Sustainable Environmental Monitoring: Multistage Fusion Algorithm for Remotely Sensed Underwater Super-Resolution Image Enhancement and Classification

Wad Ghaban, Jawad Ahmad, *Senior Member, IEEE*, Ali Akbar Siddique, Mohammad S. Alshehri, Anila Saghir, Faisal Saeed, Baraq Ghaleb, and Mujeeb Ur Rehman

***Abstract*—Oceans and seas cover more than 70% of the Earth's surface. If compared with the land mass there are a lot of unexplored locations, a wealth of natural resources, and diverse ocean creatures that are inaccessible to us humans. Underwater rovers and vehicles play a vital role in discovering these resources, yet limited visibility in deep waters and technological constraints impede underwater exploration. To address these issues, advanced image super-resolution and enhancement techniques are crucial for reliable resource identification, species recognition, and underwater ecosystem study. This will ultimately bridge the current gap in environmental monitoring by facilitating resource tracking and underwater waste assessment. This article proposes a novel multistage fusion algorithm for underwater image super-resolution, designed to enhance the quality and spatial resolution of low-resolution underwater images toward a more accurate object characterization. The effectiveness of the proposed super-resolution technique is demonstrated using multiple performance metrics including accuracy, *f*1-score, recall, and precision. By enhancing the spatial resolution of underwater images, our approach meets the increasing demand for detailed and accurate information in underwater earth observation applications.**

***Index Terms*—Image classification, image enhancement, image processing, image recognition, remote sensing.**

Received 26 August 2024; revised 20 October 2024 and 24 November 2024; accepted 21 December 2024. Date of publication 25 December 2024; date of current version 21 January 2025. *(Corresponding author: Jawad Ahmad.)*

Wad Ghaban is with the Applied College, University of Tabuk, Tabuk 47512, Saudi Arabia (e-mail: wghaban@ut.edu.sa).

Jawad Ahmad is with the Cybersecurity Center, Prince Mohammad Bin Fahd University, Al Khobar 31952, Saudi Arabia (e-mail: jahmad@pmu.edu.sa).

Ali Akbar Siddique is with the Department of Computer Science, Iqra University, Karachi 75300, Pakistan (e-mail: ali.siddiqui124k@hotmail.com).

Mohammad S. Alshehri is with the Department of Computer Science, College of Computer Science and Information Systems, Najran University, Najran 61441, Saudi Arabia.

Anila Saghir is with the Department of Telecommunication Engineering, Sir Syed University of Engineering and Technology, Karachi 75300, Pakistan (e-mail: asaghir@ssuet.edu.pk).

Faisal Saeed is with the College of Computing, Birmingham City University, B4 7XG Birmingham, U.K. (e-mail: faisal.saeed@bcu.ac.uk).

Baraq Ghaleb is with the School of Computing, Engineering and the Built Environment, Edinburgh Napier University, EH105DT Edinburgh, U.K. (e-mail: b.ghaleb@napier.ac.uk).

Mujeeb Ur Rehman is with the Institute of Artificial Intelligence, School of Computer Science and Informatics, De Montfort University, LE1 9BH Leicester, U.K. (e-mail: mujeeb.rehman@dmu.ac.uk).

Digital Object Identifier 10.1109/JSTARS.2024.3522202

I. Introduction

OCEANS are crucial components of our planet, hosting an immense variety of marine life and natural resources that remain largely unexplored due to their inaccessibility [1]. While numerous organizations are working to identify these resources, their efforts are hindered by the challenges of underwater environments [2]. Consequently, the utilization of underwater rovers and robots has increased as they can investigate underwater environments for long durations, and identify marine resources without posing any danger to human life [3]. These rovers are equipped with imagery sources capable of capturing images and recording videos such as cameras [4]. Due to this, they are used in many other fields including exploration, surveillance, monitoring, etc. [5]. However, these robots face a multitude of operational challenges even if they are equipped with cutting-edge camera technology [6] [7], [8]. These challenges include low visibility, light refraction, and optical scattering [9]. Enhancing the capability of underwater rovers through image enhancement techniques is essential for effective resource discovery and identification. These techniques can significantly improve image quality, aiding in better classification and informed decision-making, and ultimately enhancing the effectiveness of underwater scientific research and exploration [10]. The underwater environment presents unique challenges for image acquisition, similar to those encountered in atmospheric remote sensing. As light travels through water, it attenuates, causing underwater images to have low contrast, extensive blurriness, and uneven illumination [11], [12], [13]. The intensity of light rapidly decreases as it travels underwater, the rate at which it loses its intensity depends upon the wavelength of the color spectrum [14], [15]. Additionally, particles in the water cause light scattering and reflection, introducing haziness to underwater images [16], [17]. The image enhancement techniques aim to improve the quality of the image by smoothing uniform regions and at the same time maintaining edges to highlight small changes in the brightness in the greyscale channel [18], [19]. Several image enhancement techniques used in underwater settings are similar to those employed in remote sensing SR. Techniques such as color correction, contrast enhancement, and noise reduction are commonly employed to improve underwater

image quality [17], [20], [21], [22], [23]. Many traditional algorithms, including histogram equalization, gamma correction, and adaptive histogram equalization, have been used to enhance images [24], [25], [26]. Additionally, dehazing methods, which reduce haze and fog effects, have proven effective in underwater imaging. The most used approaches for improving underwater image quality are contrast enhancement and restoration with contrast enhancement is preferred by the research community due to its straightforward nature [20], [21].

As discussed, images underwater often contain haziness, muted hues, and limited visibility caused by several factors including light absorption, scattering, and suspended particles. These factors can be tackled with dehazing approach which is extensively used in image processing and computer vision applications capable of reducing the haze and fog effect from the images ultimately improving their quality [25]. Single images often provide limited information for accurate object recognition and scene interpretation in challenging environments. This limitation applies to all remote sensing scenarios including underwater settings. To address this, the integration of multiple image sources through image fusion techniques can significantly enhance our ability to extract meaningful information and improve object recognition in complex scenes [27], [28]. Indeed, the utilization of an appropriate multiscale fusion technique for image correction and enhancement is a key factor that cannot be overlooked and can be useful in overcoming the difficulties posed by underwater settings [29], [30]. Many researchers have introduced multiple fusion techniques that are specifically designed for underwater environments [24], [31]. A few of these techniques include principal component analysis, guided fusion filter, discrete wavelet transform, pyramid fusion, deep learning (DL) fusion, and nonsubsampled contour transform [32], [33], [34]. The Laplacian pyramid fusion and wavelet transform algorithms are the two of the most utilized fusion techniques among researchers and they are well known due to their ability to generate reliable results [35], [36]. The pyramid fusion technique creates image pyramids that represent the original image at multiple resolutions, allowing for a comprehensive understanding of all valuable aspects of the captured image. This method involves producing several down-sampled copies of the original image through an iterative down-sampling process, which are then effectively used in image enhancement. In the context of super-resolution, these image pyramids serve as valuable intermediate representations. They allow the algorithm to capture and preserve details at multiple scales, which is crucial for accurately reconstructing high-resolution imagery from low-resolution inputs. The fusion of information from different levels of the pyramid contributes significantly to enhancing image quality, recovering fine details, and improving overall spatial resolution in earth observation applications [37]. In recent years, artificial intelligence (AI) has emerged as a powerful tool in the field of remote sensing image super-resolution and has been utilized extensively to improve the quality and clarity of images captured in underwater environments. Researchers have developed various DL algorithms to comprehend the complex characteristics of low-resolution underwater images and create high-fidelity reconstructions for enhanced analysis and interpretation [38]. In the domain of DL models, convolutional neural networks (CNNs) and generative adversarial networks (GANs) are the most widely used for image enhancement [39]. CNNs are particularly well-suited for applications that involve classification or object identification due to their ability to recognize distinct patterns and features from an input image [40]. However, it is crucial to note that while CNNs trained on specific types of remote sensing imagery excel at super-resolution tasks for similar data, they may not generalize well to drastically different scenes or sensor types [41]. For instance, a model trained in underwater coral reef imagery might struggle with super-resolving images of deep-sea environments or shipwrecks. Indeed, it is crucial to understand the limitations of the algorithms used, as they may not provide exact solutions for every specific task. Therefore, it is more logical to apply them to applications for which they are best suited [42]. Fig. 1 represents the key enabling technologies in the field of remote sensing and highlighted boxes depict their utilization in the proposed work. In the context of this research, remote sensing refers to the acquisition of information from areas from a distance, mostly without making physical contact. In most cases, it is used in environmental monitoring, meteorology, and mapping applications. Remote sensing generally employs airborne or underwater sensory technologies such as underwater unmanned vehicles in this case for acquiring images using cameras in underwater settings. In this work, the term "remote sensing" is applied more broadly to signify the collection of data from distant or inaccessible locations.

1) The proposed algorithm leverages the multistage fusion technique on a remotely acquired image in an underwater environment.
2) Further enhancement of an acquired image is done using CLAHE and WaterNet to prepare it for classification.
3) Implementation of a DenseNet201 classifier for the prediction.

The rest of the article is organized as follows. Section II reviews recent literature on remote sensing image super-resolution, focusing on underwater applications. Section III details our proposed multistage fusion algorithm for underwater image super-resolution. Section IV discusses the process of data acquisition and augmentation. Section V presents the results, demonstrating our algorithm's performance across various underwater scenarios. Section VI concludes this article and outlines future research directions in underwater remote sensing super-resolution.

II. Literature Review

The exploration of underwater environments through remote sensing is crucial for marine resource discovery, ecosystem monitoring, and climate change studies. However, the complex nature of ocean habitats poses significant challenges for image acquisition and analysis. Consequently, there has been considerable research focused on enhancing underwater remote sensing imagery through super-resolution techniques. This section provides a summary of the latest advancements in this field

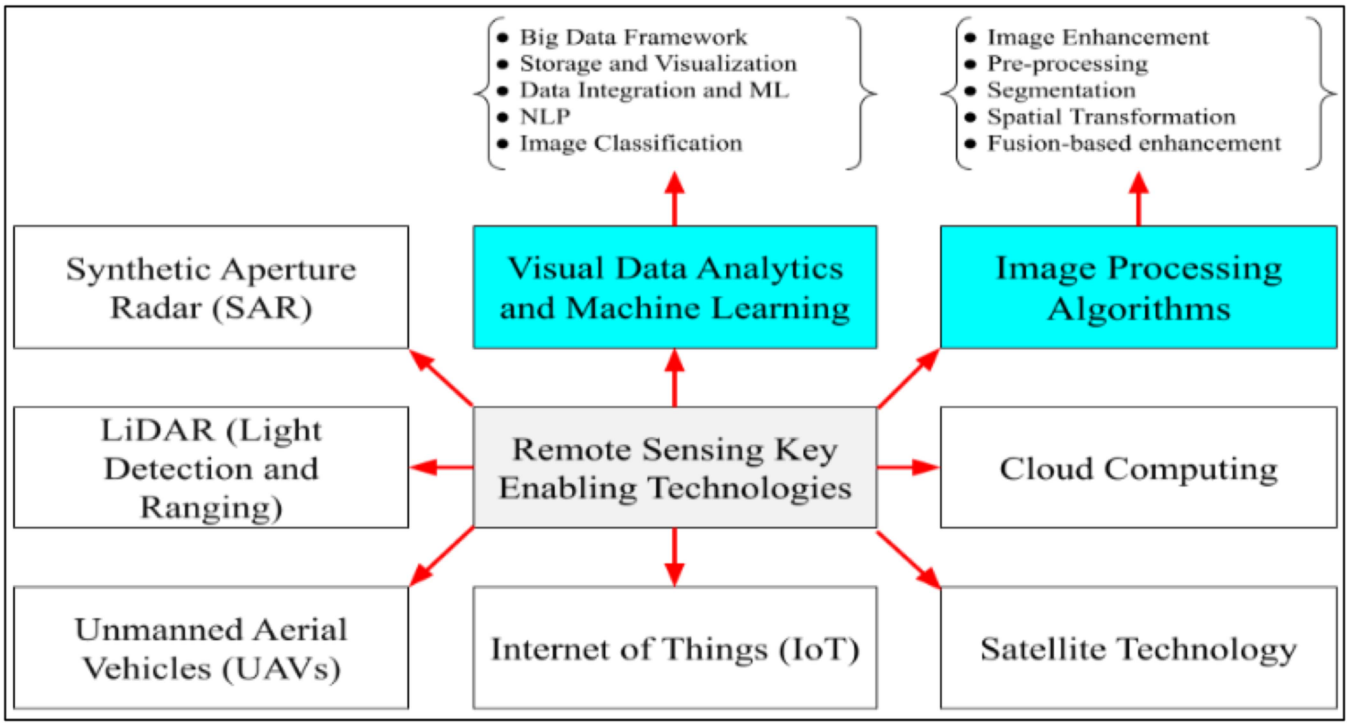


Fig. 1. Computer vision key enabling technologies.

highlighting methods that improve the spatial resolution and quality of underwater images.

Current methods for improving underwater images can be broadly classified into three categories: contrast limited adaptive histogram equalization (CLAHE) [43], white balance correction [44], and Retinex algorithms [45]. Notably, the authors in [46] proposed a fusion-based method for underwater image super-resolution, where inputs and weight measurements are derived solely from the degraded image. An optimized version of this approach is detailed in [47] and incorporates a white-balancing mechanism to further optimize the super-resolution performance in underwater remote sensing applications. In [48], an alternative enhancement strategy that leverages Retinex principles was developed to improve underwater images. The articles [49] and [50] include further investigations that are pertinent to the topic in which they introduced a method for improving images, acquired in lowlight conditions such as at night, underwater, or under artificial lighting with very low contrast known as the 3C. The method eliminates the issue where one-color channel tends to lose most of its information in these circumstances. This is possible by replacing the lost channel with its opponent color channel that minimizes the noise and color shift depending on the image enhancing tasks such as white balancing, dehazing, and underwater image enhancement. The authors are able to provide ample qualitative and quantitative support to explain the use of 3C approach. The second group is comprised of prior-based approaches, which utilize physical imaging models to predict features for enhancing underwater remote sensing imagery. One such method proposes a dehazing algorithm [51], while [52] presents a modified dark channel prior using red channel information to estimate depth maps in underwater scenes. In the context of DL techniques, computers can independently extract representations and generate enhanced mappings by utilizing both paired and unpaired training data. For instance, Li et al. [53] pioneered the use of GANs to generate synthetic underwater photographs, which were then used to train an augmentation network. Additionally, they introduced a weakly supervised method for underwater image super-resolution that does not require paired data. This is achieved through a hybrid loss function that tackles both illumination and detection issues in underwater imagery. There is a lack of accurate historical records about the disposal techniques that were used for DDT acid waste, which has resulted in doubts over whether or not dumping took place via large-scale discharge or inside containers [54], [55], [56]. To shed light on the mapping and characterization of deep-water disposal sites, a systematic technique was used to locate 74 000 garbage targets inside studied areas utilizing Autonomous Underwater Vehicles. This approach utilized advanced image and signal processing techniques in conjunction with machine learning to map and characterize deep-water disposal sites.

In a related research [57], the use of mask region-based CNNs for the identification of garbage in marine environments with little data was investigated. Utilizing artificially produced marine sceneries and the evaluation of situations with varied amounts of training data availability, the study was able to outperform the prior standards in terms of garbage detection performance. Nevertheless, it also brought to light the possibility of further enhancements, while simultaneously noting the difficulties that are related to the restricted availability of data. The study in [58] utilized YOLOV7 and instance segmentation neural networks for object recognition. The study utilized postprocessing techniques to increase the accuracy of the applied masks, and

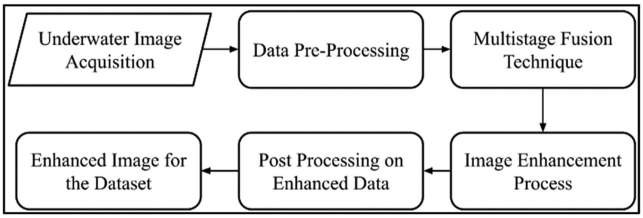


Fig. 2. Proposed image enhancement framework based on multistage fusion technique.

the Monte Carlo method to address exact area approximation within an image. This approach proved effective in applications involving complex objects and high variability, enhancing the segmentation mask application.

The study in [59], leveraged the EfficientNets algorithm to improve the effectiveness of unmanned underwater vehicle vision for the real-time detection of garbage in marine environments. In addition to the provision of a bespoke dataset for the identification of plastic bottles and bags, a significant increase in average accuracy was obtained without the introduction of GPU processing delays. To address the difficulty of properly measuring and removing waste that is buried beneath the sea, researchers in [60] suggested a solution that utilizes a customized dataset and an efficient detection approach leveraging powerful DL architectures. The proposed method presented in this article meets this demand and is equipped with a Graphical User Interface (GUI) that is user-friendly and easy to use. This system can display multistage improvements in an underwater image along with its classification. Usually, images taken in underwater environments are not up to the mark for classification purposes or for any analytical process. Image classification requires a dataset that is more suited for the object or environment as it affects the accuracy of the model if the dataset has images that are dark and unrecognizable. Applying the enhancement procedure proposed in this article improves the quality of the dataset prepared for the classification with results discussed in Table III of Section V.

III. Image Enhancement Framework for Remotely Sensed Data

The multistage fusion technique involves merging the different levels of image pyramid scales. This algorithm is mostly used by researchers to enhance the quality of an image acquired in underwater settings [61]. Fig. 2 depicts the proposed framework for image enhancement, comprising six distinct steps aimed at enhancing image quality and visibility, crucial for classification purposes. The process begins with image acquisition from a sensory source, followed by preprocessing to eliminate noise and several artifacts, such as blockiness arise after the process of compression, color shifting, blurriness, and banding effects. This process is also used to align the image for better classification output. Alignment is the process used to shape the image dataset and prepare it for the classification task. This is done by identifying common features in the acquired image and applying necessary transformations. Subsequently, the preprocessed image data undergoes a multistage fusion to enhance its overall quality, which is repeated for all images in the dataset. Following image enhancement, the acquired data undergoes postprocessing to further refine the image dataset.

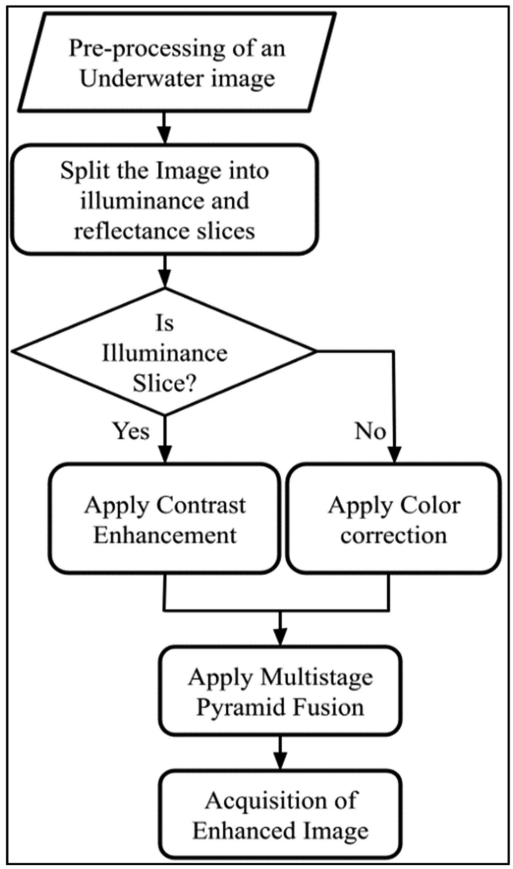


Fig. 3. Control flow representing multistage fusion algorithm.

A. Multistage Fusion Technique

The process of image fusion is performed on the preprocessed data to produce an enhanced version of the image as depicted in Fig. 3. After the acquisition of the preprocessed image, it is split into two components: illuminance and reflectance. The process will check for the components and if it is an illuminance component, contrast enhancement is applied to it, and if not, color correction is applied. These processes are mandatory for enhancement as they are key processes required before the application of a multistage pyramid fusion scheme. Once the contrast-enhanced or color-corrected image completes the multistage pyramid fusion process, the result is a significantly enhanced image.

1) Illuminance and Reflectance Slices: Illuminance refers to the amount of light that falls on the surface per unit area influencing the visibility of objects within a scene or image [62]. It is the representation of the visible light illuminating an area or location when it strikes a specific object. It can be measured by dividing the luminous flux of the light source by the area upon which the light is distributed. This is an important feature that is widely used in fields such as architecture, photography, and workspace design due to its ability to generate appropriate lighting for proper visibility, safety, or even comfort. There are

some factors that can affect the illuminance level including the distance between the light source and the illuminated surface, the characteristics of the light source itself, and environmental conditions in underwater settings. The level of illuminance required mostly depends on the application in which they are used, and often they are introduced as the guideline or the standard for the specific required to optimize lighting used on various occasions. Illuminance serves as the primary concept in lighting design, which can help in producing a well-illuminated environment that can be tailored for specific needs.

Reflectance is the proportion of incident light that is reflected by a surface, influencing the perceived brightness of objects. It quantifies the intensity of light reflected by an object, typically expressed as a percentage or decimal fraction. Reflectance plays an important role in determining the brightness and color of surfaces, as well as the overall lighting quality in a given location. A higher level of reflectance means that the surface is capable of reflecting more light producing a brighter environment and on the other hand low reflectance means the surface is unable to reflect enough light resulting in a much darker environment. Just like illuminance, reflectance is also affected by various factors such as the material properties of the object, surface treatment or coatings, and the angle of incident upon which the light strikes the object [63].

Illuminance and reflectance slices play a vital role in identifying objects as they are extensively utilized for image enhancement tasks in underwater settings [64]. These components are acquired by splitting underwater environment images, in most cases enhancement is performed to gain an image dataset that helps in challenging underwater settings [65]. Lux (lx) is a unit of illuminance denoted by (E), it can be defined as the amount of luminous flux denoted by (Φ) as an incident point on a unit area (A) of the surface where the light hits, this is expressed in (1). Luminance intensity is denoted by (L) given in (2) and it is a key matrix that is mostly utilized to quantify the brightness of the surface perceived by the human visual systems or even from an imagery source such as a camera. It measures the amount of visible light either reflected or emitted from the surface [66].

$$E = \frac{\Phi}{A} \tag{1}$$

$$L = \frac{dI.\cos(\theta)}{DA_p.d\Omega}. \tag{2}$$

Typically, the acquired images are converted into the LAB color space in order to adjust the illuminance in underwater settings. Followed by the process of luminance that is used to control the brightness of the acquired images. Components of the LAB channels given in (3)–(5) are recombined, and they are converted into a BGR color space. At this point, a histogram is calculated to correlate the luminance applied enhanced image and the original image. The correlated output displayed in the heatmap shows that the brighter region indicates a higher degree of correlated information.

YCbCr and LAB color spaces have been incorporated in the proposed algorithm intentionally to maximize the strengths of each. YCbCr intends to make better separation of luminance and chrominance information so as to directly manipulate the brightness. LAB maintains the color adjustment in the perceptual space because of human vision. This approach of using both forms of functions guarantees a more refined enhancement process, thus improved visual quality but at the cost of some extra computational complexities. Therefore, the application of these color spaces improves both spatial and chromatic resolution in the underwater imagery, which justifies their integration into the algorithm. Fig. 4 represented the block representation of Multiscale Pyramid Fusion scheme.

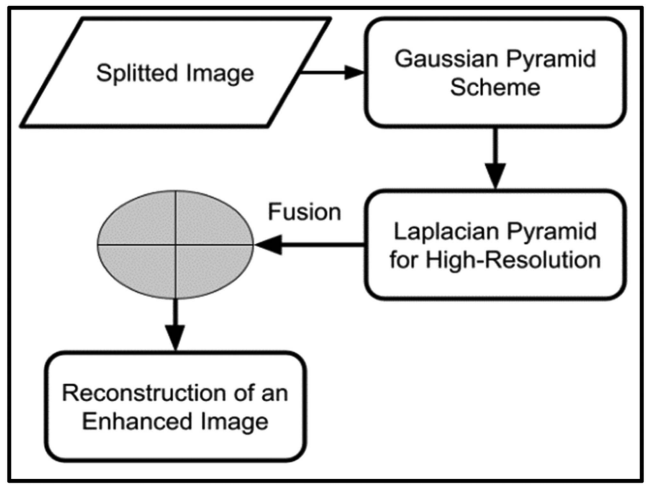


Fig. 4. Multiscale pyramid fusion.

$$L^* = 116 \times f\left(\frac{Y}{Y_n}\right) - 16 \tag{3}$$

$$a^* = 500 \times \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right) \tag{4}$$

$$b^* = 200 \times \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right) \tag{5}$$

wher $L*$ is the component based on lightness. $a*$ is the representation of the green and red components of the image. $b*$ is the representation of the blue and yellow components of the image. X, Y, Z are representations of the tristimulus values in RGB color format. X_n, Y_n, Z_n are representations of the tristimulus values for the reference white point. f(t) is the function.

In (6), $f(t)$ represents the key feature of the LAB color space transformation usually used in image processing applications. Its goal in LAB color space conversion is to normalize the color information based on light intensity for calculating L$*$, a$*$, and b$*$ values.

$$f\ (t) = \begin{cases} t^{1/3} & \text{if } t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{29}{6}\right)^2 t + \frac{4}{29} & \text{Otherwise} \end{cases}. \tag{6}$$

2) Linear Piecewise Color Correction: Color correction is a technique that is widely used in the field of image processing to adjust the colors of an image and achieve the desired contrast. This can be done by manipulating color balance, contrast, and brightness.to achieve a more visually appealing image suitable for classification. The key aim of color correction is to ensure

that the colors in an image are equally distributed, accurate, and consistent. Color correction can be performed on various color schemes such as RGB (red, green, blue) color space, CMYK (cyan, magenta, yellow, black), etc. Each color has a different representation and can be used in a variety of distinct applications that are most suited for them. A common method for color correction involves using a color correction matrix to linearly transform the color scheme of an image, as described in (7). This process involves two techniques: reflectance and illuminance, as outlined in (8). The method accepts reflection slices as input and applies a transformation based on the individual pixel intensity represented in the images. For the intensity value less than 50, it scales the intensity factor of 1.8 and applies linear transformation for the intensity of 2.15 plus 20 for the value greater than 50.

$$\begin{bmatrix} R_{out} \\ G_{out} \\ B_{out} \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} R_{in} \\ G_{in} \\ B_{in} \end{bmatrix} \tag{7}$$

$$y\ (x) = \begin{cases} 1.8x & \text{if } x < 50 \\ 2.15 + 20 & \text{if } x > 49 \end{cases}. \tag{8}$$

OpenCv library is used to apply contrast enhancement at this stage which is the most widely used Python tool for image processing tasks. Histogram equalization is performed to elevate the parts of the image with respect to its surrounding pixels. The enhanced intensity value z(x) can be easily determined using the following equation:

$$z\ (x) = \frac{H\,(x) - H_{\min}}{M \times N - 1}\ \times (L-1)\,. \tag{9}$$

In the initial stage, a Gaussian pyramid is created for the high-resolution images. This pyramid usually contains the down-sampled version of the original image to provide better results using enhancement techniques. Following the same process, a Gaussian pyramid is created for the low-resolution image. Gaussian blurring and downsampling are applied iteratively to create this pyramid for low-resolution images as depicted in (10) and (11). The mathematical representation of Gaussian blurring is given in (12). Each pixel in the blurred image is computed as a weighted sum of the surrounding pixels in the original image. Pixels farther from the central pixel (x, y), have smaller weights, resulting in a smooth, blurred effect.

$$G_{high}\ (i,j) = GaussianBlurr\,(G_{high}\,(i,j)) \tag{10}$$

$$G_{low}\ (i,j) = GaussianBlurr\,(G_{low}\,(i,j)) \tag{11}$$

$$I_{Blurr}\ (x,y) = \sum_{i=-k}^{k} \sum_{j=-k}^{k} \frac{1}{2\pi\sigma^2} e^{-\frac{i^2+j^2}{2\sigma^2}}.I\,(x+i, y+j) \tag{12}$$

where $I_{Blurr}(x, y)$ is the intensity of the blurred image at the pixel location (x, y). $I(x + i, y + j)$ represents the intensity of the original image at the pixel location $(x + i, y + j)$. k is the extent of the kernel in both dimensions. σ is the standard deviation of the Gaussian distribution, controlling the amount of blur. e is the base of the natural logarithm.

The Laplacian pyramid is created at this stage for high-resolution images by calculating the difference between each level of the Gaussian pyramid and its expanded version, as shown in (13). This acquires the exact set of Laplacian images that captures distinct levels of scales. The Laplacian pyramids for both high- and low-resolution images are then combined in the fusion stage using a weighted sum, with a weight of 0.5 for each individual image, to produce an enhanced final image as shown in (14). The final image is obtained by utilizing the fused Laplacian pyramid as described in (15) and (16). The Gaussian pyramid relies on applying Gaussian filtering to an image to smooth it, followed by down-sampling, which reduces the image size at each level. This process creates a multiscale representation of the image, allowing the algorithm to process features at different resolutions. The pyramid structure helps in handling varying levels of detail, which is particularly useful in image enhancement and analysis tasks like those in the proposed algorithm.

$$L_{high}\ (i,j) = G_{high}\ (i,j) - Expand(G_{high}\,(i,j)) \tag{13}$$

$$F_{lap}\ (i,j) = 0.5 \times L_{high}\,(i,j) + 0.5 \times L_{low}\,(i,j) \tag{14}$$

$$R_{lap}\ (i,j) = Expand(F_{lap}\ (i,j)) \tag{15}$$

$$Enhance\ d_{Image} = \sum_{i} R_{lap}\,(i,j)\,. \tag{16}$$

B. Postprocessing (Multistage Image Enhancement)

Postprocessing is the final stage of the multistage fusion technique. In this phase, additional enhancement algorithms are applied to improve image quality and eliminate any remaining noise.

1) CLAHE Based Image Enhancement: CLAHE is an enhancement technique used by researchers to enhance underwater images. Literature on CLAHE-based algorithms proves its effectiveness in underwater settings [67], [68], [69], [70].

2) WaterNet-Based Image Enhancement: The challenges associated with identifying water bodies in images arise from the diverse nature of water types, resulting in distinct scattering effects that may hide boundaries and key features. To tackle these problems, the proposed framework employs channel-wise normalization algorithms. Typically, image channels tend to vary depending on their location and may contain useful information regarding the underlying structure. By using channel-wise normalization, this technique helps to minimize image deformation and degradation [71].

WaterNet, also known as the water body detection network, is a sophisticated DL framework specifically developed to identify water bodies precisely and effectively in satellite or aerial images. It is also used to enhance image quality for subsequent analysis or classification. The proposed algorithm utilizes CNNs to accurately classify objects in underwater environments on which it is trained [72], [73].

C. Image Classification Based on DenseNet201

After the acquisition of the enhanced image, the proposed algorithm implements DenseNet201 [74], a CNN-based pre-trained classifier to identify objects in the water. Basically, DenseNet201 is a conventional DL approach that makes use

Algorithm 1: Algorithm for CLAHE-Based Image Enhancement.

1. **Input**: Image Acquisition *I(x, y)*
2. **Pre-Processing:** Greyscale conversion
   $Y = 0.299 \times R + 0.587 \times G + 0.114 \times B$
3. **Tile Division:** Division of the tiles in an image into a non-overlapping, similar to a macroblock

```
def divide_into_tiles(image, grid_size):
# Get the dimensions of the image
image_height, image_width = image.shape
# Calculate the number of tiles in each dimension
num_tiles_vertical = image_height // grid_size
num_tiles_horizontal = image_width // grid_size

tiles =[]

for i in range(num_tiles_vertical):
  for j in range(num_tiles_horizontal):
    # Calculate the coordinates for the current tile
    tile_y_start = i * grid_size
    tile_y_end = (i + 1) * grid_size
    tile_x_start = j * grid_size
    tile_x_end = (j + 1) * grid_size

# Extract the current tile from the image
Tile = image [tile_y_start:tile_y_end,
 tile_x_start:tile_x_end]

# Append the tile to the list of tiles
tiles.append(tile)
return tiles
```

4. **Tile-level Processing:**
   for *i* in *Q[I]*: # Q is an array of images and I represent an image of that array.
   - Calculate Histogram to identify pixel intensity at each level.
   - Histogram Equalization
     - Identify Cumulative Distribution Function (CDF).
     - Assign a clip limit and clip the CDF.
     - Mapping of pixel values present in the tile utilizing the modified CDF.

5. **Recombination:** Assemble all the tiles to begin the reconstruction of an enhanced image.
6. **Enhanced Image acquired**.

Algorithm 2: WaterNet Basic Algorithm for Image Enhancement.

1. **Input:** Image Acquisition *I(x,y)*
2. **Pre-Processing:** Channel adjustment, resizing, and color conversion to RGB

3. **CNN:** Initialize CNN Architecture

```
for i in range(1, N): # N is the number of Layers
  F_i = Convolution(F_{i-1}, W_i) + b_i
  F_i = ReLU(F_i)
  F_i = MaxPooling(F_i) if downsampling
  F_i = Upsample(F_i) if upsampling
  # F_i represents feature map
```

4. **Feature Extraction**

   for i in range(1, N):
   $Z_i = F_{i-1} * W_i + b_i$ # W_i are learnable weights
   $F_i = Activation\ (Z_i)$ # b_i is the bias for i^{th} layer

5. **Skip Connections:**
   - Initialize skip connections between encoder and decoder paths.

6. **Computing Loss Function:**
   $L = MSE(I_{enhanced}, I_{ground\ truth})$

7. **Optimization:** Initialize network parameters (weights '*W*' and biases '*b*')**.**

   for i in t: # t is the number of iterations.
   $y_{pred} = ForwardPropagation(I_{input},\ W,\ b)$
   $W = W - \alpha * \partial W$
   $b = b - \alpha * \partial b$
   # Check for the convergence

8. **Output:** Obtain the Enhanced Image $I_{enhanced}$

of neural network architecture and widely known for its dense connectivity and robust parameter utilization [75]. It features dense blocks where each layer is connected to every other layer in a feed-forward manner, using feature map concatenation as inputs for subsequent layers rather than the traditional addition and concatenation methods. This facilitates the promotion of feature reuse and the flow of information [76], [77]. The variable (k) in (17) demonstrates the growth rate which determines the inclusion of new features in the network to identify.

$$H_l = H_{l-1} \quad BN\left(ReLU\left(Conv_3 \circ \right.\right. \\ \left.\left. \times BN\left(ReLU\left(Conv_{1x1}\left(H_{l-1}\right)\right)\right)\right)\right). \tag{17}$$

Batch normalization and activation layers are the two elements of the training process provided to the network. ReLU is used to introduce nonlinearity in the data that further expands the feature map for better classification as given in (18). Table I represents the architecture of the DenseNet201 model used for the classification.

$$H_{l+1} = Transition\ (H_l) \\ = BN\left(ReLU\left(Conv_{1x1}\left(H_l\right)\right)\right) \circ AvgPool\left(2x2\right)\left(H_l\right). \tag{18}$$

IV. Data Acquisition and Preprocessing With Respect to Ocean Dynamics

Typically, a dataset requires augmentation as it may initially not be enough for training. A few steps are involved in the

TABLE I
ARCHITECTURE OF A DENSENET201 CLASSIFIER

Layer	Composition	Output size
Input	–	224×224
Convolution	Conv(7×7), stride 2	112×112
Pooling	(3×3), stride 2	56×56
Dense block 1	Conv(1×1) Conv(3×3)	56×56×(6)
Transition layer 1	Conv(1×1) Avg pool (1×1), stride 2	56×56 28×28
Dense block 2	Conv(1×1) Conv(3×3)	28×28×(12)
Transition layer 2	Conv(1×1) Avg pool (1×1), stride 2	28×28 14 ×14
Dense block3	Conv(1×1) Conv(3×3)	14×14×(48)
Transition layer 3	Conv(1×1) Avg pool (1×1), stride 2	14 ×14 7×7
Dense block4	Conv(1×1) Conv(3×3)	7×7×(32)
Classification layer	7×7 Avg pool, 1000D fully connected layer, softmax	1×1

process of augmentation, such as scaling, rotations, flipping, and noise suppression. Utilization of these preprocessing methods enhances the performance of the trained model and makes it much more flexible and accurate [78]. The dataset used to train and validate the model is publicly available on Kaggle [79]. It comprises over 3500 images categorized into two distinct classes: images depicting plastic waste and those showing no plastic waste in water. This classification highlights the environmental threat posed by plastic waste, which impacts marine life and contaminates their habitats. The dataset is divided into training and testing sets, with 880 images in the no-plastic category and 840 in the plastic category. We utilized 2900 images for the DL model's training and 300 images for validation as detailed in Tables II and III. Augmentation plays a crucial role in expanding the dataset, which significantly aids in developing a more generalized and effective model. It involves various techniques to generate different images using the same dataset including rotation, alteration in brightness, flipping, etc. [80]. Addressing class imbalance is another challenge in model training, and augmentation can help mitigate this issue by balancing the dataset.

Fig. 5 demonstrates the augmented datasets used to train the Denset201 Classifier. Table II shows the number of instances in

TABLE II
ACTUAL AND AUGMENTED DATASET

Classes	Actual dataset	Augmented dataset
Plastic	840	1600
No-plastic	880	1600
Total	1720	3200

TABLE III
DISTRIBUTION BETWEEN TRAINING AND VALIDATION SET

Classes	Training set	Validation set
Plastic	1450	150
No-plastic	1450	150
Total	2900	300

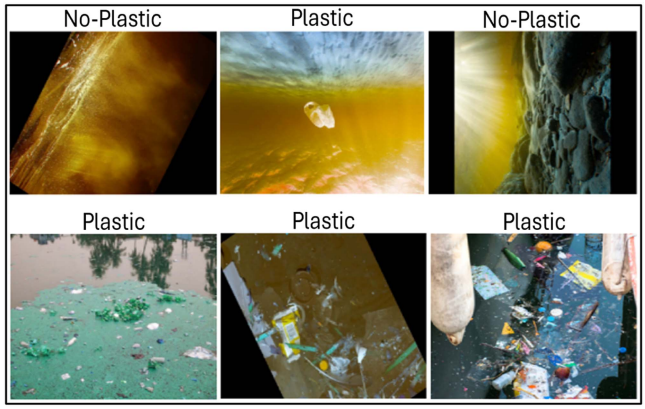


Fig. 5. Sample images of the augmented data for the purpose of classification.

the dataset before and after the augmentation process showing an increase in the number of the total instances from 1720 images to 3200 images. To ensure consistency, it is crucial to uniformly scale the dataset through normalization, which ensures that all images have the same resolution. This can be achieved using the method described in the following equation:

$$Z_N = (Z - Z_{\min}) / (Z_{\max} - Z_{\min}). \tag{19}$$

The process of rotation is applied to an image to augment it and acquire images containing the same content but slightly rotated at a certain angle, this process is depicted in the following equation:

$$r(\theta) = \begin{bmatrix} \mathrm{Cos}(\theta) & -\mathrm{Sin}(\theta) & 0 \\ \mathrm{Sin}(\theta) & \mathrm{Cos}(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{20}$$

In the domain of image processing, it is possible to apply techniques given in (21) to transform it in the horizontal or

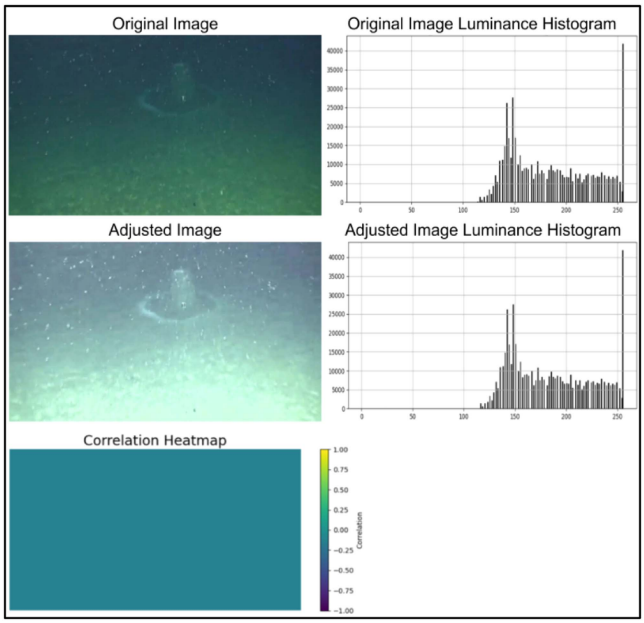


Fig. 6. Original images and acquired enhanced images based on luminance using the factor of 2.15.

vertical direction while keeping all of its features.

$$t\ (w,h) = \begin{bmatrix} 1 & 0 & w \\ 0 & 1 & h \\ 0 & 0 & 1 \end{bmatrix}. \tag{21}$$

In the proposed work, all computations were performed using a workstation with 2.9 GHz Intel Core I5-9499F CPU (6 cores), 32 gigabytes of random access memory, and an NVIDIA GeForce GTX 1660 SUPER. Since the algorithm requires both enhancement and classification tasks, which are computationally intensive, it is necessary to use high-end hardware to perform adequately. For the current scenario, the computations were handled efficiently by the hardware and no bottlenecks were reported during the entire process. However, if we want to scale the proposed algorithm to much larger dataset, then the process may require more advanced hardware or distributed computing solution to maintain the processing needs.

V. Results

To obtain an enhanced image, preprocessing techniques are first used to align the images, which are then fused together. The image mentioned in Fig. 6 describes the outcomes of an image enhancement process. The "original image" taken in an underwater environment has a low luminance and is blurry, the graph under the luminance histogram shows a large number of pixel brightness values at the lower end of the spectrum. On the other hand, the "adjust image" center part demonstrates a much brighter and balanced colored image, free from hazy or foggy appearance, making the visual much sharper. The histogram of the corresponding luminance for the adjusted image shows that there is a shift of the luminance toward the higher end of the scale which means that the distribution of luminance within the image will be much more even. Furthermore, the correlation heatmap

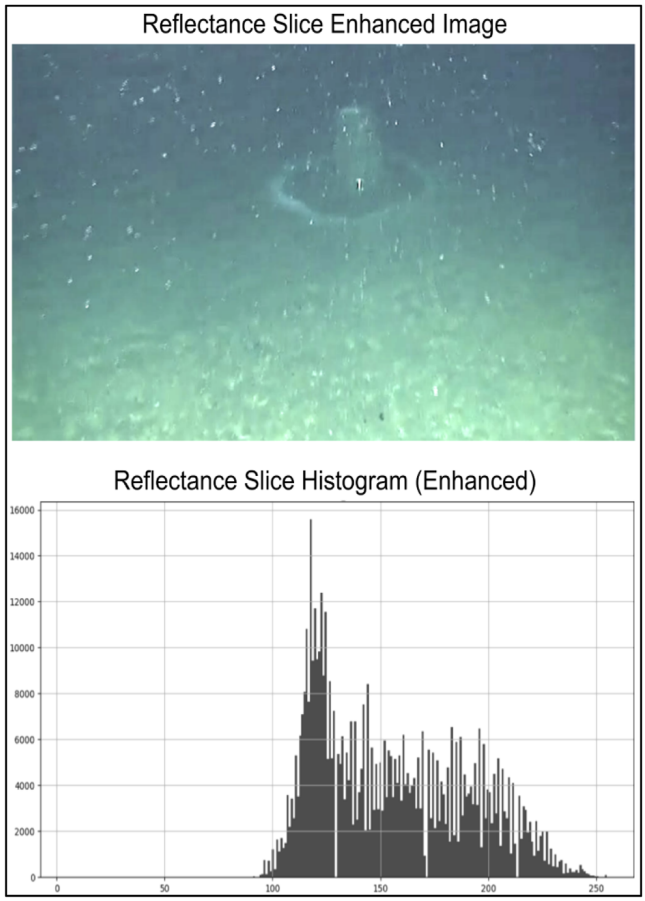


Fig. 7. Enhanced reflectance slice along with its histogram.

provided at the bottom of the visual shows the correlation of color channels before and after the adjustment, indicating the efficacy of the enhancement technique applied in the current project for image enhancement. The luminance factor is set to 2.15 in which each individual value of the luminance image gets multiplied by this value and in return, the acquired image is brighter making the dark areas in the underwater environment more visible.

In Fig. 7, the image that has been provided showcases the underwater image after applying a reflectance slice enhancement. The image at the top, which is the "reflectance slice enhanced image," demonstrates a great improvement in the visibility and color balance as compared to the unenhanced underwater image providing a better view of the underwater territory. Below this is the "reflectance slice histogram (enhanced)" which shows the histogram of pixel intensities of the enhanced image. Analyzing the histogram, it can be stated that the image has high contrast and is characterized by a wide range of luminosity while having a peak of mid-tones (100–150), which means that the reflectance is distributed equally. Such distribution can be explained by the efficiency of the enhancement technique to restore details and fix color distortions, thus delivering an image that is more aesthetically pleasant to the viewer.

Real-time management of the proposed algorithm has specific difficulties. First, the dual optimization of enhancement and

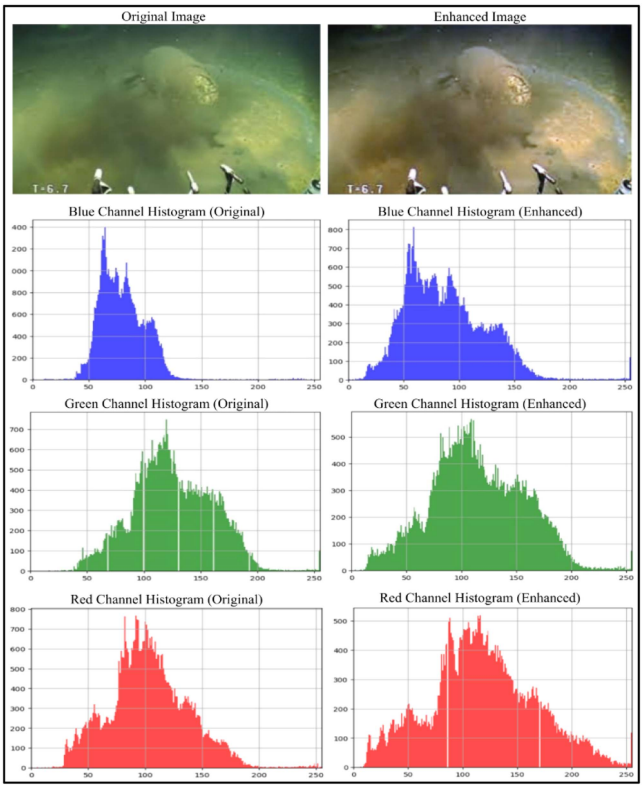


Fig. 8. Gaussian pyramid-based enhanced image acquired and individual histograms of each RGB channel for both enhanced and original images.

classification tasks exponentially increases the computational load and needs resource management. Making sure the results of data processing by the algorithm meet the real-time requirement, particularly when dealing with super-resolution images, requires optimal use of resources in both software and hardware aspects.

Another defining factor is maintaining the balance between processing speed and the quality of the image produced by the enhancement process. It is a fact that real-time applications like the proposed technique require low latency but at the cost of the algorithm's performance if not properly addressed. The ability of the proposed algorithm to adjust to the changing environment in real time, without manual intervention, is essential to maintain robustness in practical applications.

To effectively address these challenges in real-time applications, it is crucial to optimize the code for faster execution, leverage parallel processing, and ensure that the algorithm scales efficiently as data complexity increases or input volumes grow larger.

To acquire the final enhanced output image, we applied a multistage fusion algorithm incorporating both Laplacian and Gaussian Pyramid. It is done by reshaping the construction of the pyramids and applying Gaussian upsampling to each of its levels. At this point, all the levels can be added to make an enhanced image as shown in Fig. 8.

Fig. 9 represents the training and validation accuracy as well as the loss for the DenseNet201 model trained on the enhanced dataset. The model is trained over the period of 30 epochs and at every epoch full dataset is used for training. During training, a

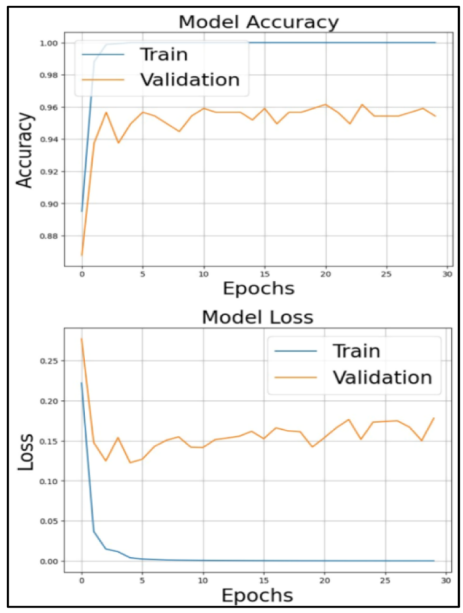


Fig. 9. Model training and validation accuracy and loss for the enhanced image dataset.

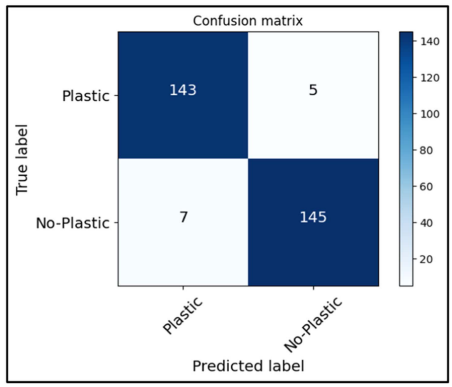


Fig. 10. Confusion matrix of the fusion-enhanced model.

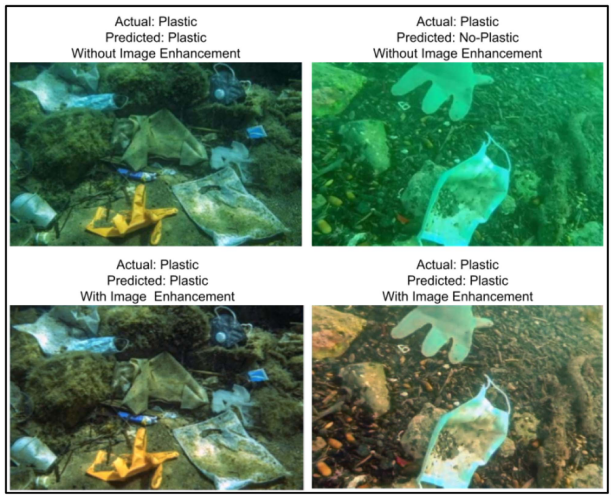


Fig. 11. Depiction of the classification process both with and without the application of image enhancement process.

TABLE IV
PERFORMANCE OF THE MODEL WITH AND WITHOUT FUSION ENHANCED DATASET

Classes	With enhanced dataset	Without enhanced dataset
Accuracy	0.9633	0.9103
***f*1-score**	0.9636	0.907
Recall	0.954	0.893
Precision	0.973	0.912

trend of decreasing loss and increasing accuracy was observed, reflecting the model's performance. The model achieved an accuracy of 0.954 which is quite good for the images used in this particular application based on an underwater environment.

In Table IV, it is evident that the model performed significantly better on the fusion-enhanced dataset compared to the original, nonenhanced dataset. The proposed algorithm achieved an accuracy of 0.9633 on the enhanced dataset, outperforming the 0.9103 accuracy obtained on the original dataset. During training, a clear trend of decreasing loss and increasing accuracy was observed, further demonstrating the model's effectiveness. Notably, the model successfully predicted 143 instances for the "no-plastic" class, as illustrated in the confusion matrix in Fig. 10, highlighting its robust classification capabilities.

The performance breakdown of the proposed algorithm with and without the enhancement step for underwater images is shown in Fig. 11. The first row of images demonstrates the results of the classification with no image enhancement, for several images the classifier predicts the existence of nonplastic materials while in fact there is plastic material present in the frame which is clearly visible. However, in the bottom row, figures represent images processed by the enhancement algorithm where the classifier correctly recognizes the plastic materials. The enhancement step is shown to be greatly beneficial in increasing visibility and overall image clarity of underwater settings, thus enhancing classification results.

VI. CONCLUSION

A major obstacle for underwater exploration robots or rovers is capturing clear images in these environments challenging conditions, thus, equipping them with image enhancement tools is crucial. To overcome this, this article presents a novel multistage fusion algorithm for underwater image super-resolution, addressing the existing gaps in current solutions. Our approach combines Gaussian and pyramid-based fusion techniques with advanced postprocessing methods, including CLAHE and WaterNet-based enhancement, to significantly improve the spatial resolution and quality of underwater images resulting in high-quality images suitable for classification. The enhanced images are then classified using DenseNet201, a state-of-the-art DL model showing promising results demonstrating the effectiveness of our super-resolution method. This innovative approach offers researchers a valuable tool for studying diverse underwater ecosystems, facilitating more effective marine resource exploration, and environmental monitoring. We aim to explore further optimizations in the future to improve the real-time performance, especially in constrained environment. Also, integrating state-of-the-art machine learning algorithm for adaptive enhancement may enhance the versatility across many other similar applications.

REFERENCES

[1] R. N. Hughes, D. Hughes, and I. P. Smith, "Oceans and marine resources in a changing climate," *Oceanogr. Mar. Biol.: Annu. Rev.*, vol. 51, pp. 71–192, 2013.

[2] B. C. McLellan, "Sustainability assessment of deep ocean resources," *Procedia Environ. Sci.*, vol. 28, pp. 502–508, 2015.

[3] D. Chiarella et al., "A novel gesture-based language for underwater human–robot interaction," *J. Mar. Sci. Eng.*, vol. 6, no. 3, pp. 91–109, 2018.

[4] X. Lin, J. Wu, and Q. Qin, "A novel obstacle localization method for an underwater robot based on the flow field," *J. Mar. Sci. Eng.*, vol. 7, no. 12, pp. 437–449, 2019.

[5] X. Chen et al., "Underwater image enhancement based on deep learning and image formation model," 2021, *arXiv:2101.00991*.

[6] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robot. Automat. Lett.*, vol. 5, no. 2, pp. 3227–3234, Apr. 2020.

[7] H. Lu, Y. Li, and S. Serikawa, "Underwater image enhancement using guided trigonometric bilateral filter and fast automatic color correction," in *Proc. IEEE Int. Conf. Image Process.*, 2013, pp. 3412–3416.

[8] M. J. Islam, M. Ho, and J. Sattar, "Understanding human motion and gestures for underwater human–robot collaboration," *J. Field Robot.*, vol. 36, no. 5, pp. 851–873, 2019.

[9] S. Zhang et al., "Underwater image enhancement via extended multi-scale Retinex," *Neurocomputing*, vol. 245, pp. 1–9, 2017.

[10] I. Bae and J. Hong, "Survey on the developments of unmanned marine vehicles: Intelligence and cooperation," *Sensors*, vol. 23, no. 10, 2023, Art. no. 4643.

[11] J.-D. Tournier et al., "MRtrix3: A fast, flexible and open software framework for medical image processing and visualisation," *Neuroimage*, vol. 202, 2019, Art. no. 116137.

[12] M. Salvi et al., "The impact of pre- and post-image processing techniques on deep learning frameworks: A comprehensive review for digital pathology image analysis," *Comput. Biol. Med.*, vol. 128, 2021, Art. no. 104129.

[13] M. Jian et al., "The extended marine underwater environment database and baseline evaluations," *Appl. Soft Comput.*, vol. 80, pp. 425–437, 2019.

[14] I. Ullah et al., "Efficient and accurate target localization in underwater environment," *IEEE Access*, vol. 7, pp. 101415–101426, 2019.

[15] J. R. V. Zaneveld and W. Scott Pegau, "Robust underwater visibility parameter," *Opt. Express*, vol. 11, no. 23, pp. 2997–3009, 2003.

[16] E. Trucco and A. T. Olmos-Antillon, "Self-tuning underwater image restoration," *IEEE J. Ocean. Eng.*, vol. 31, no. 2, pp. 511–519, Apr. 2006.

[17] C. Li et al., "An underwater image enhancement benchmark dataset and beyond," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 4376–4389, Nov. 2019.

[18] A. A. Siddique, M. Tahir Qadr, and Z. Mohy-Ud-Din, "Masking of temporal activity for video quality control, measurement and assessment," *Meas. Control*, vol. 53, no. 9-10, pp. 1817–1824, 2020.

[19] O. Almutiry et al., "Underwater images contrast enhancement and its challenges: A survey," *Multimedia Tools Appl.*, vol. 83, pp. 1–26, 2024.

[20] J. Čejka et al., "Detecting square markers in underwater environments," *Remote Sens.*, vol. 11, no. 4, pp. 459–482, 2019.

[21] A. Jalal et al., "Fish detection and species classification in underwater environments using deep learning with temporal information," *Ecological Inform.*, vol. 57, 2020, Art. no. 101088.

[22] W. Zhang et al., "Color correction and adaptive contrast enhancement for underwater image enhancement," *Comput. Elect. Eng.*, vol. 91, 2021, Art. no. 106981.

[23] C. Li, S. Anwar, J. Hou, R. Cong, C. Guo, and W. Ren, "Underwater image enhancement via medium transmission-guided multi-color space embedding," *IEEE Trans. Image Process.*, vol. 30, no. 1, pp. 4985–5000, May 2021.

[24] S. Hao, X. Han, Y. Guo, X. Xu, and M. Wang, "Low-light image enhancement with semi-decoupled decomposition," *IEEE Trans. Multimedia*, vol. 22, no. 12, pp. 3025–3038, Dec. 2020.

[25] T. Li et al., "Underwater image enhancement using adaptive color restoration and dehazing," *Opt. Express*, vol. 30, no. 4, pp. 6216–6235, 2022.

[26] B. Zhang et al., "Progress and challenges in intelligent remote sensing satellite systems," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, no. 1, pp. 1814–1822, Feb. 2022.

[27] R. Raskar, A. Ilie, and J. Yu, "Image fusion for context enhancement and video surrealism," in *Proc. 3rd Int. Symp. Non-Photorealistic Animation Rendering*, 2004, pp. 85–152.

[28] S. An et al., "HFM: A hybrid fusion method for underwater image enhancement," *Eng. Appl. Artif. Intell.*, vol. 127, 2024, Art. no. 107219.

[29] G. Choudhary and D. Sethi, "Mathematical modeling and simulation of multi-focus image fusion techniques using the effect of image enhancement criteria: A systematic review and performance evaluation," *Artif. Intell. Rev.*, vol. 56, no. 11, pp. 13787–13839, 2023.

[30] H. Ghanbari, M. Mahdianpari, S. Homayouni, and F. Mohammadimanesh, "A meta-analysis of convolutional neural networks for remote sensing applications," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 3602–3613, Mar. 2021.

[31] C. H. Chong, A. S. A. Ghani and K. Z. M. Azmi, "Dual image fusion technique for underwater image contrast enhancement," in *Proc. 11th Nat. Tech. Seminar Unmanned Syst. Technol.*, 2021, pp. 57–72.

[32] A. A. Siddique, Z. Mohy-Ud-Din, and M. Tahir Qadri, "Real time image encoding for fast IOT (internet of things) based video vigilance system," *Wireless Pers. Commun.*, vol. 114, no. 2, pp. 995–1008, 2020.

[33] J. Lei, X. Luo, L. Fang, M. Wang, and Y. Gu, "Region-enhanced convolutional neural network for object detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5693–5702, Aug. 2020.

[34] H. Wang et al., "Underwater image super-resolution using multi-stage information distillation networks," *J. Vis. Commun. Image Representation*, vol. 77, 2021, Art. no. 103136.

[35] S. Krishnamoorthy and K. P. Soman, "Implementation and comparative study of image fusion algorithms," *Int. J. Comput. Appl.*, vol. 9, no. 2, pp. 25–35, 2010.

[36] S. Li et al., "Pixel-level image fusion: A survey of the state of the art," *Inf. Fusion*, vol. 33, pp. 100–112, 2017.

[37] P. Hambarde, S. Murala, and A. Dhall, "UW-GAN: Single-image depth estimation and image enhancement for underwater images," *IEEE Trans. Instrum. Meas.*, vol. 70, no. 1, Oct. 2021, Art. no. 5018412.

[38] A. M. Khan, M.-Á. Luque-Nieto, and A. A. Siddique, "Underwater efficient data routing: Clustering-travel salesman protocol (CTSP)," *IEEE Access*, vol. 12, pp. 26428–26440, 2024.

[39] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 1, pp. 3735–3756, Jun. 2020.

[40] A. Dhillon and G. K. Verma, "Convolutional neural network: A review of models, methodologies and applications to object detection," *Prog. Artif. Intell.*, vol. 9, no. 2, pp. 85–112, 2020.

[41] A. A. Siddique, M. T. Qadri, N. A. Siddiqui, and Z. Mohy-ud-Din, "Temporal masking with Luma adjusted interframe coding for underwater exploration using acoustic channel," *Wireless Pers. Commun.*, vol. 116, no. 2, pp. 1493–1506, 2021.

[42] X. Du et al., "A comparative study of different CNN models and transfer learning effect for underwater object classification in side-scan sonar images," *Remote Sens.*, vol. 15, no. 3, pp. 593–606, 2023.

[43] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, and K. E. Muller, "Contrast-limited adaptive histogram equalization: Speed and effectiveness," in *Proc. 1st Conf. Visualization Biomed. Comput.*, 1990, pp. 337–345.

[44] Y.-C. Liu, W.-H. Chan, and Y.-Q. Chen, "Automatic white balance for digital still camera," *IEEE Trans. Consum. Electron.*, vol. 41, no. 3, pp. 460–466, Aug. 1995.

[45] T. Chen, Z. Lu, Y. Yang, Y. Zhang, B. Du, and A. Plaza, "A Siamese network based U-Net for change detection in high resolution remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, no. 1, pp. 2357–2369, Mar. 2022.

[46] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 81–88.

[47] C. O. Ancuti, C. Ancuti, C. D. Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE Trans. Image Process.*, vol. 27, no. 1, pp. 379–393, Jan. 2018.

[48] S.-B. Gao, M. Zhang, Q. Zhao, X.-S. Zhang, and Y.-J. Li, "Underwater image enhancement using adaptive retinal mechanisms," *IEEE Trans. Image Process.*, vol. 28, no. 11, pp. 5580–5595, Nov. 2019.

[49] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and M. Sbert, "Color channel compensation (3C): A fundamental pre-processing step for image enhancement," *IEEE Trans. Image Process.*, vol. 29, no. 1, pp. 2653–2665, Nov. 2020.

[50] T. Khan, R. Iftikhar, A. Akbar Siddique, and S. Asghar, "An algorithm to reduce compression ratio in multimedia applications," *Comput., Mater., Continua*, vol. 75, no. 1, pp. 539–557, 2023.

[51] J. Y. Chiang and Y.-C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE Trans. Image Process.*, vol. 21, no. 4, pp. 1756–1769, Apr. 2012.

[52] J. Wang, Y. Zheng, M. Wang, Q. Shen, and J. Huang, "Object-scale adaptive convolutional neural networks for high-spatial resolution remote sensing image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 283–299, Dec. 2021.

[53] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, "WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images," *IEEE Robot. Automat. Lett.*, vol. 3, no. 1, pp. 387–394, Jan. 2018.

[54] B. Huang, G. Chen, H. Zhang, G. Hou, and M. Radenkovic, "Instant deep sea debris detection for maneuverable underwater machines to build sustainable ocean using deep neural network," *Sci. Total Environ.*, vol. 878, 2023, Art. no. 162826.

[55] J. Zhang, C. Xie, X. Xu, Z. Shi, and B. Pan, "A contextual bidirectional enhancement method for remote sensing image object detection," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, no. 1, pp. 4518–4531, Aug. 2020.

[56] A. Sánchez-Ferrer, J. J. Valero-Mas, A. J. Gallego, and J. Calvo-Zaragoza, "An experimental study on marine debris location and recognition using object detection," *Pattern Recognit. Lett.*, vol. 168, pp. 154–161, 2023.

[57] Z. Zhang, L. Zhang, Y. Wang, F. Pengming, and H. Ran, "ShipRSImageNet: A large-scale fine-grained dataset for ship detection in high-resolution optical remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 8458–8472, Aug. 2021.

[58] F. Zocco, T.-C. Lin, C.-I. Huang, H.-C. Wang, M. Omar Khyam, and M. Van, "Towards more efficient efficientDets and real-time marine debris detection," *IEEE Robot. Automat. Lett.*, vol. 8, no. 4, pp. 2134–2141, Apr. 2023.

[59] J. S. Walia and K. Seemakurthy, "Optimized custom dataset for efficient detection of underwater trash," 2023, *arXiv:2305.16460*.

[60] Z. Lu et al., "An image enhancement method for side-scan sonar images based on multi-stage repairing image fusion," *Electronics*, vol. 12, no. 17, 2023, Art. no. 3553.

[61] S. Huang, Y. Shi, M. Li, J. Qian, and K. Xu, "Underwater 3D reconstruction using a photometric stereo with illuminance estimation," *Appl. Opt.*, vol. 62, no. 3, pp. 612–619, 2023.

[62] H. Huang, J. Ma, K. Zheng, S. Zhang, and P. Wu, "A novel underwater image enhancement method based on the dual-image fusion," in *Proc. Int. Conf. Sens., Meas., Data Analytics Artif. Intell.*, 2022, pp. 1–6.

[63] K. Liu and Y. Liang, "Enhancement method for non-uniform scattering images of turbid underwater based on neural network," *Image Vis. Comput.*, vol. 138, 2023, Art. no. 104813.

[64] M. Jian et al., "Underwater image processing and analysis: A review," *Signal Process.: Image Commun.*, vol. 91, 2021, Art. no. 116088.

[65] P. D. Hiscocks and P. Eng, "Measuring luminance with a digital camera," in *Proc. Syscomp Electron. Des. Limited*, 2011, vol. 2.

[66] K. He, Z. Zhang, Y. Dong, D. Cai, Y. Lu, and W. Han, "Improving geological remote sensing interpretation via a contextually enhanced multiscale feature fusion network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, no. 1, pp. 6158–6173, Mar. 2024.

[67] A. W. Setiawan, T. R. Mengko, O. S. Santoso, and A. B. Suksmono, "Color retinal image enhancement using CLAHE," in *Proc. Int. Conf. ICT Smart Soc.*, 2013, pp. 1–3.

[68] R. S. M. Alex, S. Deepa, and M. H. Supriya, "Underwater image enhancement using CLAHE in a reconfigurable platform," in *Proc. MTS/IEEE Monterey*, 2016, pp. 1–5.

[69] D. Garg, N. K. Garg, and M. Kumar, "Underwater image enhancement using blending of CLAHE and percentile methodologies," *Multimedia Tools Appl.*, vol. 77, pp. 26545–26561, 2018.

[70] R. G. Tambe, S. N. Talbar, and S. S. Chavan, "Deep multi-feature learning architecture for water body segmentation from satellite images," *J. Vis. Commun. Image Representation*, vol. 77, 2021, Art. no. 103141.
[71] O. O. Ajayi, A. B. Bagula, H. C. Maluleke, Z. Gaffoor, N. Jovanovic, and K. C. Pietersen, "WaterNet: A network for monitoring and assessing water quality for drinking and irrigation purposes," *IEEE Access*, vol. 10, pp. 48318–48337, 2022.
[72] M. A. Syariz et al., "WaterNet: A convolutional neural network for chlorophyll-A concentration retrieval," *Remote Sens.*, vol. 12, no. 12, 2020, Art. no. 1966.
[73] J. Shao, Q. Yang, C. Luo, R. Li, Y. Zhou, and F. Zhang, "Vessel detection from nighttime remote sensing imagery based on deep learning," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, no. 1, pp. 12536–12544, Nov. 2021.
[74] A. Jaiswal et al., "Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021.
[75] F. Salim et al., "DenseNet-201 and Xception pre-trained deep learning models for fruit recognition," *Electronics*, vol. 12, no. 14, 2023, Art. no. 3132.
[76] H. Wu, M. Zhang, P. Huang, and W. Tang, "CMLFormer: CNN and multi-scale local-context transformer network for remote sensing images semantic segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 17, no. 1, pp. 7233–7241, Mar. 2024.
[77] P. Chlap et al., "A review of medical image data augmentation techniques for deep learning applications," *J. Med. Imag. Radiat. Oncol.*, vol. 65, no. 5, pp. 545–563, 2021.
[78] K. Maharana, S. Mondal, and B. Nemade, "A review: Data pre-processing and data augmentation techniques," *Glob. Transitions Proc.*, vol. 3, no. 1, pp. 91–99, 2022.
[79] S. Surajit, "SouvikDataset," Kaggle, 2022. Accessed: May 3, 2024. [Online]. Available: https://www.kaggle.com/datasets/surajit651/souvikdataset
[80] G. Iglesias et al., "Data augmentation techniques in time series domain: A survey and taxonomy," 2022, *arXiv:2206.13508*.

Wad Ghaban received the B.Sc. (with Hons.) degree in computer science from King Abdul-Aziz University, Jeddah, Saudi Arabia, in 2009, the M.Sc. (with distinction) degree in advanced computer science, and the Ph.D. degree in computer science from the University of Birmingham, Birmingham, U.K., in 2015 and 2020, respectively.

She is currently an Assistant Professor with the Applied College, University of Tabuk, Tabuk, Saudi Arabia. During her studies, she worked on several projects related to human-computer interaction, survival analysis, online learning, natural language processing, and sentiment analysis.

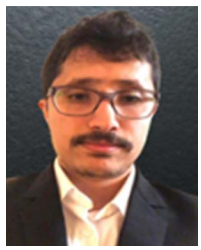
Jawad Ahmad (Senior Member, IEEE) received the Ph.D. degree in computing from Glasgow Caledonian University, Glasgow, Scotland, U.K., in 2019.

is a highly experienced Teacher with more than 13 years of teaching and research experience at prestigious institutions. He has conducted research and taught at renowned institutions, such as Prince Mohammad Bin Fahd University (KSA), Edinburgh Napier University (U.K.), Glasgow Caledonian University (U.K.), Hongik University (South Korea), among others. He has authored or coauthored more than 200 research papers in renowned journals, including IEEE Transactions, ACM Transactions, Elsevier, and Springer, with more than 6000 citations and an H-Index of 45. For the past three years, his name has appeared on the list of the world's top 2% of scientists in Cybersecurity, as published by Clarivate (a list endorsed by Stanford University, USA). In 2020, he received the U.K. endorsement as an exceptional talent candidate (Emerging Leader) for his pioneering work in Cybersecurity and AI. To date, he has secured more than £250K in research and funding grants in the U.K. and Norway as a Principal Investigator and Co-Investigator. In terms of academic achievements, he has earned a Gold Medal for his outstanding performance in M.S. degree and a Bronze Medal for his achievements in B.S. degree.

Ali Akbar Siddique received the bachelor's degree in electronics engineering from the Sir Syed University of Engineering and Technology, Karachi, Pakistan, in 2009, the master's degree in control and automation from the Usman Institute of Technology, Karachi, in 2013, and the Ph.D. degree in electronic engineering from Sir Syed University of Engineering and Technology, in 2021.

During his master's, he achieved a 3.81 CGPA. His research focuses on signal, image, and video processing.

Dr. Siddique was the recipient of Gold Medal in his respective discipline.



Mohammed S. Alshehri received the B.S. degree in computer science from the King Khalid University, Abha, Saudi Arabia, in 2010, the M.S. degree in computer science from the University of Colorado Denver, Denver, CO, USA, in 2014, and the Ph.D. degree in computer science with a concentration on information security from the University of Arkansas, Fayetteville, AR, USA, in 2021.

His research interests include cybersecurity, computer networks, blockchain, machine learning, and deep learning.

Anila Saghir received the bachelor's and master's degrees in telecommunication engineering from NED University of Engineering and Technology, Karachi, Pakistan, in 2012 and 2016, respectively. She is currently working toward the Ph.D. degree in computer engineering with the Sir Syed University of Engineering and Technology, Karachi.

Her research interests include federated learning and deep learning in IoT and 6G-enabled edge computing applications.

Faisal Saeed received the B.Sc. degree in computers (information technology) from Cairo University, Giza, Egypt, in 2005, and the M.Sc. degree in information technology management and the Ph.D. degree in computer science from Universiti Teknologi Malaysia (UTM), Malaysia, in 2010 and 2013, respectively.

He is currently an Associate Professor with the College of Computing and Digital Technology, Birmingham City University, Birmingham, U.K. He is leading the M.Sc. Artificial Intelligence course, Research Summer School, and Smart Health Lab at the College. From 2017 to 2021, he worked as an Assistant and Associate Professor with the Information Systems Department, Taibah University, KSA. From 2014 to 2017, he was a Senior Lecturer with the Information Systems Department, Faculty of Computing, UTM. His research interests include data mining, artificial intelligence, machine learning, information retrieval, and health informatics.

Dr. Saeed achieved an outstanding publication record in prestigious high-impact factor journals and presented papers at several international conferences. Moreover, he has secured 21 research funds from various funders in Malaysia, Saudi Arabia, and the U.K. He served as the General Chair of 12 international conferences and as Guest Editor of tens of indexed journals.

Baraq Ghaleb received the Ph.D. degree in computing from Edinburgh Napier University, Edinburgh, U.K., in 2019.

He is currently an Associate Professor with the Centre for Distributed Computing, Networks and Security, Edinburgh Napier University. He is currently involved in delivering and leading several modules for the undergraduate and postgraduate levels with the Cyber Security and Systems Engineering subject group. He is an experienced researcher with more than 8 years of cutting-edge research around cybersecurity, the Internet of Things, blockchain, and machine learning. He has authored or coauthored more than 35 research papers in prestigious peer-reviewed outlets. His research interests revolve around investigating the security vulnerabilities of the Internet of Things standards and utilizing cutting-edge advancements in addressing such vulnerabilities.

Dr. Ghaleb has been involved in several externally funded research projects as a Principal Investigator and Co-Investigator, with a total budget of around £435,000. In addition, he has been involved in supervising and mentoring master's and Ph.D. research students since completing his Ph.D. thesis in 2019.

Mujeeb Ur Rehman received the Ph.D. degree in engineering, with a focus on artificial intelligence and cybersecurity, in 2022.

He is a highly experienced Educator with more than ten years of teaching and research experience at prestigious institutions, including De Montfort University, Leicester, U.K., University of Glasgow, Glasgow, U.K., York St John University, York, U.K., and Riphah International University, Islamabad, Pakistan, among others. He was a Supervisor for numerous postgraduate and undergraduate students, providing valuable guidance and support for their dissertations. He has authored or coauthored more than 40 research papers in renowned journals, including IEEE transactions, IET transactions, Elsevier, and Springer.

Dr. Rehman was the recipient of the U.K. endorsement as an Exceptional Talent (Emerging Leader) candidate for his pioneering work in Artificial Intelligence and Cybersecurity in 2022 and a Gold Medal for his outstanding performance in his M.S. degree and earned a distinction in his B.S. degree. He has secured multiple research grants as both a Principal Investigator and Co-Investigator.