A Mini Project Report On

Real Time Recognition Of Underwater Images Using Deep Learning Techniques

Submitted in partial fulfillment of the requirement for the 6^{th} semester

Bachelor of Engineering

in

Artificial Intelligence and Machine Learning

DAYANANDA SAGAR COLLEGE OF ENGINEERING

(An Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE & ISO 9001:2008 Certified) Accredited by National Assessment & Accreditation Council (NAAC) with 'A' grade

Shavige Malleshwara Hills, Kumaraswamy Layout, Bengaluru-560078

Submitted By

AYUSH ADITYA 1DS20AI015

PRAVEEN YADAV 1DS20AI042

YASH RATHI 1DS20AI057

OM PRAKASH 1DS21AI401

Under the guidance of

Prof Ramya K , Assistant professor, DSCE

2022 – 2023

Department of Artificial Intelligence and Machine Learning

DAYANANDA SAGAR COLLEGE OF ENGINEERING

Bangalore - 560078

VISVESVARAYA TECHNOLOGICAL UNIVERSITY

Dayananda Sagar College of Engineering

(An Autonomous Institute affiliated to VTU, Belagavi, Approved by AICTE & ISO 9001:2008 Certified) Accredited by National Assessment & Accreditation Council (NAAC) with ‘A’ grade
Shavige Malleshwara Hills, Kumaraswamy Layout, Bengaluru-560078

Department of Artificial Intelligence and Machine Learning

CERTIFICATE

This is to certify that the project entitled **Real Time Recognition Of Underwater Images Using Deep Learning Techniques**. is a bonafide work carried out by **AYUSH ADITYA , YASH RATHI , OM PRAKASH , PRAVEEN and USN(1DS20AI015, 1DS20AI057 , 1DS21AI401 , 1DS20AI042)** in partial fulfillment of 6th semester, Bachelor of Engineering in Artificial Intelligence and Machine Learning under Visvesvaraya Technological University, Belgaum during the year 2022-23.

Prof. RAMYA	**Dr. Vindhya P Malagi**	**Dr. B G Prasad**
(Internal Guide)	HOD, AI&ML	Principal
Designation, AI&ML, DSCE	DSCE	DSCE
Signature:...........	Signature:...........	Signature:...........

Name of the Examiners:	Signature with date:
1..........................	....
2..........................	....

Acknowledgement

We are pleased to have successfully completed the mini project **Implementation of Image Recognition for Human detection in Underwater Images.** We thoroughly enjoyed the process of working on this project and gained a lot of knowledge doing so.

We would like to take this opportunity to express our gratitude to **Dr. B G Prasad**, Principal of DSCE, for permitting us to utilize all the necessary facilities of the institution.

We also thank our respected HOD of AI&ML, DSCE, Bangalore, **Dr. Vindhya P Malagi**, for her support and encouragement throughout the process.

We are immensely grateful to our respected and learned guide, **Prof. Ramya K, Assistant professor** , DSCE for their valuable help and guidance. We are indebted to them for their invaluable guidance throughout the process and their useful inputs at all stages of the process.

We also thank all the faculty and support staff of the Department of AI & ML, DSCE. Without their support, this work would not have been possible.

Lastly, we would like to express our deep appreciation towards our classmates and our family for providing us with constant moral support and encouragement. They have stood by us in the most difficult of times.

Team member names and USN

AYUSH ADITYA 1DS20AI015

PRAVEEN YADAV 1DS20AI042

YASH RATHI 1DS20AI057

OM PRAKASH 1DS21AI401

Abstract

In order to tackle the challenges of the underwater environment, data augmentation techniques are employed during training to increase model diversity and robustness. This involves augmenting the dataset with variations in lighting, blur, and distortion, enabling the model to generalize effectively to unseen underwater scenarios. The image recognition pipeline consists of preprocessing, feature extraction, and classification stages. Preprocessing techniques enhance image quality, reduce noise, and correct color distortion caused by water absorption. Feature extraction is performed using specially designed Support Vector Classifier (SVC) architectures for underwater imagery, allowing the network to learn meaningful representations. The trained model is evaluated on a separate test set, demonstrating its effectiveness in detecting and localizing humans in challenging underwater conditions. This approach has promising applications in underwater surveillance, search and rescue operations, and marine biology research, facilitating automated analysis and decision-making processes for advancements in underwater exploration and monitoring technologies.

Contents

5 IMPLEMENTATION

6 TESTING

7 EXPERIMENTATION AND RESULTS

CONCLUSION

REFERENCES

List of Figures

1 INTRODUCTION

1.1 OBJECT DETECTION

A computer vision technology called object detection enables us to recognise and pinpoint certain things in an image or video. Using this form of localization and identification, object detection can be used to count the items in a scene, as well as to locate and track them in real time while precisely labeling them. To be more precise, object detection creates bounding boxes around the items it has found, allowing us to determine their location inside (or how they move across) a scene.

Object detection models based on deep learning typically contain two components. An encoder receives an image as input and processes it through a number of layers and blocks that teach them to extract statistical features that are used to identify and locate things. A decoder receives the encoder's outputs and determines the bounding boxes and labels for each item. A pure regressor serves as the simplest decoder. The regressor directly predicts the location and size of each bounding box by connecting to the encoder's output. The model's output is the object's and its area in the image's X, Y coordinate pair.

A region proposal network is an expansion of the regressor method. The model in this decoder suggests areas of an image where it thinks an object might be present. A classification subnetwork is then used to assign a label to the pixels in these locations (or reject the proposal). The pixels that contain those regions are subsequently sent through a classification network. This approach has the advantage of providing a more precise, adaptable model that can suggest any number of locations that might include a bounding box. But the reduced computing efficiency comes at the expense of the increased precision.

A final word about accuracy. The location and label of each object are output by object detectors, but how can we determine how well the model is performing? The most used measure for determining the position of an object is intersection-over-union (IOU). We calculate the intersection's area and divide it by the union's area given two bounding boxes.

1.1.1 SUPPORT VECTOR CLASSIFIER

SVC stands for Support Vector Classification, which is a machine learning algorithm used for classification tasks. It belongs to a class of algorithms known as Support Vector Machines (SVMs). In SVC, the algorithm creates a hyperplane or a set of hyperplanes in a high-dimensional space to separate different classes of data points. The goal is to find the hyperplane that maximizes the margin between the classes, effectively creating a decision boundary.

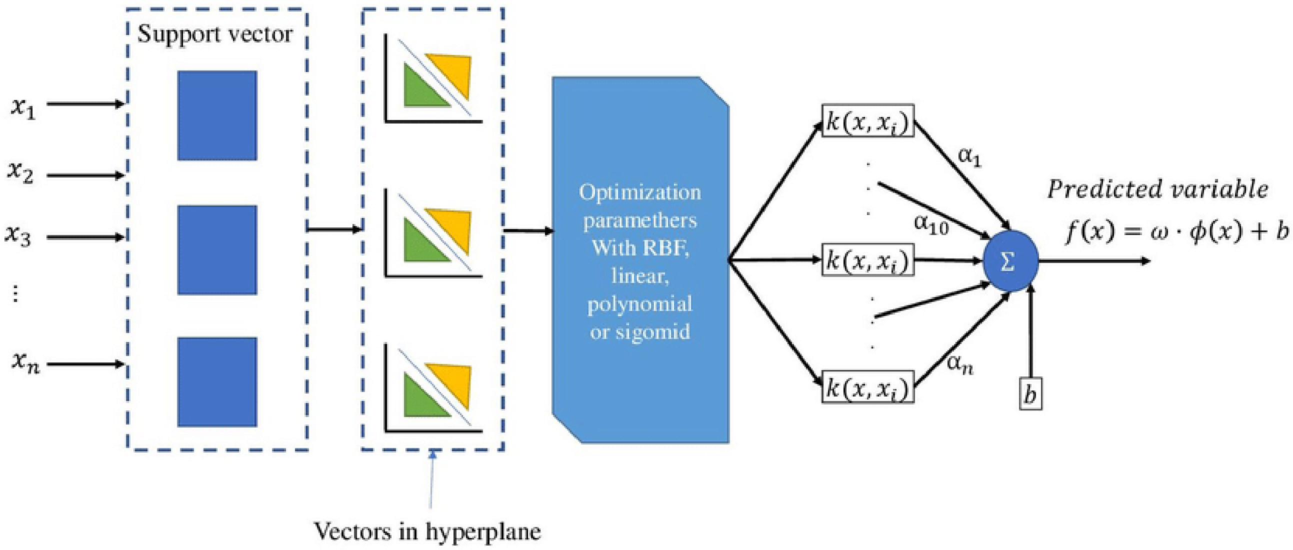


Figure 1: SVC Model Architecture

It classifies data by finding the optimal hyperplane that separates different classes. It aims to maximize the margin between data points and the decision boundary. SVM can handle both linear and non-linear classification tasks through the use of different kernel functions.For example, the popular RBF (radial basis function) kernel can be used to map data points into a higher dimensional space so that they become linearly separable. Once the data points are mapped, SVM will find the optimal hyperplane in this new space that can separate the data points into two classes.The Support vector machine algorithm is also known as a max-margin classifier. Support vector machine is a powerful tool for machine learning and has been widely used in many tasks such as hand-written digit recognition, facial expression recognition, and text classification.An SVM classifier, or support vector machine classifier, is a type of machine learning algorithm .

1.1.2 YOLO - YOU ONLY LOOK ONCE

The ”You Only Look Once” (YOLO) object detection technique divides images into a grid layout. Each cell in the grid is in charge of identifying the objects that are contained within. The object identification procedure in YOLO, which is carried out as a regression problem, provides the class probabilities of the observed images. The YOLO approach uses Support Vector Machine(SVC) to recognise objects instantaneously. This shows that the entire image is subjected to a single algorithm run for prediction. The SVC is used to forecast several bounding boxes and class probabilities at once.

Speed: This method helps speed up object detection because it can identify objects in real-time.

High degree of accuracy: The YOLO prediction method produces accurate results with low background errors.

Learning Capabilities: The algorithm has excellent learning capabilities that enable it to recognize - noise and use object representations for object detection.

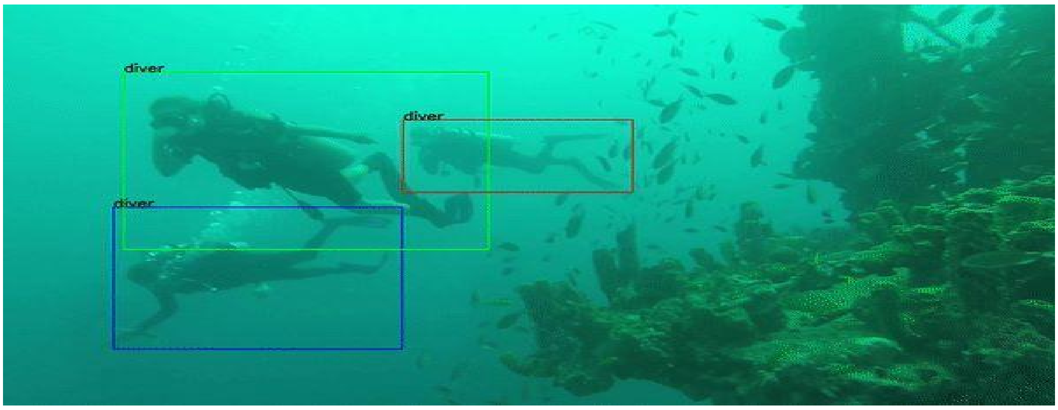


Figure 2: Object Detection from a YOLO model

Fig 2 shows that this is detecting the human in underwater.This is using SVC classifier .Yolo model is also used in it.YOLOv8 is a new state-of-the-art computer vision model built by ultralytics, the creators of YOLOv5. The YOLOv8 model contains out-of-the-box support for object detection, classification, and segmentation tasks, accessible through a Python package as well as a command line interface.

1.2 REAL TIME APPLICATION

Conservation Efforts: Image recognition can be used to identify different species of marine life, which can help conservation efforts. By recognizing different types of marine creatures and their behaviors, researchers can better understand the ocean ecosystem and make more informed decisions about conservation measures.Environmental Monitoring: Underwater images can be used to monitor the health of the ocean, including the impact of pollution, climate change, and other environmental factors. Image recognition technology can help automate this process by identifying changes in the underwater environment, such as the presence of harmful algal blooms or changes in water temperature

Search and Rescue: Image recognition can be used to locate lost or missing people in the water. By analyzing underwater images, rescue teams can identify objects or people that are not visible to the naked eye, which can greatly improve search and rescue efforts.

Commercial and Recreational Activities: Image recognition can also be used for commercial and recreational activities such as fishing, scuba diving, and ocean tourism. By analyzing underwater images, businesses can better understand the behavior of marine creatures and make more informed decisions about where to fish or how to interact with marine life.

1.3 ORGANIZATION OF REPORT

The project report is set up as follows:

We talk about the problem statement and our solution in Chapter (2). The other present tech- nologies are covered in the same chapter as well. Brief descriptions on the literature review of the publications that relate to the problem statement and the suggested solutions are mentioned in the next chapter, chapter (3). The System Outline, architecture and design are presented in Chapter(4) in the form of data flow diagrams and sequence diagrams. The requirements and details on the implementation of the suggested system are provided in the following chapter, chapter (5).In Chapter (6), the product's testing and the anticipated outcomes are covered. The contribut- ing factors are and how they impact the system are covered in Chapter (7). The same chapter includes defining the ideal system parameters. The paper is concluded next, with a remark of future improvement. Details about the references made during the creation of the system are cited in the following chapter. The source code and additional supporting data are put together in the Appendix.

2 PROBLEM STATEMENT AND PROPOSED SOLUTION

2.1 PROBLEM STATEMENT

Given a set of underwater images with varying degrees of noise, distortion, and lighting conditions, the task is to develop an accurate and efficient image recognition system using artificial neural networks. The system should be able to classify the images into predefined categories such as different species of marine organisms, underwater objects, or environmental conditions.

2.2 EXISTING SYSTEMS

2.2.1 SWIPENET: OBJECT DETECTION IN NOISY UNDERWATER IMAGES.

There are multiple variations of wearable systems which have been proposed in the past.The following describes these systems:

- Deep learning based object detection methods have achieved promising performance in controlled environments. However, these methods lack sufficient capabilities to handle underwater object detection due to these challenges: images in the underwater datasets and real applications are blurry whilst accompanying severe noise that confuses the detectors and objects in real applications are usually small. In this paper, we propose a novel Sample-WeIghted hyPEr Network (SWIPENET), and a robust training paradigm named Curriculum Multi-Class Adaboost (CMA), to address these two problems at the same time .

- Firstly, the backbone of swipenet produces multiple high resolution and semantic-rich Hyper Feature Maps, which significantly improve small object detection. Secondly, a novel sample-weighted detection loss function is designed for SWIPENET, which focuses on learning high weight samples and ignores learning low weight samples. Moreover, inspired by the human education process that drives the learning from easy to hard concepts, we here propose the CMA training paradigm that first trains a clean detector which is free from the influence of noisy data

- Then, based on the clean detector, multiple detectors focusing on learning diverse noisy data are trained and incorporated into a unified deep ensemble of strong noise immunity. Experiments on two underwater robot picking contest datasets (URPC2017 and URPC2018) show that the proposed SWIPE NET+CMA framework achieves better accuracy in object detection against several state-of-the-art approaches.

2.2.2 UNDERWATER FISH DETECTION USING DEEP LEARNING FOR WATER POWER APPLICATIONS.

- Clean energy from oceans and rivers is becoming a reality with the development of new technologies like tidal and instream turbines that generate electricity from naturally flowing water. These new technologies are being monitored for effects on fish and other wildlife using underwater video. Methods for automated analysis of underwater video are needed to lower the costs of analysis and improve accuracy. A deep learning model, YOLO, was trained to recognize fish in underwater video using three very different datasets recorded at real-world water power sites. Training and testing with examples from all three datasets resulted in a mean average precision (mAP) score of 0.5392.
- To test how well a model could generalize to new datasets, the model was trained using examples from only two of the datasets and then tested on examples from all three datasets. The resulting model could not recognize fish in the dataset that was not part of the training set. The mAP scores on the other two datasets that were included in the training set were higher than the scores achieved by the model trained on all three datasets. These results indicate that different methods are needed in order to produce a trained model that can generalize to new data sets such as those real world

2.3 PROPOSED SOLUTION

2.3.1 DATA COLLECTION

In the data collection phase, a comprehensive dataset of underwater images is collected, which includes a diverse range of human bodies and fishes. This dataset is carefully curated and annotated to ensure accurate training and testing of the neural network model. The collected images cover various underwater conditions, lighting variations, and different poses of humans and fishes, enabling the model to learn from a wide range of scenarios.

2.3.2 NEURAL NETWORK ARCHITECTURE DESIGN

During the neural network architecture design stage, a suitable architecture is carefully crafted to address the specific task of human body detection in underwater images. This involves considering the unique challenges of the underwater environment, such as water absorption and lighting variations. The architecture may include specialized layers, such as convolutional layers, pooling layers, and potentially incorporating Support Vector Machine (SVM) network hyperparameters to optimize the network's performance.

2.3.3 MODEL TRAINING

The preprocessed dataset is then utilized to train the neural network model. This training process involves feeding the dataset through the network and iteratively adjusting the model's internal parameters using optimization techniques such as stochastic gradient descent (SGD) or the Adam optimizer. The goal is to minimize the difference between the predicted output of the model and the ground truth labels in the training dataset, enabling the model to learn the patterns and features associated with human body detection in underwater images.

2.3.4 MODEL EVALUATION

After training, the performance of the trained model is evaluated on a separate test dataset. This evaluation phase aims to assess how well the model generalizes to unseen data and how accurately it can detect human bodies in underwater images. Metrics such as precision, recall, and F1 score are calculated to measure the model's performance, providing insights into its accuracy and effectiveness in differentiating human bodies from other objects or fishes in the underwater environment.

2.3.5 MODEL REFINEMENT

Based on the evaluation results, the model may undergo refinement. This involves fine-tuning the model by adjusting hyperparameters, such as learning rate or regularization parameters, to improve its performance. Additionally, modifications to the network architecture may be considered, such as adding or removing layers or changing their configurations. This iterative refinement process aims to enhance the model's accuracy and robustness in detecting human bodies in various underwater scenarios.

2.3.6 DEPLOYMENT

Once the model has been refined and optimized, it is ready for deployment in a real-world underwater environment. This stage involves setting up specialized hardware and communication systems to facilitate the model's operation. The deployment may include integrating the model into underwater robotics or surveillance systems, allowing it to analyze and detect human bodies in real-time, contributing to applications such as underwater surveillance, search and rescue operations, or marine biology research.

3.0 LITERATURE SURVEY

3.1 SWIPENET AND CMA(CURRICULUM MULTI-CLASS ADABOOST)

Authors: Long Chen, Feixiang Zhou , Shengke Wang **, "**SWIPENET: Object detection in noisy underwater scenes".

Description:- WeIghted hyPEr Network (SWIPENET), and a robust training paradigm named Curriculum Multi-Class Adaboost (CMA), to address these two problems at the same time. Firstly, the backbone of SWIPENET produces multiple high resolution and semantic-rich Hyper Feature Maps, which significantly improve small object detection. Secondly, a novel sample-weighted detection loss function is designed for SWIPENET, which focuses on learning high weight samples and ignores learning low weight samples. Moreover, inspired by the human education process that drives the learning from easy to hard concepts, we here propose the CMA training paradigm that first trains a clean detector which is free from the influence of noisy data. Then, based on the clean detector, multiple detectors focusing on learning diverse noisy data are trained and incorporated into a unified deep ensemble of strong noise immunity.

3.2 FISH DETECTION

Authors: Wenwei Xu, Shari Matzner , "Underwater Fish Detection using Deep Learning for Water Power Applications".

Description:- Clean energy from oceans and rivers is becoming a reality with the development of new technologies like tidal and instream turbines that generate electricity from naturally flowing water. These new technologies are being monitored for effects on fish and other wildlife using underwater video. Methods for automated analysis of underwater video are needed to lower the costs of analysis and improve accuracy. A deep learning model, YOLO, was trained to recognize fish in underwater video using three very different datasets recorded at real-world water power sites. Training and testing with examples from all three datasets resulted in a mean average precision (mAP) score of 0.5392. To test how well a model could generalize to new datasets, the model was trained using examples from only two of the datasets and then tested on examples from all three datasets.

The resulting model could not recognize fish in the dataset that was not part of the training set.

The mAP scores on the other two datasets that were included in the training set were higher than the scores achieved by the model trained on all three datasets. These results indicate that different methods are needed in order to produce a trained model that can generalize to new data sets such as those encountered in real world applications.

3.3 RolMix

Authors: Wei-Hong Lin, Jia-Xing Zhong, Shan Liu, Thomas Li, Ge L**i** , " Fusion among Multiple Images for Underwater Object Detection"

Description:- Generic object detection algorithms have proven their excellent performance in recent years. However, object detection on underwater datasets is still less explored. In contrast to generic datasets, underwater images usually have color shift and low contrast; sediment would cause blurring in underwater images. In addition, underwater creatures often appear closely to each other on images due to their living habits. To address these issues, our work investigates augmentation policies to simulate overlapping, occluded and blurred objects, and we construct a model capable of achieving better generalization. We propose an augmentation method called RoIMix, which characterizes interactions among images. Proposals extracted from different images are mixed together. Previous data augmentation methods operate on a single image while we apply RoIMix to multiple images to create enhanced samples as training data. Experiments show that our proposed method improves the performance of region-based object detectors on both Pascal VOC and URPC datasets.

4.0 ARCHITECTURE AND DESIGN

4.1 SYSTEM OVERVIEW

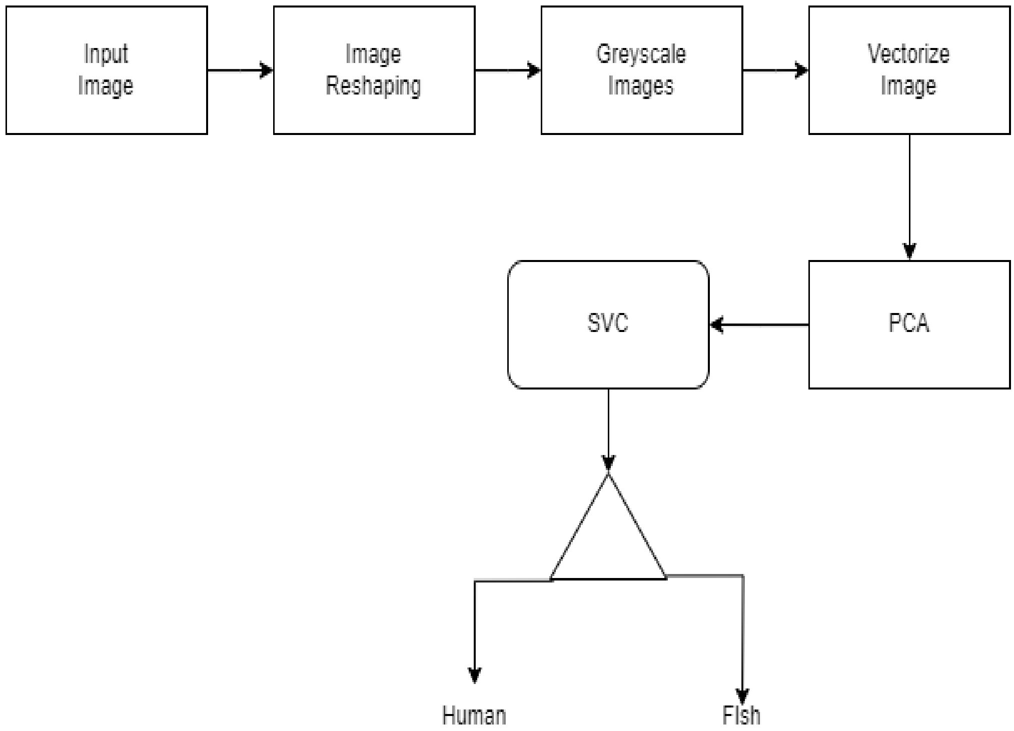


Figure 3: System Overview

The overview of the system is represented in Fig 3. It shows the modules involved in building the system i.e:

1. Object to be detected.
2. Model training.
3. AI Model(YOLOv8).
4. Camera for detection.
5. Object Detection.

4.2 SOFTWARE ARCHITECTURE

4.2.1 YOLOV8 BLOCK DIAGRAM

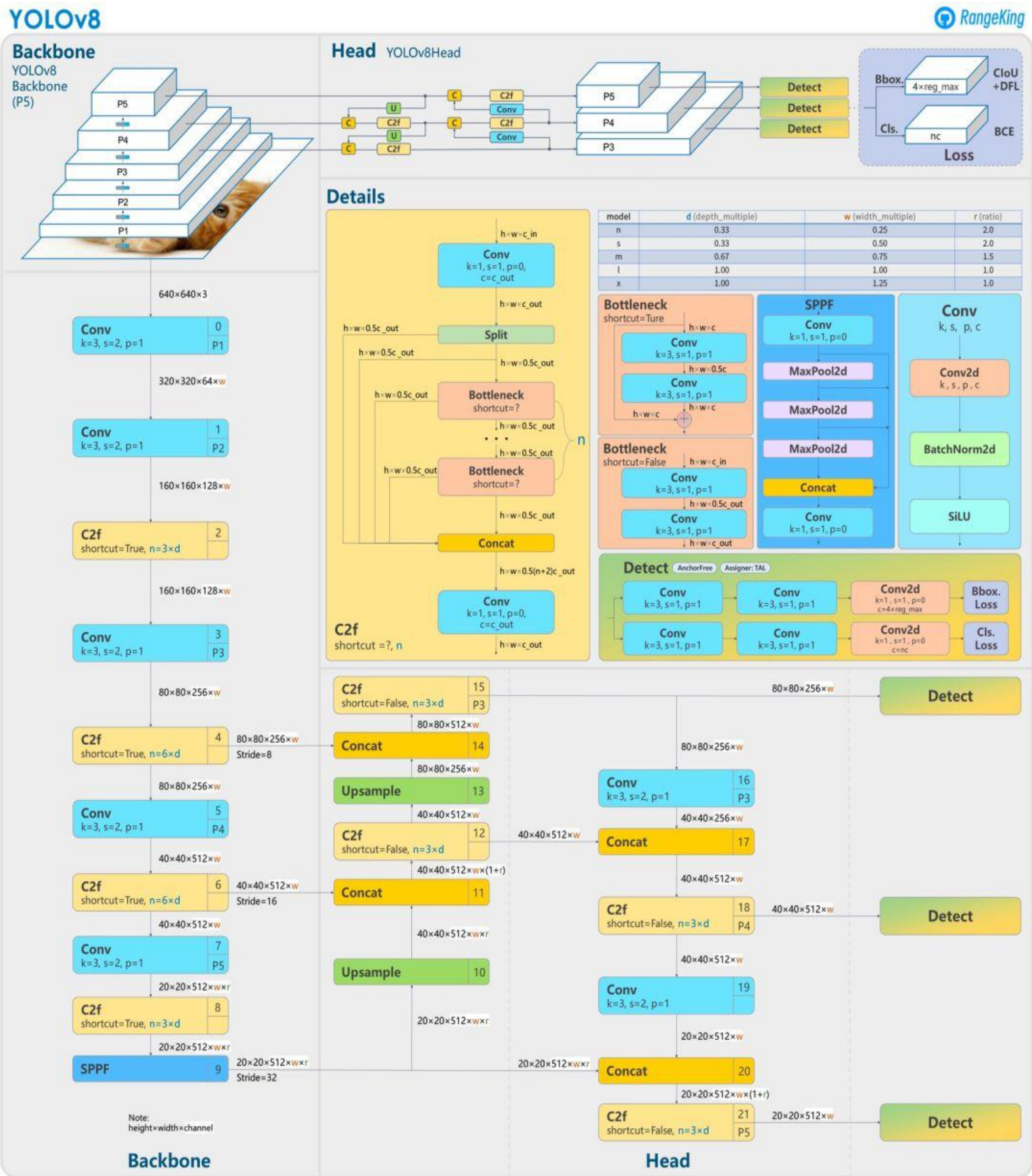


Figure 4 : System Block Diagram

The overall block diagram of the proposed system is shown in Figure.4.

1. Improved architecture
2. Enhanced object detection performance
3. Multi-scale feature fusion
4. Efficient inference
5. Open-source implementation and community support

4.2.2 DATA FLOW DIAGRAM

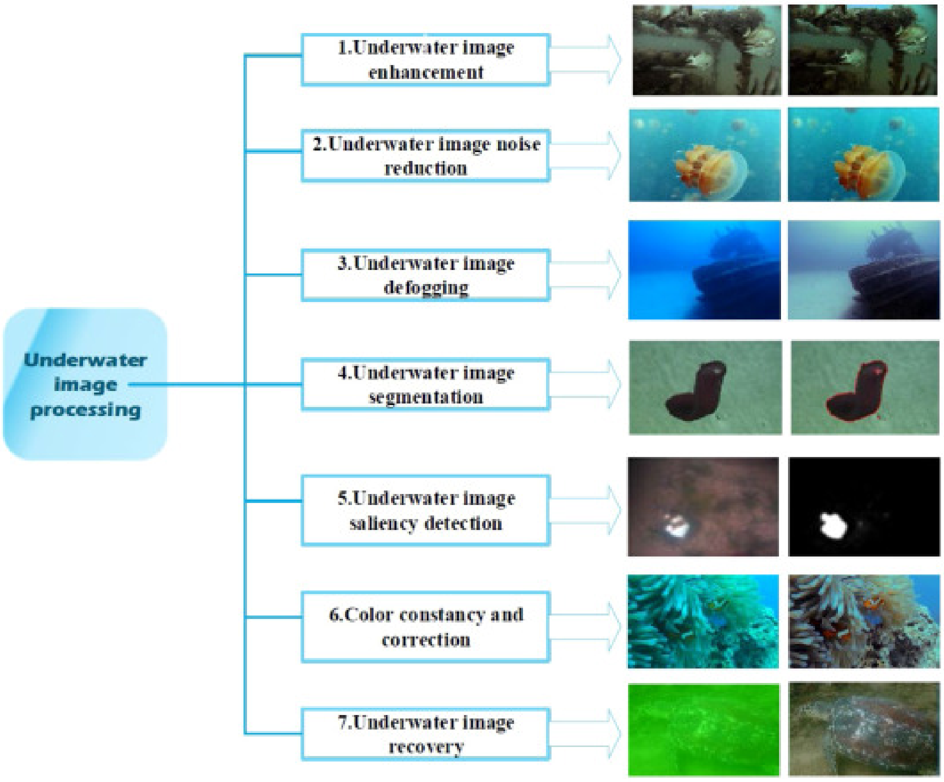


Figure 5: Data Flow Diagram

The flow of data right from its input stage to its final output stage is represented by the Data Flow diagram. It gives an overall overview of the system implementation without going deeply into the intricacies involved. The flow of data in the system is as follows:

1. Images are captured from the video stream using the binocular cameras.

2. Images are preprocessed and cleaned in order to be ready to be sent to the YOLOv8 model.

3. Preprocessed images are sent to the AI model to be operated on.

4. Bounding boxes are created for the images to find their centroids and calculate their outliers.

5 IMPLEMENTATION AND TESTING

5.1 IMPLEMENTATION PLATFORM

5.1.1 SOFTWARE

- **Operating System:** Windows

○

- **Software Used:** Jupyter Notebook,YoloV8
- **Programming Languages:** Python 3.10
- **Server:** Local Server

5.1.2 HARDWARE

- Laptop
- Graphical Processing Unit (GPU if available)
- RAM, ROM

5.2 IMPLEMENTATION DETAILS

5.2.1 ORGANIZATION OF IMPLEMENTATION FILES

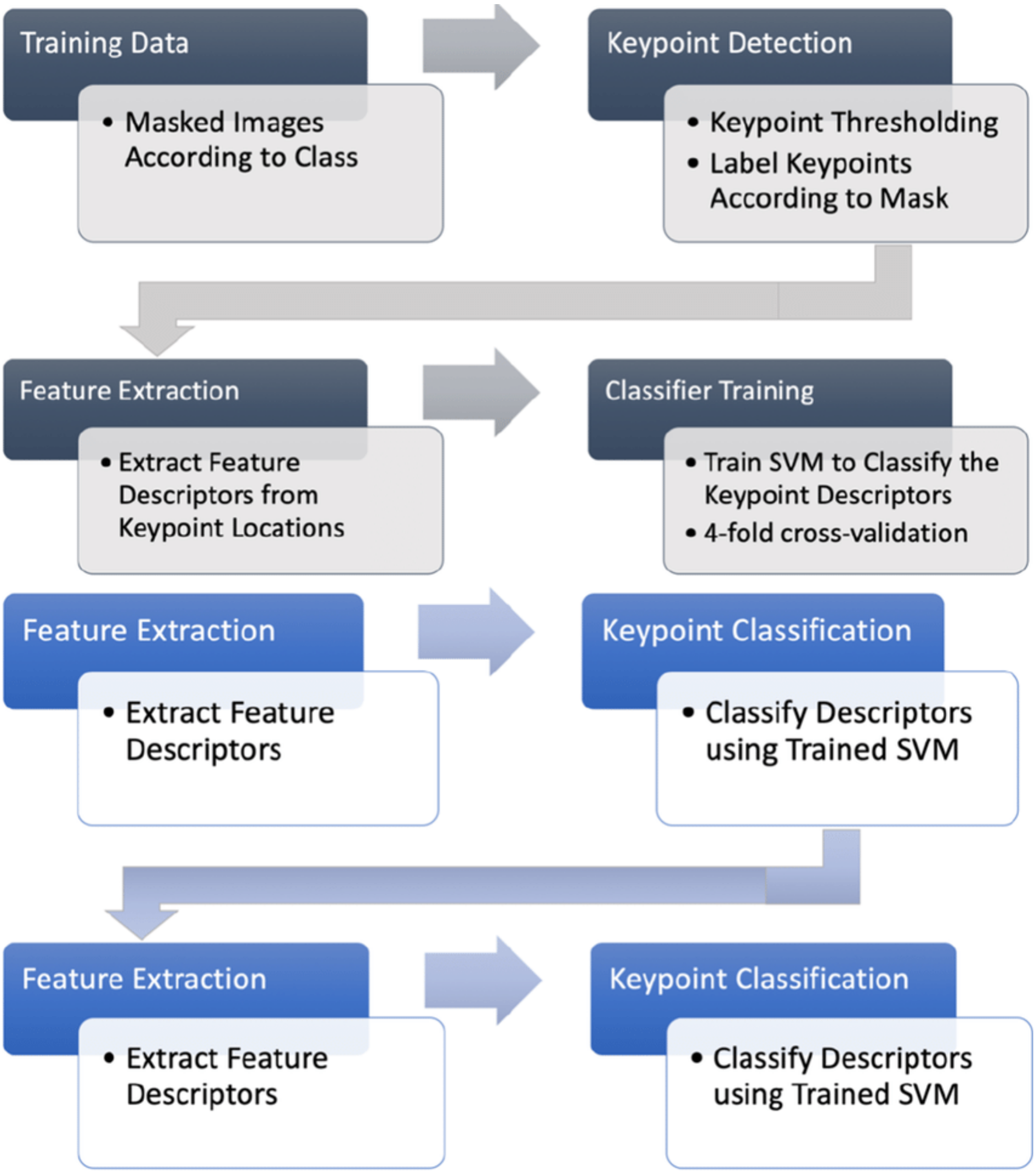


Figure 6 : Flowchart for the model

5.2.2 TRAINING A MACHINE LEARNING MODEL

Train Model

```
.]: from sklearn.svm import SVC
```

```
:]: sv = SVC()
    sv.fit(xtrain, ytrain)
```

```
:]: SVC()
```

Evaluation

```
:]: print("Training Score:", sv.score(xtrain, ytrain))
    print("Testing Score:", sv.score(xtest, ytest))
```

```
Training Score: 0.9594838709677419
Testing Score: 0.921733895243829
```

Figure 7: Training Diagram

Use a large dataset of images. The more images you have, the better the model will be able to learn.Use a variety of images. The model should be trained on images of different fish species, as well as images of other underwater objects.Use a powerful machine learning algorithm. CNNs are a good choice for underwater classification, but other algorithms may also be effective.Evaluate the model thoroughly. Make sure the model is able to classify the test images accurately.Deploy the model in a production environment. Once the model is deployed, you can use it to classify underwater images in real time.. This dataset should include images of different fish species, as well as images of other underwater objects that you want to classify. The images should be of high quality and should be properly labeled.This may involve removing noise, cropping the images, and resizing them to a consistent size. There are many different machine learning algorithms that can be used for underwater classification.Our choices include support vector machines (SVMs).Once the model is evaluated and you are satisfied with its performance, you can deploy it to a production environment.Prediction of Underwater Images Using DL Techniques Department ofAI&ML, 2022-2023

5.3 DATASET

The dataset that we are using is from Kaggle and roboflow.

This dataset can be used for the following purposes:

(Fish)

- Train object detection model to recognize underwater species.
- Prototype fish detection system.
- Identifying fish with computer vision.
- Free fish dataset.
- Free fish identification dataset.
- Scuba diving object detection dataset.
- Fish bounding boxes.
- Fish species annotations.
- These images have been listed in the public domain.

(Human)

- Train object detection model to recognize underwater human
- Prototype human detection system.
- Identifying humans with computer vision.
- Free human dataset.
- Free human identification dataset.
- Scuba diving object detection dataset.
- Human bounding boxes.
- Human species annotations.
- These images have been listed in the public domain.

DATASET LINKS

Roboflow:- https://public.roboflow.com/object-detection/fish

Roboflow: -https://universe.roboflow.com/search?q=underwater%20human

Kaggle:-https://www.kaggle.com/datasets/aungpyaeap/fish-market

Fig 7 Fish data

Fig 8 Human data

6 TESTING

6.1 SOFTWARE TESTING

We first ran the Yolov8 object detection algorithm on two stereo images stored on the system. On successfully getting an output, we connected the algorithm and programmed it in such a way that it runs continuously, taking a video stream as input. Once the algorithm was working successfully with the hardware, we tried training the model specific to a college classroom. For this, we took a few pictures of the classroom from different angles, and captured different objects. We then ran these weights through the Yolov8 algorithm, in order to test the algorithm in a unique environment.

The algorithm was successfully able to take input through the cameras and detect the objects in both cases. Once object detection worked, we moved to depth estimation. We worked on calculating the parallax between the two images and sent it through a triangulation function to calculate the depth. The latest version of YOLO by Ultralytics. As a cutting-edge, state-of-the-art (SOTA) model, YOLOv8 builds on the success of previous versions, introducing new features and improvements for enhanced performance, flexibility, and efficiency. YOLOv8 supports a full range of vision AI tasks, including detection, segmentation, pose estimation, tracking and classification. This versatility allows users to leverage YOLOv8's capabilities across diverse applications and domains

7 EXPERIMENTATION AND RESULTS

7.1 CAMERA TESTING

When the cameras are powered and successfully connected to the Wifi, the camera stream can be watched on the screen.The difference in the images in terms of parallax can be clearly seen on desktop.

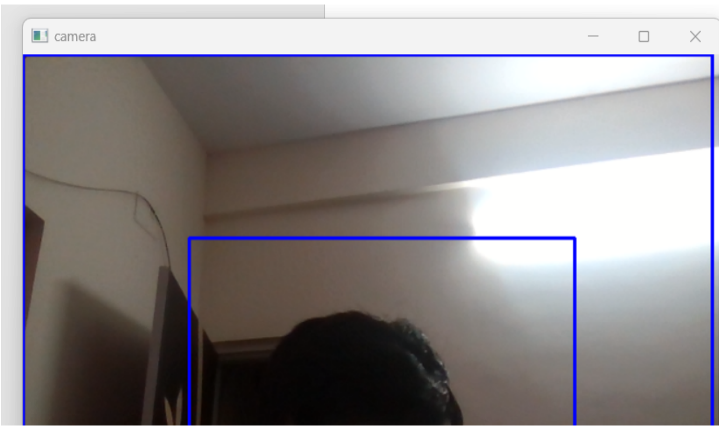


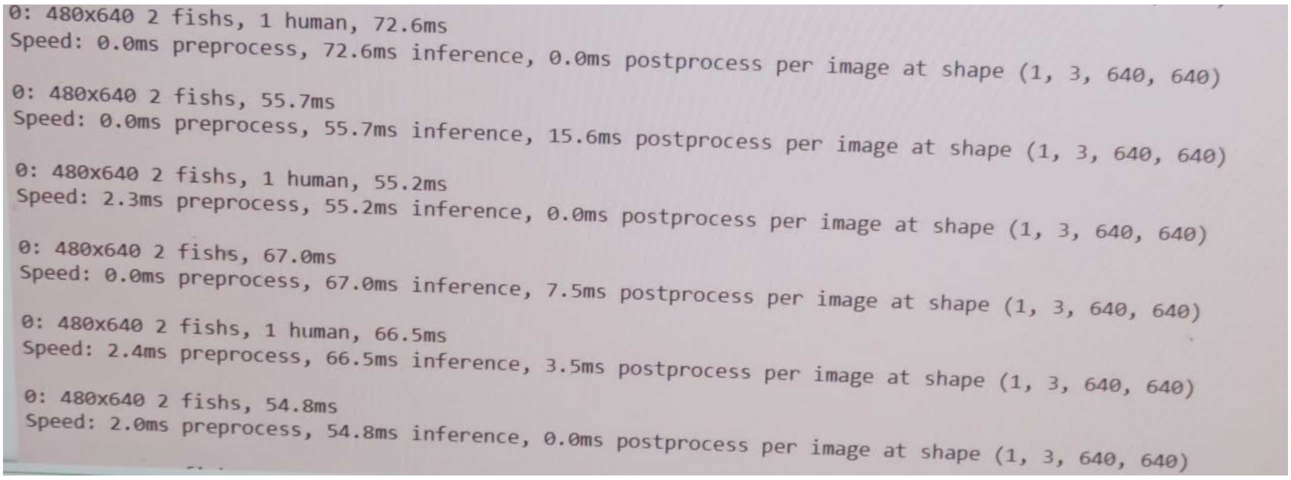
Figure 8: Camera Testing

7.2 ALGORITHM ANALYSIS

The time complexity of SVC depends on the size of the training dataset and the complexity of the chosen kernel function. In general, the training time complexity of SVC can be approximated to be between O(n^2) and O(n^3), where n is the number of training samples. However, the actual training time may vary based on the implementation and optimization techniques employed.

The prediction time complexity of SVC is typically O(m), where m is the number of support vectors, which is usually smaller than the total number of training samples. This makes SVC efficient during the prediction phase.

SVC is a powerful and versatile classification algorithm known for its ability to handle both linear and non-linear classification problems. The performance of SVC depends on several factors, such as the choice of kernel function, regularization parameter (C), and kernel-specific parameters.



```
0: 480x640 2 fishs, 1 human, 72.6ms
Speed: 0.0ms preprocess, 72.6ms inference, 0.0ms postprocess per image at shape (1, 3, 640, 640)

0: 480x640 2 fishs, 55.7ms
Speed: 0.0ms preprocess, 55.7ms inference, 15.6ms postprocess per image at shape (1, 3, 640, 640)

0: 480x640 2 fishs, 1 human, 55.2ms
Speed: 2.3ms preprocess, 55.2ms inference, 0.0ms postprocess per image at shape (1, 3, 640, 640)

0: 480x640 2 fishs, 67.0ms
Speed: 0.0ms preprocess, 67.0ms inference, 7.5ms postprocess per image at shape (1, 3, 640, 640)

0: 480x640 2 fishs, 1 human, 66.5ms
Speed: 2.4ms preprocess, 66.5ms inference, 3.5ms postprocess per image at shape (1, 3, 640, 640)

0: 480x640 2 fishs, 54.8ms
Speed: 2.0ms preprocess, 54.8ms inference, 0.0ms postprocess per image at shape (1, 3, 640, 640)
```

Figure 9: Terminal output

Figure 10: Human Prediction

Figure 11: Fish Prediction

CONCLUSION

In conclusion, underwater detection can be challenging due to various factors such as low visibility, poor lighting conditions, and color distortion. However, there are several approaches that can be used to improve the accuracy of image recognition for underwater images, such as data augmentation, preprocessing, transfer learning, object detection, ensemble learning, domain-specific datasets, and sensor fusion.

A combination of these approaches can be used to develop a robust image recognition system that can accurately recognize objects in underwater environments. It is important to note that the specific approach used will depend on the specific requirements of the application, and further research and development are needed to improve the accuracy and reliability of image recognition for underwater images.

FUTURE WORK

Future work in underwater detection focuses on enhancing technologies and techniques for marine exploration, environmental monitoring, defense, and resource management. It involves improving sonar systems for better resolution and accuracy, advancing autonomous underwater vehicles (AUVs) with improved sensors and navigation capabilities, developing advanced imaging technologies for high-resolution underwater imaging, creating monitoring systems for environmental assessments, improving underwater communication, enhancing threat detection systems, and refining methods for resource exploration. These advancements aim to deepen our understanding of underwater environments, protect marine ecosystems, and efficiently utilize underwater resources.

REFERENCES

[1] SWIPENET: Object detection in noisy underwater images 19 Oct 2020 · Long Chen, Feixiang Zhou, Shengke Wang, Junyu Dong, Ning li, Haiping Ma, Xin Wang, Huiyu Zhou ·

[2]Underwater object detection using Invert Multi-Class Adaboost with deep learning 23 May 2020 · Long Chen, Zhihua Liu, Lei Tong, Zheheng Jiang, Shengke Wang, Junyu Dong, Huiyu Zhou

[3]Underwater Fish Detection using Deep Learning for Water Power Applications 5 Nov 2018 · Wenwei Xu, Shari Matzner ·

[4]. Chen, Z., Zhang, Z., Dai, F., Bu, Y., Wang, H.: Monocular vision-based underwater object detection. Sensors 17(8), 1784 (2017)

[5]. Cong, Y., Fan, B., Hou, D., Fan, H., Liu, K., Luo, J.: Novel event analysis for human-machine collaborative underwater exploration. Pattern Recognition 96, 106967 (2019) novel method of training a multi-species seagrass classifier using a dataset of single species photos, the emphasis is on classifying seagrass into key morphological super-classes, together with background.

[6]Wang The accuracy of the four-class categorization was 88.2% overall, and it was increased to 92.4% by breaking down the Background class into the Substrate and Water Column sub-classes. The method could be improved with more training data and field testing in order to divide morphological super-classes into distinct seagrass species. The suggested method might also be used as a basis for calculating species-specific percentage cover and density. The dataset and analytic code have been made available by the authors to support more research in

[7].In this paper, SWIPENet, a neural network architecture created for the detection of small underwater objects, is introduced. To deal with the problem of noise, the authors suggest a sample reweighting technique dubbed Invert Multi-Class Adaboost (IMA). Although the suggested method's time complexity is M times higher since it is an ensemble of M deep neural networks, it nevertheless performs at the cutting edge on difficult datasets. The goal of next research should be to make the method's computations simpler. To further enhance SWIPENet, the authors propose combining attention mechanisms and unique loss functions from existing deep models, as these strategies have shown promise in resolving noise and small object detection difficulties.

[8]. This study provides a universal tracking-by-detection method for robots with limited resources that makes use of optical sense of unengineered targets and lightweight neural network architectures, including recurrent extensions. The authors used supervised learning with little training data to successfully show multi-robot convoying in open sea conditions. Additionally, they carried out a thorough evaluation of tracker variations for visual convoying activities. By permitting end-to-end training of the complete architecture, the scientists hope to enhance temporal-based bounding box detection in the future. They also intend to expand the research to visual servoing with several bounding boxes per frame and, by using more powerful predictive models, to improve target tracking robustness despite breaks in visual contact.

[9]Dual Refinement Underwater Object Detection Network ECCV 2020 · Baojie Fan, Wei Chen, Yang Cong, Jiandong Tian ·Underwater target detection based on improved YOLOv7 14 Feb 2023 · Kaiyue Liu, Qi Sun, Daming Sun, Mengduo Yang, Nizhuan Wang · Excavating RoI Attention for Underwater Object Detection 24 Jun 2022 · Xutao Liang, Pinhao Song .CDNet is all you need: Cascade DCN based underwater object detection RCNN 25 Nov 2021 · Di Chang.