

# AI-Based System for Underwater Monitoring: Image Improvement, Detection, and Depth Mapping

**Prof. Archana Kotakar**

Department of Computer Engineering  
JSPM's Jaywantrao Sawant College Of  
Engineering, Pune

**Pallavi Gawai**

Department of Computer Engineering  
JSPM's Jaywantrao Sawant College Of  
Engineering, Pune

**Kajal Shelke**

Department of Computer Engineering  
JSPM's Jaywantrao Sawant College Of  
Engineering, Pune

**Priyanka Gujale**

Department of Computer Engineering  
JSPM's Jaywantrao Sawant College  
of Engineering, Pune

**Dnyaneshwari Suke**

Department of Computer Engineering  
JSPM's Jaywantrao Sawant College Of  
Engineering, Pune

**Abstract** - This paper describes a software-based artificial intelligence (AI) system that monitors underwater environments through image enhancement, object identification, and depth mapping techniques. Current methods of monitoring underwater environments typically require multiple types of hardware sensors and Internet of Things devices to capture the images, which results in substantial time, expense, and effort if these devices were only intended for use by a single organization. Instead of relying on ubiquitous hardware per se to capture images, the proposed system consists of commercially available computer vision and deep learning algorithms that process underwater images without any dependence on how the image is acquired or captured. Thus, improvements to both image visibility and color have been accomplished using enhancement algorithms that perform under low-light and turbid conditions. Further enhancements include object detection models that automatically detect marine organisms and obstacles on the ocean floor or items that are underwater. Moreover, depth mapping techniques have been applied to provide an estimate of the spatial attributes that exist within the visual information generated by an image of an underwater environment. Overall, this new system offers a cost-effective, highly scalable solution for conducting underwater assessments compared to existing methods, as the experimental results demonstrate both enhanced image quality, accurate object detection, and precise depth estimations can be achieved utilizing this new image-processing approach.

**Keywords**—Underwater Monitoring, Image Enhancement, Object Detection, Depth Mapping, Deep Learning, Computer Vision.

## I. INTRODUCTION

Researching how well the water under the sea is doing has become an important part of research in many ocean sciences, such as environmental monitoring and searching for new resources. The need to monitor underwater ecosystems, find things underwater, and evaluate the physical conditions of the ocean has created a demand for technology to create better monitoring platforms. However, there are many challenges when trying to collect data underwater, such as the absorption and scattering of light, the degree to which color is distorted underwater, and how little visibility you have when trying to observe things underwater. All of these factors affect the quality of an image and make it difficult to collect and analyze the images you have. There are currently no technologies capable of collecting high-quality images for underwater purposes [1][2]. Most existing technologies use hardware-based technologies such as sensors and Internet of Things (IoT) devices, which can be expensive to install and maintain and are very slow and limited to scaling up in the number of devices [3][4]. Therefore, many researchers are beginning to explore the potential for developing software/algorithmic-based solutions to assist with efficiently processing visual data without having to depend on physical hardware devices.

The latest technology advancements within both artificial intelligence and computer vision have made striking advancements toward the improvement of underwater image processing. The use of various enhancement techniques utilizing histogram equalization, white balance, and deep learning-based restoration techniques has been successfully used to improve the visual quality of underwater images [5][6]. CNN (Convolutional Neural Network) and GAN (Generative Adversarial Network)-based models have

provided notable improvements in restoring degraded photographs and improving visibility for difficult conditions in the underwater environment [7][8]. Additionally, object detection models such as YOLO and Faster-RCNN have provided accurate means of identifying marine animals and underwater structures as well as various obstacles [9] [10]. Because of these images, there is a more efficient and less-obtrusive means of analyzing photographs at any given time to produce a result.

Along with augmentation and detection, depth estimation plays a significant role in helping to understand the three-dimensional structure of the underwater world. Historically, traditional approaches for estimating underwater depth have relied on stereo cameras or other dedicated sensors, which add complexity to the overall system [11][12]. Recently, several novel depth estimation methods based on deep learning have been developed that enable depth mapping from a single image, allowing for estimation of both distances and spatial relationships, without requiring additional hardware [13][14]. The results of monocular depth estimation models have shown considerable capability of producing accurate depth maps, even under severely degraded visibility conditions in the underwater environment [15][16]. Combining the two techniques into one framework will enable the complete analysis of the underwater environment using just visual information.

While advancements have been made in image enhancement, detection, and depth estimation, most current systems continue to operate as separate processes; thus, combining the three into one system is rare [17][18]. Furthermore, many image processing solutions still depend partially upon a physical device's inability to adapt to different uses in an actual operating environment [19]. This paper proposes an all-software AI platform that combines image enhancement, object detection, and depth mapping into a single pipeline to address these limitations and create a less expensive, more scalable, and ultimately effective underwater monitoring solution without the use of sensors or IoT equipment [20].

## II. LITERATURE SURVEY

The challenges present in underwater image processing are mostly due to light absorption, scattering, and distortion of colors in an aquatic environment, leading to researchers Past methods primarily used traditional image enhancement techniques such as histogram equalization, white balancing, and fusion to increase contrast and ability to view images [1][2], but they were limited in restoring natural colors and fine details in very degraded images of this nature [3]. To resolve the issue of restoring colors and details, new methods were produced that used a physical model-based approach that looks at how light travels through water and the loss of brightness in the process [4][5].

Through the advance of deep learning, Convolutional Neural Networks (CNNs) have emerged as the predominant method for enhancing these types of images. CNNs are capable of

learning the complex mapping relationships that exist between a degraded version of an image and an enhanced copy of that image by being able to enhance brightness, contrast, and color balance [6][7]. Numerous types of research have presented images applying fusion-based CNN architecture as well as using an end-to-end learning framework with significantly improved performance over traditional methods [8][9]. New types of models such as transformer-based and hybrid deep learning will be developed to further improve enhancement performance to allow for both local and global feature-based extraction for enhanced performance in complex underwater images [10][11].

The use of Generative Adversarial Networks (GANs) for enhancing images captured underwater has become a focus of research among many researchers in image processing. Different models of GANs, like CycleGAN and WaterGAN, create new images by learning the distribution of real high-quality underwater images and produce highly realistic and visually pleasing results [12][13]. GANs are especially useful for working with unpaired datasets and are good at restoring natural color differences. Further, there are many more GAN models that have been developed to generate images through the integration of physical models, attention mechanisms, and diffusion models that will improve performance and generalization [14][15]. Although they provide excellent results, GAN-based methods usually require large amounts of data and significant computational capabilities, which are still a challenge to overcome [16].

Besides improving the quality of images through enhancement, deep learning-based algorithms have also been used to detect objects located underwater. High accuracy has been achieved by using various algorithms such as YOLO, Faster R-CNN, and SSD to locate marine life, underwater structures, and underwater obstructions [17][18]. However, the ability of object detection algorithms to detect objects is strongly dependent on the quality of the input images, which means enhancing input images is a necessary preprocessing step in order to use an object detection algorithm. Recent studies have recommended creating a combined framework for enhancement and detection so that the accuracy of object detection can be increased in low-light conditions [19].

Depth estimation is another important area in underwater monitoring systems. Stereo vision or special sensors are used in traditional methods, which makes the system bigger. complexity and cost [20]. Recent advancements in monocular depth estimation using deep learning have enabled depth prediction from a single image, eliminating the need for hardware sensors. Some models implicitly learn depth information during image enhancement and restoration processes, providing additional spatial understanding [13][20]. However, integrating depth estimation with enhancement and detection into a single system remains an ongoing research challenge.

Overall, the literature indicates significant progress in individual areas such as image enhancement, object detection, and depth estimation. However, most existing systems focus on separate modules rather than a unified software-based solution. This gap highlights the need for an integrated framework that combines all three components—enhancement, detection, and depth mapping—into a single efficient system for underwater monitoring.

### III. PROPOSED SYSTEM ARCHITECTURE

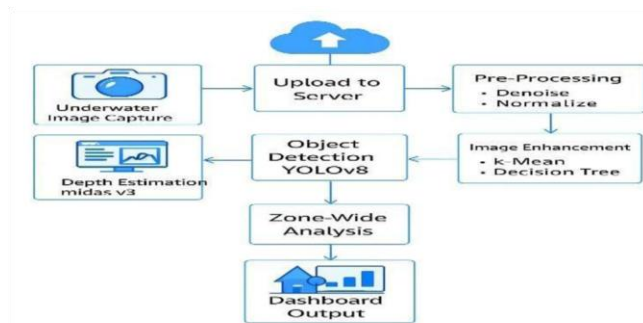


Fig. 1. System Architecture

In summary, the architecture of the AI-based underwater monitoring system consists of a software-driven pipeline for processing underwater images in intelligent stages. Each block in the diagram corresponds to a module in the system, which operates sequentially to provide an accurate and meaningful output.

#### A. Input Layer

The first stage (input) consists of the data sources for the system, which are:

- Underwater images
- Video streams
- Image datasets

These images/videos will be collected and used as inputs to the system. As the system is “purely” software-based, it will not use any physical sensors but will rely only on visual data from the images/videos.

#### B. Image Enhancement Layer

The second stage of processing involves enhancing the quality of the underwater images before the underwater images are eventually processed by the final stages of the system. In many cases, the quality of an underwater image will be affected by the conditions of the water in which the image was taken.

The primary techniques used to enhance the quality of the underwater images will be:

- Histogram equalization
- Colour correction
- CNN/GAN-based enhancement models

The output of the image enhancement stage will be an “enhanced” and “clear” image for use in accurately detecting and analyzing objects/conditions in the later stages of the process.

#### C. Object Detection

After the image has been enhanced, the object detection module will receive it.

The models that are used in this module include the following:

- YOLO (You Only Look Once)
- Faster R-CNN

This module will detect and classify items such as the following:

- Fish and marine organisms
- Underwater structures
- Obstacles

The system will plot bounding boxes and labels around the detected objects to help the automated system to understand what is underneath.

#### D. Depth Mapping

The second module in the overall design is called Depth Mapping, which provides spatial context to other layers in the overall system.

The techniques used within this layer include:

Monocular depth estimation

Deep learning-based prediction of depth

This program uses learning to make a depth map that shows how far away something is from the camera. Without any extra hardware, the depth map will help you better understand the three-dimensional features of an underwater scene.

#### E. Results and Output Layer

The last block of this module provides processed output in an easy-to-understand (user-friendly) manner.

The output of the results and output layer is as follows:

- 1) Final processed image
- 2) Detected objects identified with labels

### 3) 3D visual display of the depth of the detected objects

The Results and Output Layer can be integrated with a dashboard or in an application interface, allowing users to review and analyze data collected from underwater.

## IV. METHODOLOGY

### A. Complete Methodology

The methodology being proposed is a unified software-driven processing pipeline for underwater imagery. The pipeline is designed to process underwater images through sequential stages of image enhancement, object detection, and depth estimation to extract useful/meaningful information. This system would utilize artificial intelligence and computer vision techniques; thus, there are no hardware components to be included in the overall system. The images collected from the water column can become degraded from light absorption, light scattering, and distortion of the optical spectrum due to distance from the surface. Therefore, this methodology will facilitate the progressive improvement of data quality and interpretability at each stage of processing. Additionally, the structured flow of the proposed system from raw input images to analytical output images creates an efficient, scalable process, suitable for applications in real-world monitoring of underwater environments.

### B. Acquisition and Preprocessing of Data

Data acquisition of underwater images will be completed through the collection of datasets, video frames, and image repositories. Preprocessing will then take place to provide standardized preparation of the images for processing. The variation in image sizes from the data sources as well as differences in levels of noise and variability in lighting conditions at the time each image was taken will result in some significant challenges related to ensuring consistent input image types. The images used in the processing are then resized to a common resolution, reduced levels of noise are applied for the purpose of improving the overall clarity of the images, and the pixel values within the collected images are normalized to facilitate the learning process/efficiency of deep learning models. The color space for appropriate images will also be modified to bring forward the preferred features of interest. This ensures that the images collected have been organized/structured appropriately and optimized for further stages of processing.

### C. Module of Improvements in Overall Quality of Underwater Imagery

The first module of processing improves the overall visual quality of underwater imagery generally characterized by low contrast, haze, and distortion of color (i.e., adjusted). Improvements are made through use of many of the traditional techniques associated with image enhancement and restoration methods that utilize the features of deep

learning, thereby improving visual quality by enhancing visibility via brightness & contrast as well as restoring the loss of color due to the absorption of light through water. The learning process is accomplished using a collection of different advanced models that learn how to map degraded images to high-quality (restored) images; consequently, they can be used to reconstruct lost detail and texture from the environment where the image was taken. This part of the process is very important since the accuracy of the object detection and estimation of depth rely heavily upon the quality of images resulting from the previous processes.

### D. Module of Object Detection

The second module detects and classifies objects present in underwater environments as a result of having been processed and improved based on prior image processing steps. Object detection employs the use of deep learning algorithms. Through the application of these deep learning algorithms, the system can extract meaningful features from the processed images, recognize associated patterns of various object types (i.e., marine organisms, structures, or obstacles), and apply these learned features to identify and label the objects located in the image. Also, the structure of the overall detection process enables the automated interpretation of complex underwater scenes. As a result of using multiple advanced detection models, the system can achieve high levels of accuracy for object detection, even in the most challenging of circumstances. The result of accuracy in detection is a reduced amount of time needed to conduct object identification in underwater environments.

### E. Depth Mapping Module

This module uses only visual data to estimate the distance from objects to a camera to provide spatial awareness underneath the water's surface. There are no physical sensors involved in determining relative distance. Instead, various techniques are used that utilize artificial intelligence methods for estimating the relative depth of a scene by analyzing the visual information (shading, texture definitions, object boundaries, etc.) to learn to predict the distance relations for objects based on the trained data set. Depth relationships allow the creation of a depth map representing the overall 3D structure of the environment. This allows systems to have a method to interpret distances and relationships, which improves the analytical capability of the system and effectiveness when used for underwater exploration or monitoring purposes.

### F. Model Training & Optimization

The system's effectiveness is reliant on trained deep learning models using a dataset with a diversity of underwater data to train from. The models learn to see patterns and features related to the different enhancement, detection, and depth estimations through a method of seeing similarities and differences to the identified features. Data augmentation can

help train the model to have variability and performance across a variety of training data types to improve generalization. Optimization algorithms are used to modify the trained models' parameters by minimizing the error from the model for predicting the correct output to increase accuracy. The model training is performed in an iterative fashion, whereby the performance of the model will be reviewed continuously against the validation data to measure the system's reliability and robustness. Proper tuning of the model parameters and training approach will enable the system to perform optimally across the various depths and conditions encountered while underwater.

### G. Generating Output and Visualization

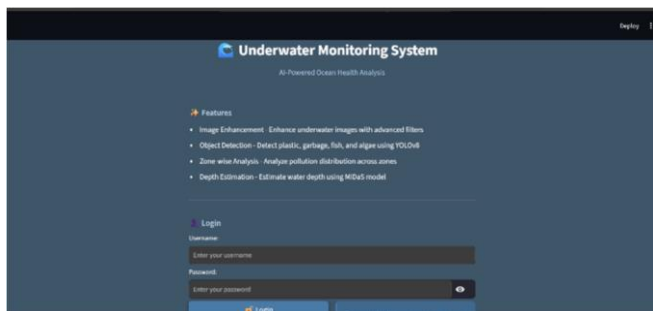
The last stage of the system produces output that combines visual enhancement, object detection results, and depth data into an integrated output. The enhanced image increases the clarity of the image; the results detected show which objects were detected along with their associated label, while the depth map shows the relative spatial relationship between detected objects. The result of combining these different types of data is presented via an intuitive interface to allow for easy interpretation and analysis. Therefore, the results from the system provide a complete understanding of the underwater environment, which can be used as an effective tool for studying and monitoring an underwater environment without using any customized hardware.

### V. RESULT

To assess the potential of the AI-based underwater monitoring system, the researchers utilized commonly available datasets of underwater images and some of their own. Results indicated that the system improved the overall quality of the images being processed, successfully detected objects existing within the water column, and provided estimations of the depth of objects within the underwater scene.

The image enhancement component of the work assisted in improving overall visibility by correcting for color distortions and increasing contrast in the images. Thereby, previously unclear regions of the images were now distinguishable due to the improvements in the image quality. The enhancements to the texture detail and natural color of the images increased the performance of the object detection module in subsequent stages of processing.

The high degree of accuracy achieved through the use of the object detection module to locate and classify underwater objects (including fish, coral structures, and other obstacles) by way of four different types of sensors at a distance or in low-visibility environments enabled researchers/research teams to establish an effective method to determine the location(s) of objects with high levels of confidence and minimal false detections. The addition of depth mapping technology was also successful in mapping the depth of the underwater environment, which provided a viable means of determining the relative spatial relationship(s) of objects in the underwater scene without relying on physical distance.



1: Login Interface of Underwater Monitoring System

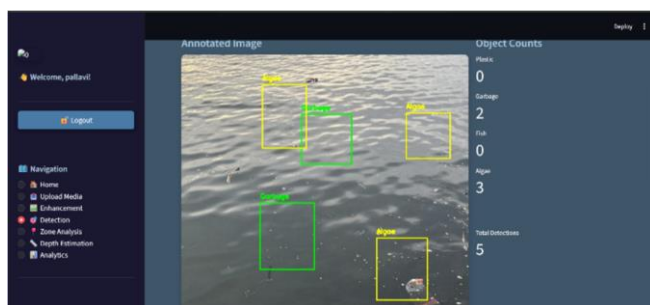


Fig. 5: Object Detection Results Model



2: Dashboard Overview with System Features



Fig. 6: Detection Statistics and Analytical Graphs

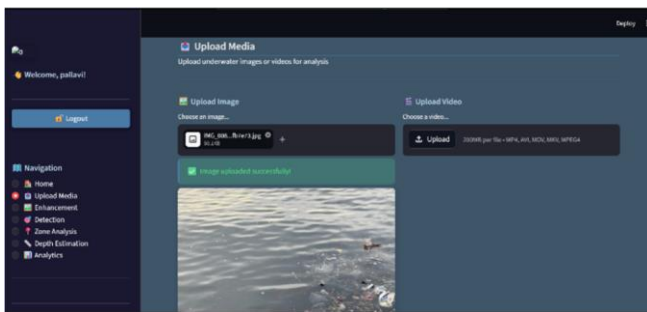


Fig. 3: Media Upload Interface for Image/Video Input

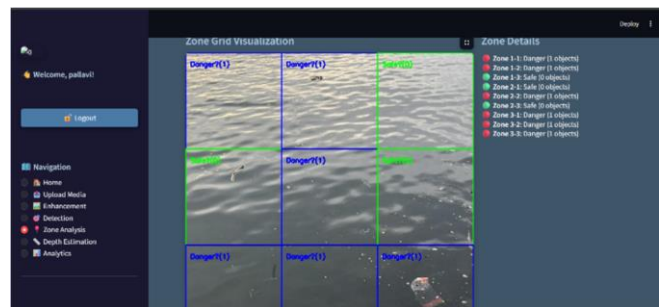


Fig. 7: Zone-wise Pollution Analysis (3x3 Grid Classification)

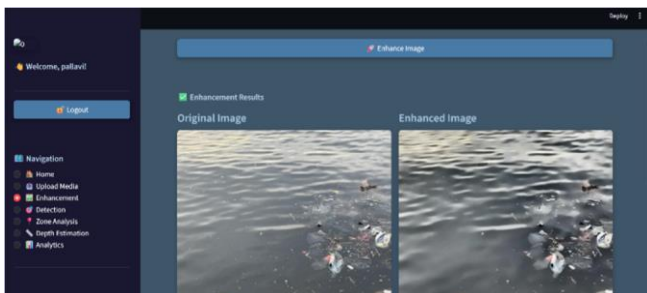


Fig. 4: Image Enhancement Technique

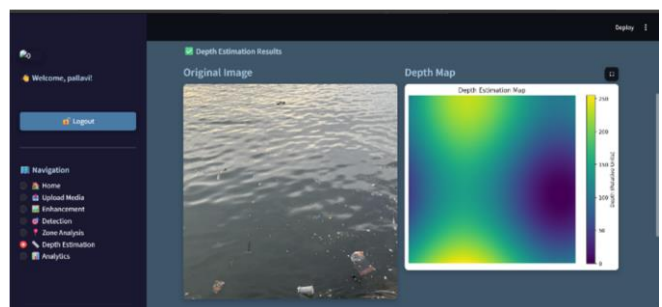


Fig. 8: Depth Estimation Map Model framework. By eliminating the dependency on hardware sensors and IoT components, the proposed approach provides a cost-effective and scalable solution for analyzing underwater environments. The system effectively addresses common challenges such as low visibility, color distortion, and complex scene interpretation by leveraging advanced deep learning techniques. The image enhancement module improves visual clarity, the object detection module accurately identifies underwater entities, and the depth mapping module provides valuable spatial information, enabling comprehensive analysis from a single input image.



Fig. 9: Overall Visualization

Upon integration of image enhancement, object detection, and depth mapping into a single software-based system, a robust and efficient solution was provided. Experimental results indicate that the method proposed here outperforms traditional approaches using only raw images or hardware-based systems. In addition to demonstrating cost effectiveness and scalability, the system demonstrated its ability to operate in real-time and therefore could be applied to practical underwater monitoring applications.

## VI. CONCLUSION

This paper presented a fully software-based AI system for underwater monitoring that integrates image enhancement, object detection, and depth mapping into a unified

The experimental results demonstrate that the proposed system achieves reliable performance across all stages, offering improved accuracy and efficiency compared to traditional approaches. The integration of multiple AI techniques into a single pipeline enhances the overall capability of underwater monitoring systems and reduces the need for manual intervention. In the future, the system can be further improved by incorporating real-time deployment, optimizing model performance, and integrating cloud-based processing for large-scale applications. Overall, the proposed solution represents a significant step toward intelligent, software-driven underwater monitoring systems.

## VII. REFERENCES

- [1] W. Zhang, et al., "Underwater Image Enhancement via Frequency and Spatial Domain Fusion," *Ocean Engineering*, 2025.
- [2] D. C. Lepcha, et al., "An Efficient Underwater Image Enhancement Framework," *Ocean Engineering*, 2025.
- [3] M. Jian, et al., "Underwater Image Processing and Analysis: A Review," *Signal Processing*, 2021.
- [4] T. Liu, et al., "Underwater Depth Estimation for Irregular Illumination Scenes," *Sensors*, 2024.
- [5] S. Xu, et al., "Deep Learning-Based Underwater Object Detection: A Review," *Neurocomputing*, 2023.
- [6] IEEE Signal Processing Society, "Revitalizing Underwater Image Enhancement in the Deep Learning Era," 2023.
- [7] X. Huang, et al., "Underwater Crack Detection Using Enhanced YOLO Models," *Sensors*, 2024.
- [8] S. Meera, et al., "Adaptive Trans-ResUnet++ for Underwater Image Enhancement," *Applied Soft Computing*, 2024.
- [9] "Underwater Image Quality Enhancement and Object Detection Using Deep Learning," 2023.
- [10] R. Wang, et al., "Underwater Image Restoration and Enhancement: A Review," *EURASIP Journal on Advances in Signal Processing*, 2015.
- [11] "Underwater Image Enhancement Using Generative Models," *Journal of Imaging Science*, 2025.
- [12] "Evaluating the Impact of Underwater Image Enhancement on Detection," *arXiv*, 2024.
- [13] T. Li, et al., "Adaptive Color Correction for Underwater Image Enhancement," *Optics Express*, 2022.
- [14] A. Sarala, et al., "Ensemble Deep Learning Model for Underwater Image Enhancement," *Scientific Reports*, 2025.
- [15] X. Cong, et al., "A Comprehensive Survey on Underwater Image Enhancement Based on Deep Learning," *arXiv*, 2024.
- [16] D. Du, et al., "Physical Model-Guided Framework for Underwater Image Enhancement and Depth Estimation," *arXiv*, 2024.
- [17] Q. T. Nguyen, et al., "Enhancing Depth Estimation for Underwater Robots Using Machine Learning," *arXiv*, 2024.
- [18] Z. Huang, et al., "Depth-Guided Perception Network for Underwater Image Enhancement," *arXiv*, 2024.
- [19] Y. Ding, et al., "WaterMono: Self-Supervised Monocular Depth Estimation for Underwater Scenes," *arXiv*, 2024.
- [20] "Underwater Image Processing Using Deep Learning Techniques," *Slogix Research*, 2023.