Binary classification Model's performance metrics, including accuracy, the confusion matrix, sample output, accuracy and loss graph, as well as precision, recall, and F1-score, are presented in Table 3, Figure 15, Figure 16, Figure 17, and Table 4, respectively.

Table 3: Accuracy Results of binary classification

Algorithm	Accuracy Score
CNN	60%

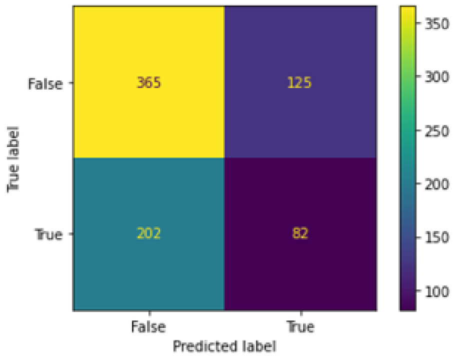


Fig-15 Confusion matrix for CNN binary classification

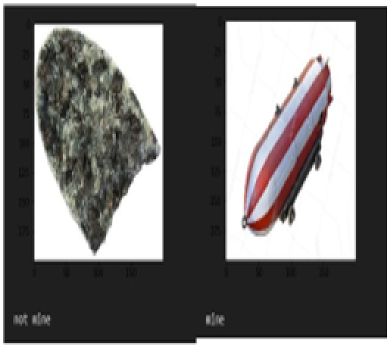


Fig-16 CNN Binary Classification sample input and output

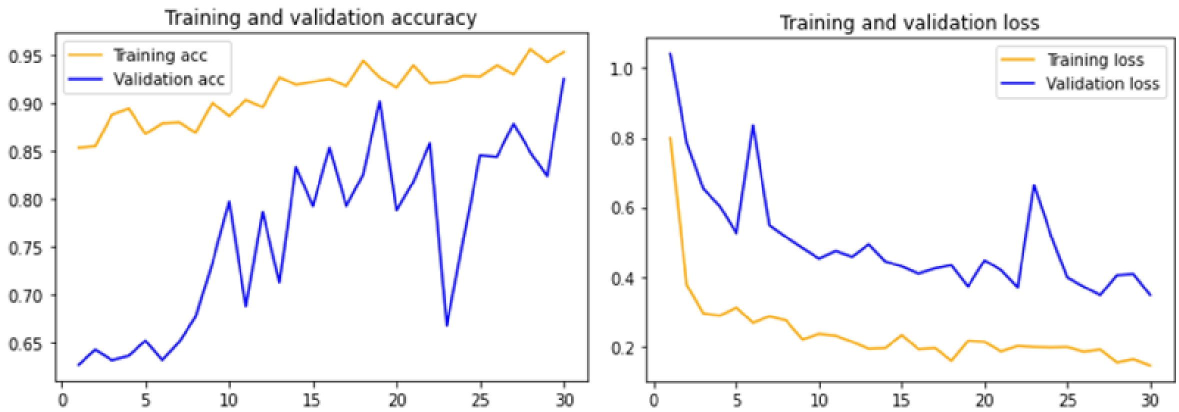


Fig.17. Model Accuracy & Loss Graph for CNN Binary classification

Table 4 CNN Precision, recall and f1-score for binary classification

Algorithm	Label	Accuracy metric	Results
CNN	Mine	precision	0.552
		recall	0.577
		F1-score	0.559
	Not mine	precision	0.552
		recall	0.577
		F1-score	0.559

CNN Multi Class Classification Model

This multi-class CNN classification model is designed to process input images, detect mines, and classify them among 13 distinct classes. With an impressive accuracy score of 98%, the model effectively categorizes images into the specified classes. Figure 18 illustrates the Confusion Matrix of the CNN Multi-classification model, while sample outputs, including correct and incorrect predictions, are depicted in Figure 19 and Figure 20, respectively. Additionally, the accuracy and loss graph of the model are presented in Figure 21.

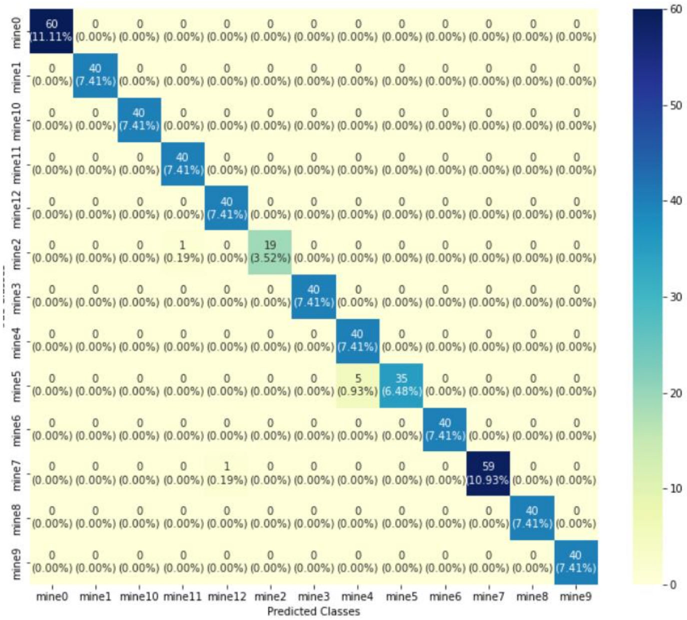


Fig-18 Confusion matrix for CNN multi class classification

Fig.19 CNN Accurate Predictions on the Test Dataset

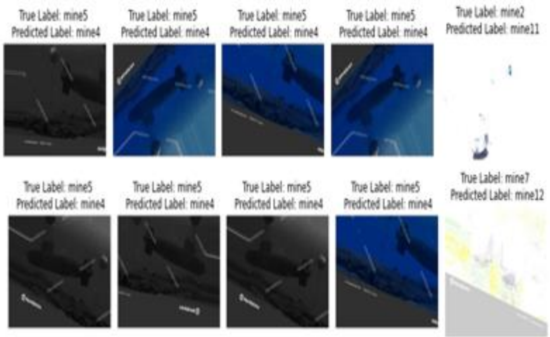


Fig.20 CNN Wrong predictions on test dataset

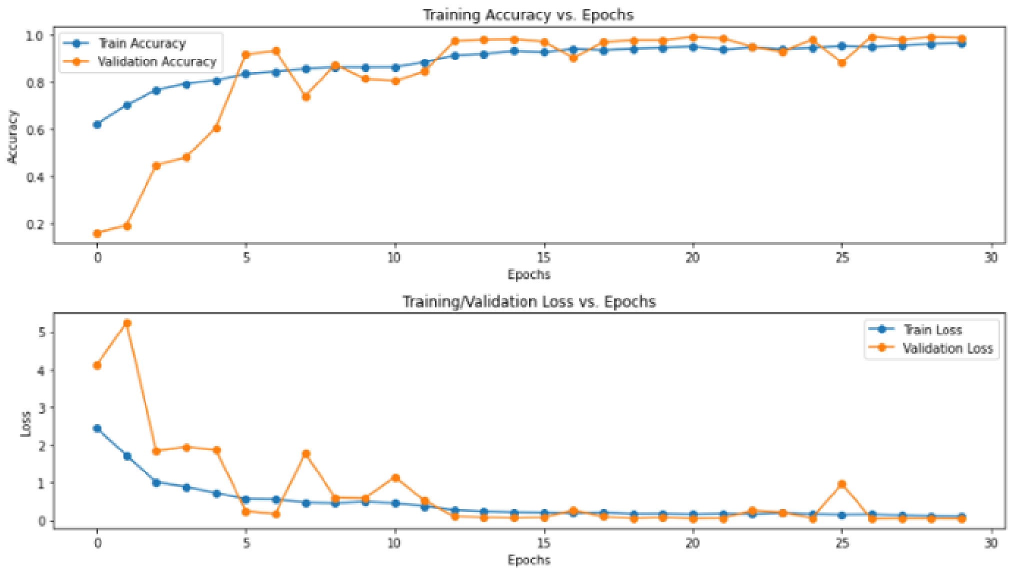


Fig.21. Model Accuracy & Loss Graph for CNN Multi class classification

VGG-19 Model

The VGG-19 model operates by analyzing input images to detect and classify mines among the 13 predefined classes, achieving an impressive accuracy score of 98%. Detailed insights into the VGG-19 model's performance are provided through its Confusion Matrix, sample outputs, including correct and incorrect predictions, and the accuracy and loss graph, which are respectively showcased in Figure 22, Figure 23, Figure 24, and Figure 25.

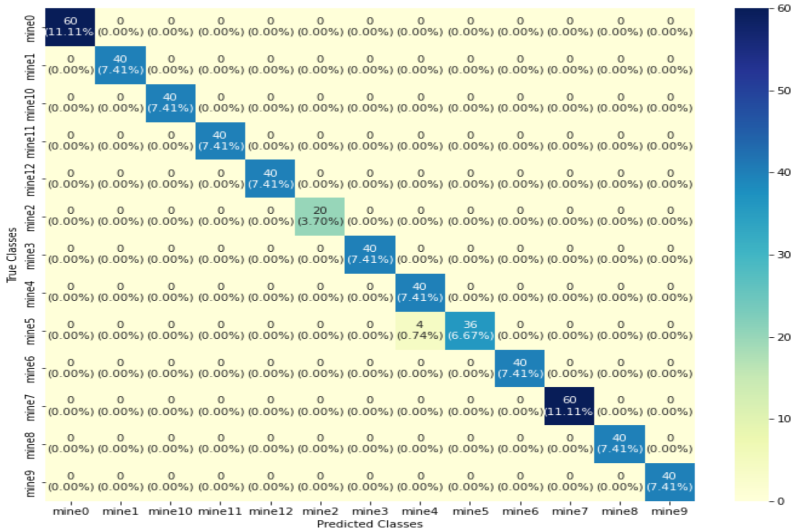


Fig-22 VGG-19 Confusion Matrix for multiple classification

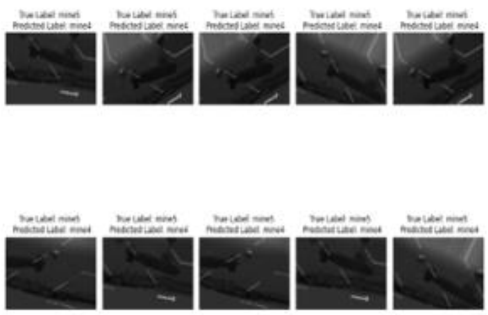


Fig.23 VGG-19 accurate predictions on test dataset **Fig.24 VGG-19 Wrong predictions on test dataset**

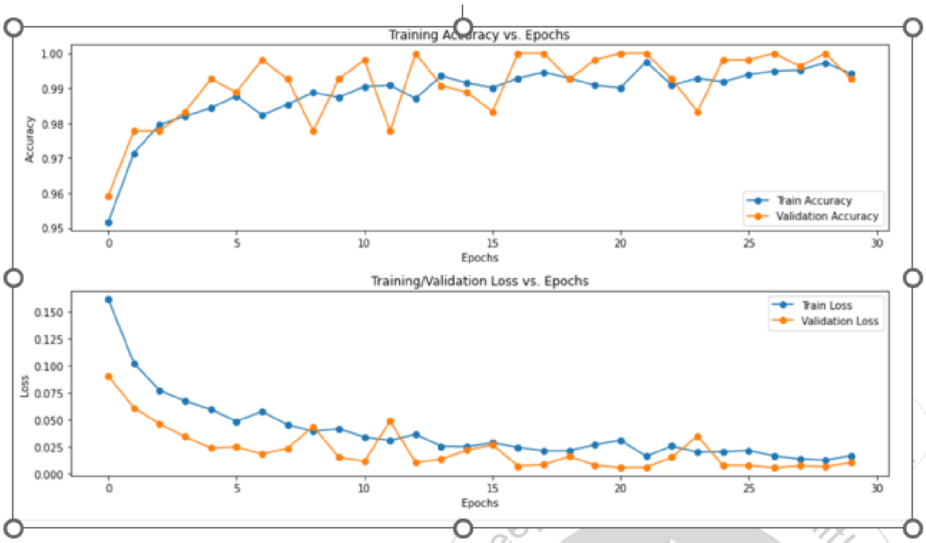


Fig.25. Model Accuracy & Loss Graph for VGG-19 Model

CNN-VGG-19

The CNN-VGG-19 hybrid model, a fusion of CNN and VGG-19 architectures, is engineered to process input images, detect mines, and categorize them among the 13 predefined classes. Demonstrating a commendable accuracy score of 86%, this hybrid classification model provides valuable insights into its performance through its Confusion Matrix, sample output for incorrect predictions, and accuracy and loss graph, depicted respectively in Figure 26, Figure 27, and Figure 28.

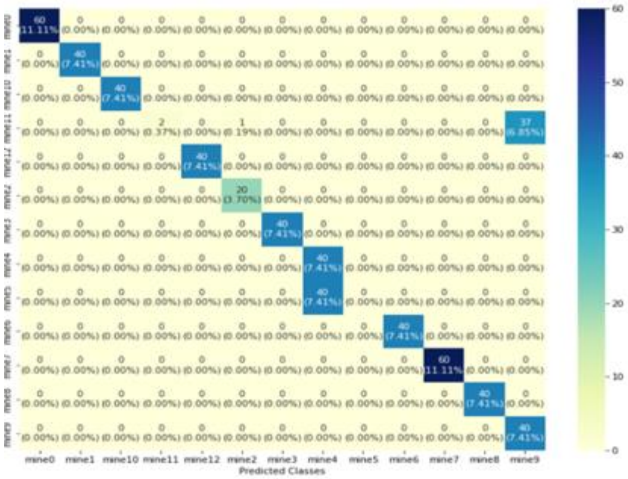


Fig.26 CNN-VGG-19 Confusion Matrix for multiple classification

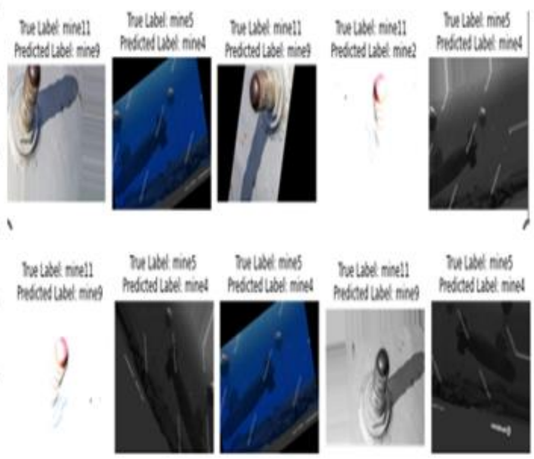


Fig.27 CNN-VGG-19 Wrong predictions on test dataset

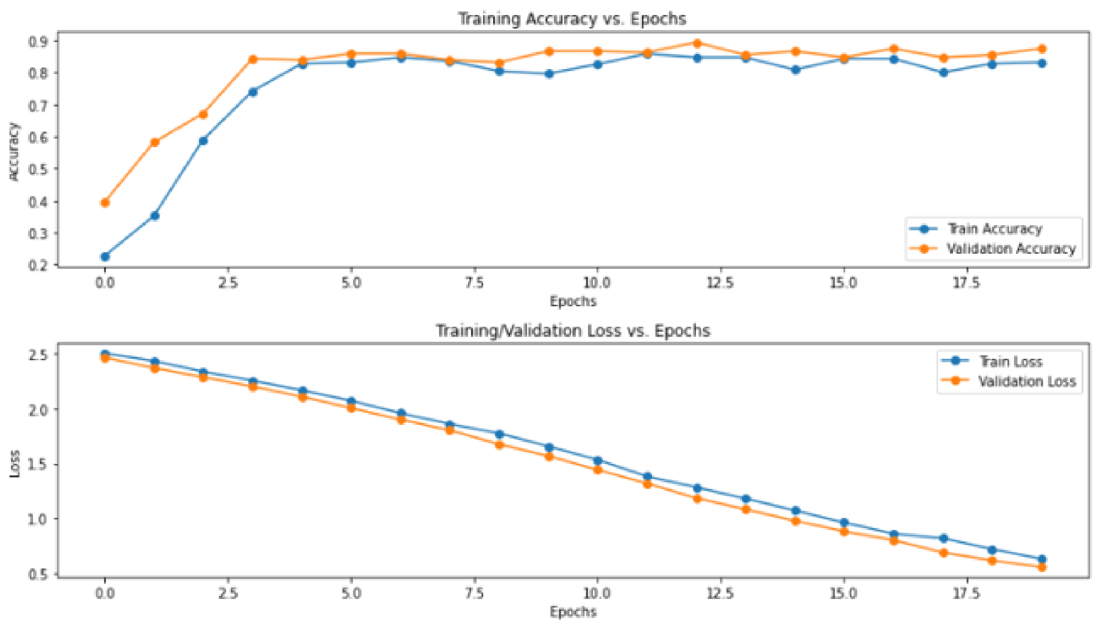
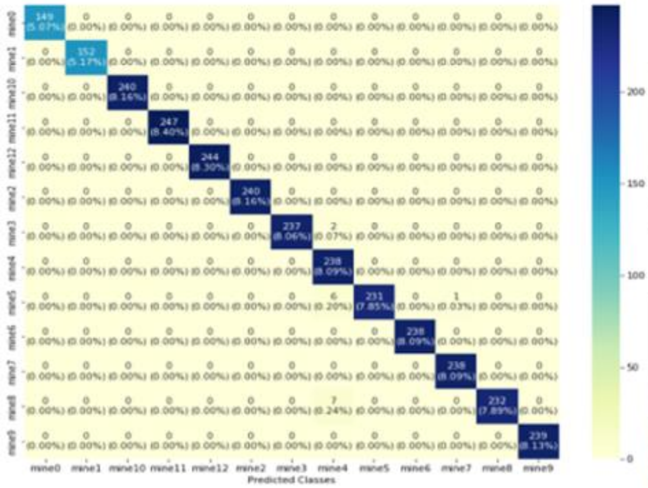


Fig.28. Model Accuracy & Loss Graph for CNN-VGG-19 Hybrid model

CNN-Mobilenet

The CNN-Mobilenet model operates by analyzing input images to detect and classify mines among the 13 designated classes, achieving an exceptional accuracy score of 99% on a test dataset comprising 668 samples. This hybrid classification model provides comprehensive insights into its performance through various metrics. Specifically, its Confusion Matrix, sample output for incorrect predictions, and accuracy and loss graph are respectively depicted in Figure 29, Figure 30, and Figure 31.



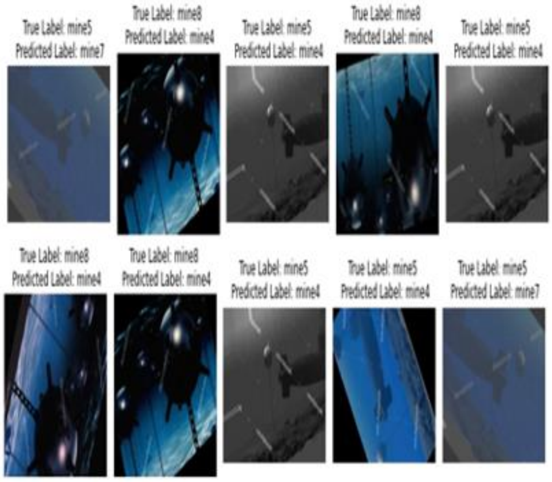


Fig.29 CNN-Mobilenet Confusion Matrix for multiple classification

Fig.30 CNN-Mobilenet Wrong predictions on test dataset

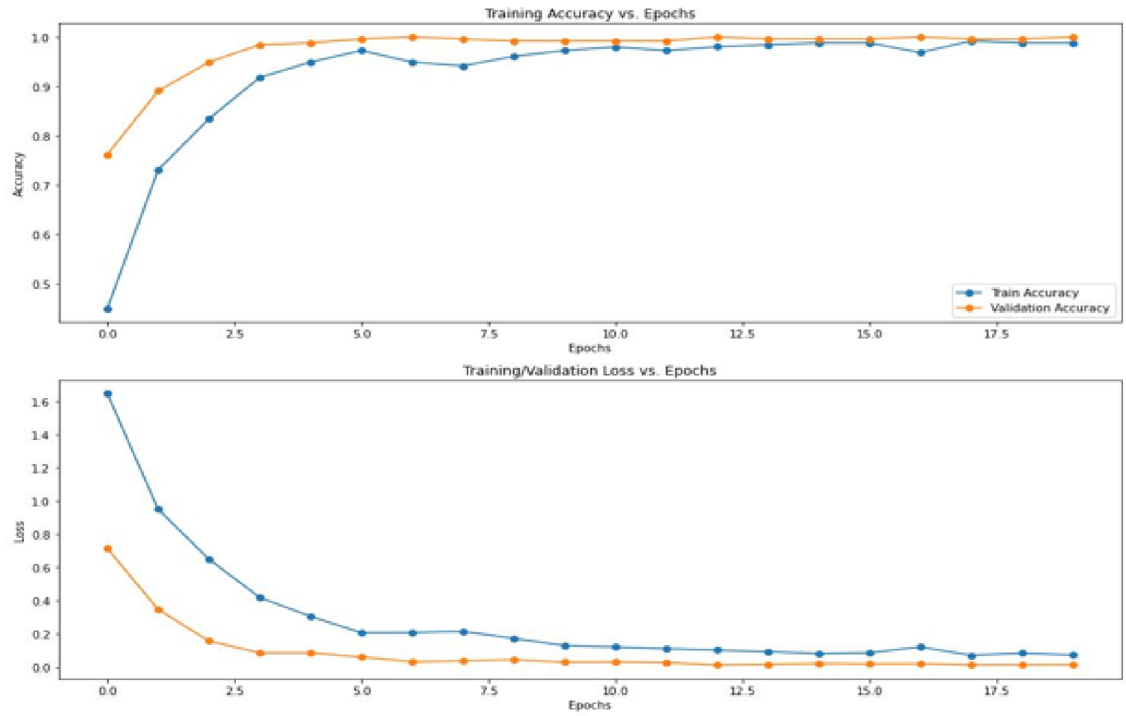


Fig.31. Model Accuracy & Loss Graph for CNN-MobileNet Hybrid model

YOLOV5

For YOLOv5, an annotated dataset comprising 173 naval mine images was prepared using the annotation tool makesense.ai, facilitating multiple classification. Within this dataset, five distinct classes of sea mine images were categorized, with the model achieving an impressive mean average precision (mAP) of 97%. Training the YOLOv5 model involved utilizing 143 images, while the model's performance was evaluated on the remaining 30 images. Figure 32 provides a visual representation of the sample output depicting the detection and classification of naval mines by YOLOv5**.**

Fig-32 Sample output -Detection and Classification of naval mines by YOLOV5

Table 5: Accuracy Results of Multiple classification

Algorithm	Accuracy Score
CNN	98
YOLOV5	97
VGG-19	98
CNN-VGG-19	86
CNN-MobileNet	99

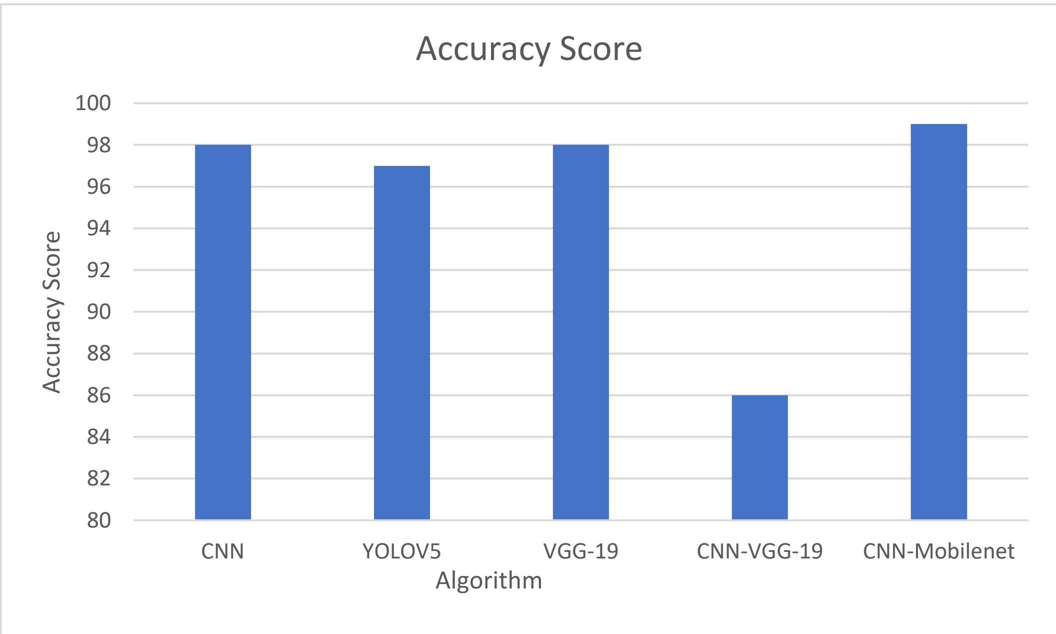


Fig-33 Accuracy graph of different deep learning and hybrid model

Table 5 provides a comprehensive overview of the results obtained for multiple classifications across various deep learning models and hybrid models, showcasing their respective accuracies. Additionally, Figure 33 illustrates the accuracy graph depicting the performance of different deep learning models and hybrid models. Upon analysis of Table 5, it is evident that CNN-MobileNet emerges as the top-performing hybrid model for detecting naval mine objects, achieving the highest accuracy score among all models considered.

5. Conclusion

The proposed system effectively identifies and classifies underwater naval mine objects. However, the dataset's size is a crucial factor in model performance. Since our dataset is relatively small, comprising self-combined images from diverse sources, the risk of overfitting increases. This occurs because outlier features may disproportionately influence the model's weights. To address this issue, data augmentation techniques are employed to expand the dataset's size. Additionally, attention is directed towards implementing dropout layers in each CNN architecture, including hybrid CNN models like CNN-VGG-19 and CNN-MobileNet.

Upon evaluating the metrics, it becomes apparent that the CNN-MobileNet hybrid model yields superior accuracy compared to other models utilized in this study. Furthermore, the inclusion of the yolov5 model, particularly noteworthy for its robust performance in processing video input, achieves a remarkable accuracy rate of 98%.

6. Future Work

The potential enhancements for the proposed system are numerous. For instance, upon detecting a mine, the system could provide additional information such as its depth, angle, and distance from the detection point. Currently, the model is tailored for identifying stationary, sphere-shaped underwater mines. However, contemporary mines come in diverse shapes. Therefore, expanding the training dataset to encompass images of mines with varying shapes and properties—such as different colors, depths, and configurations—is imperative for enhancing the model's robustness and adaptability.

References

[1] Dhiraj Neupane and Jongwon Seok, 'A Review on Deep Learning-Based Approaches for Automatic Sonar Target Recognition', International Journal of Soft Computing and Engineering (IJSCE), 22 November 2020.
[2] N Abhishek, Arjun R, Bharathesh R , Kavitha K S , Prof. Manonmani S , Dr. Shanta Rangaswamy, 'Underwater detection of mines using image Processing' International Research Journal of Engineering and TechnologyMay 2020.
[3] S. N. Geethalakshmi, P. Subashini, S. Ramya, A Study on Detection and Classification of Underwater Mines using Neural NetworksInternational Journal of Soft Computing and Engineering (IJSCE), November 2011.
[4] Killian DENOS, Mathieu RAVAUT, Antoine FAGETTE, LIM Hock-Siong,'Deep Learning applied to Underwater Mine Warfare, IEEE Xplore, 26 October 2017.
[5] Shivam S. Thakare, Amit Sahu, 'Comparative Analysis of Various Underwater Image Enhancement Techniques', International Journal of Computer Science and Mobile Computing, Publication, April 2014.
[6] Chanakya Hosamani, Sumukha Adiga,Satya Sujan,Sohan G,Srinidhi MS,Manonmani S,Shanta Rangaswamy,' Detection and feature extraction of naval mines using CNN architecture', 2018 IJSRED.
[7] Manonmani S, Akshita L, Annette Shajan L, Aneesh Sidharth L, Shanta Rangaswamy,' Underwater Mine Detection Using Histogram of oriented gradients and Canny Edge Detector',25 February IEEE Xplore.
[8] Deshpande Susharad, Dhonddev Rushikesh,' Naval Mines and Their High Blast Explosivesikes', International Research Journal of Engineering and Technology (IRJET),7 July 2017.
[9] Amrutha Kulkarni, Shanta Rangaswa my and Manonma ni S,' Comparative Study of Major Image Enhancement Algorithms', EJERS, European Journal of Engineering Research and Science, July 2017.
[10] WEI CHEN 1, 2, QIANG SUN, JUE WANG, (Member, IEEE),'A Novel Model Based on AdaBoost and Deep CNN for Vehicle Classification',IEEE Xplore, November 8 2018.
[11] Moniruzzaman, M.; Islam, S.M.S.; Bennamoun, M.; Lavery, P. Deep learning on underwater marine object detection: A survey. In Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems, Antwerp, Belgium, 18–21 September 2017; pp. 150–160.

[12] Teng B, Zhao H. Underwater target recognition methods based on the framework of deep learning: A survey. International Journal of Advanced Robotic Systems. 2020;17(6). doi:10.1177/1729881420976307

[13] Gomes, D.; Saif, A.S.; Nandi, D. Robust Underwater Object Detection with Autonomous Underwater Vehicle: A Comprehensive Study. In Proceedings of the International Conference on Computing Advancements, Dhaka, Bangladesh, 10–12 January 2020; pp. 1–10.

[14] Er, Meng Joo; Jie, Chen; Zhang, Yani; Gao, Wenxiao (2022): Research Challenges, Recent Advances and Benchmark Datasets in Deep-Learning-Based Underwater Marine Object Detection: A Review. TechRxiv. Preprint.

[15] Boris Gašparović, Jonatan Lerga, Goran Mauša & Marina Ivašić-Kos (2022) Deep Learning Approach For Objects Detection in Underwater Pipeline Images, Applied Artificial Intelligence, 36:1

[16] N. Jiang, J. Wang, L. Kong, S. Zhang, J. Dong, Optimization of underwater marker detection based on YOLOv3, Procedia Comput. Sci. 187 (2021) 52–59

[17] L. Chen, et al., Underwater object detection using Invert Multi-Class Adaboost with deep learning, in: Proc. Int. Jt. Conf. Neural Networks, 2020

[18] Ning Wang, Yuanyuan Wang, Meng Joo Er, Review on deep learning techniques for marine object recognition: Architectures and algorithms, Control Engineering Practice, Volume 118,2022,104458,

[19] Fenglei Han, Jingzheng Yao, Haitao Zhu, Chunhui Wang, "Underwater Image Processing and Object Detection Based on Deep CNN Method", Journal of Sensors, vol. 2020, Article ID 6707328, 20 pages, 2020.

[20] Shubo Xu, Minghua Zhang, Wei Song, Haibin Mei, Qi He, Antonio Liotta, A systematic review and analysis of deep learning-based underwater object detection, Neurocomputing, Volume 527, 2023, Pages 204-232, ISSN 0925-2312