DOI: 10.21917/ijsc.2023.0415

BUILD AN EFFECTIVE DEEP LEARNING MODEL FOR UNDERWATER IMAGE ENHANCEMENT BASED ON EUVP DATA

Sugandha Jaitly and Sudha Bhati
Department of Computer Science, NGF College of Engineering and Technology, India

Abstract

Images taken underwater often suffer from colour distortion as well as a reduction in visibility because of the light absorption and scattering that occurs. Existing approaches for underwater image enhancement make use of numerous assumptions/constraints to produce decent results. However, these solutions all have the same disadvantage in that the assumptions used may not work for certain scenarios. To solve this issue, this research provides an end-to-end system for underwater image augmentation that includes a CNN-based network dubbed VGG and UNet. There have been several studies in recent years proving the efficiency of deep learning approaches in various application fields. Color correction and haze removal tasks are used to train the VGG and UNet. This combined training technique allows for the simultaneous learning of a robust feature representation for both tasks. A pixel disrupting method is used in the suggested learning framework to better extract the intrinsic characteristics in local patches, which considerably enhances convergence speed and accuracy. We used EUVP dataset training images based on the underwater imaging model to handle VGG and UNet training. The testing findings on several real-world underwater settings show that the suggested strategy yields aesthetically pleasing outcomes. The experimental results for the full-reference measures SSIM, PSNR, RMSE, UCIQE, and UIQM demonstrate the reliability and efficacy of the proposed technique.

Keywords:

Image Enhancement, Underwater Image Enhancement, Deep Learning, VGGNET Model

1. INTRODUCTION

Over the course of the last several years, underwater image enhancement (UIE) has garnered a significant amount of interest in the fields of image processing (IP) as well as underwater vision [1][2]. Image enhancement underwater is a tough subject to solve because of the challenging habitat and lighting circumstances that exist underwater. In most cases, the quality of an image captured underwater would suffer as a result of wavelength-dependent absorption and scattering, which includes both forward and backward scattering [3]. In addition to this, the marine snow contributes to the overall level of noise and amplifies the scattering effects. These unfavorable effects lower visibility, lessen contrast, and even add colour casts, all of which restrict the practical uses of underwater images and videos in fields such as marine biology and archaeology, marine ecology, to mention a few. Newer In order to solve this problem, algorithms only use information from a data image, in contrast to older approaches that relied on numerous underwater images or polarisation filters to overcome the problem.

Underwater images are unavailable to the general public due to the lack of a publicly available dataset, a thorough study of UIE algorithms as well as an insightful analysis of these algorithms have proven to be largely unsatisfactory, despite the large amount of work that has been put into them [4][5]. In addition, it is very hard to take a snapshot of a true underwater scene while also capturing the ground truth image that corresponds to that scenario for each of the many kinds of water. The performance of learned in the classroom UIE algorithms is not comparable to the accomplishment of recent deep having to learn low - and high vision issues. A lack of appropriate and efficient training data is to blame [6] [7]. To further the progress of the development of UIE, we create an efficient approach for the enhancement of underwater images [8]. The Fig.1 displays a few sample images together with the reference images that correspond to them from the EUVP (Enhancing Underwater Visual Perception) dataset. These were used in the process of constructing this model. It has been shown that the raw underwater images included inside the EUVP contain a wide variety of Color ranges and contrast are reduced. While this may be true for the originals, the corresponding images have been enhanced to make them more visible and brighter (or at least, they should be). An investigation of numerous state-of-the-art algorithms for boosting underwater images is carried out by us using our approach. This study gives us a better understanding of how they perform and sheds light on the future research opportunities. Moreover, the created technique for enhancing underwater images includes a mechanism that makes it simple to train VGGNET to increase the overall visual quality of an underwater image. In order to show the utility of this application, we have developed a model for improving underwater images called Water-Net. This model was trained using the approach that was established for improving underwater images.

Fig.1. Images taken from the EUVP. The images in the top row are the unprocessed versions, while the images in the bottom row are the corresponding reference images.

Following is a summary of the main contributions of this study.

- Using the EUVP dataset, which provides different sets of paired and unpaired image samples of low and acceptable perceptual quality to assist supervised training of underwater image improvement models, we create an UIE approach.
- In order to train underwater tips to improve models, we use the EUVP dataset, which includes paired and unpaired image samples of low as well as acceptable perceptual quality.
- Using the developed underwater image enhancement approach, we undertake a complete analysis of the most

advanced methods for improving a single underwater image, from qualitative to quantitative. Our thorough examination and analysis reveal the strengths and limitations of current UIE methods as well as recommend new study avenues.

- Our proposed VGGNET-based UNET Machine Learning technique (Water-Net) is trained using the underwater contrast enhancement (UIE) method for UIE. This also demonstrates the generalizability of our Water-Net and its advantages, but it also serves as inspiration for the creation of a UIE based on deep learning.

2. LITERATURE REVIEW

The section that follows is a literature review on UIE. This section provides information about prior research relating to the present subject. According to the conclusions of the supplied literature analysis, it is comparable to establishing a precedent among the ways that are currently available. Utilizing deep learning (DL) methods, a variety of studies are being undertaken.

In this study [9], the correlation between deep learning-based image enhancement and underwater object recognition is investigated. First, three image enhancement techniques UWCNN as well as FUNIE-GAN are being used to dehaze images from the same URPC dataset. The URPC set of data and the three improved datasets are then trained and evaluated using the SSD method. Image quality parameters and object detection rates are also analyzed in this research. The image-enhanced dataset’s ordinary object identification accuracy has improved, but no statistically significant links exist between changes to the image quality parameter and the final detection accuracy, according to the experiments. Multiple variables may have contributed to the marginal improvement in the final object detection accuracy.

In this paper, [10] develop Only 890 out of 950 real-world immersed images have the previous relevant images inside the Underwater collection Contrast Enhancement Benchmark For the purpose of training Convolutional Neural Networks, they provide an image enhancement system (named Water-Net) that was trained using this benchmark as a starting point (CNNs). To better understand cutting-edge algorithms, the benchmark evaluations and the proposed Water-Net provide information for future research in UIE.

The paper [11] proposes using a remotely operated vehicle (ROV) to identify trash in the ocean floor. To train an object detection model based on YOLO Neural Network topologies, they compiled their own dataset of three types of marine debris. This paper made use of YOLOv4. Many noise reduction and image enhancement filters are applied to our dataset’s training images in order to improve the accuracy of our results. In their experiment, the ROV was able to collect a image and transmit it to a computer for analysis, augmentation, and identification.

This article [12] proposes a fully connected convolutional neural network for dehazing underwater images. Deep encoder-decoder architecture integrates low-level and high-level elements, enabling the recovery of a fuzzy image. The suggested approach is statistically and subjectively assessed using UIEB. In terms of SSIM, PSNR, and MSE, the model surpasses previous approaches and has the potential to dehaze underwater images while preserving small features.

In this article, [13] propose 2 perceptual augmentation models, which each uses a detection perceptor and a deep enhancement model. Patch levels that are visually appealing or detectable are created using feedback provided by the sensing perceptor in the form of gradients. It has also been proposed that because there is not enough training data available, it is possible to create underwater images can be enhanced using a combination of physical priors as well as data from real-world images to train the model. This model can be found here. The findings of their experiments demonstrate that their suggested technique is superior to a number of state-of-the-art techniques when applied to synthetic and real-world datasets including underwater environments.

In [14], Offer a method for the production of synthetic underwater images for huge databases. Real-world as well as synthetic underwater datasets are used to assess the performance of a proposed network. Additional testing is done on underwater images with different colour dominance, comparison, and illumination conditions to see how the proposed UW-GAN performs. The presented UW-GAN structure is also used to improve underwater single images. Extensive analysis of the results shows that the proposed UW-GAN is superior to the state-of-the-art (SoTA) hand-crafted and learning-based techniques for underwater single scale estimation (USIDE).

This work [15] Propose an encoder-decoder Siamese architecture-based A network of underwater image enhancers (UICoE-Net). So that the reciprocal relationship of the two branches could be transmitted, they included. The similarity feature based units use a Siamese encoder-decoder structure for joint learning. Numerous experiments utilising the Underwater Contrast Benchmark (UIEB), a Submerged Image Co-enhancement Set of data (UICoD), and the Sound system Quantitative Covered in water Image Dataset have demonstrated the efficacy of this method so far (SQUID).

3. RESEARCH METHODOLOGY

Following a thorough review of prior work, the main issue in underwater image enhancement is a lack of relatively large depth image data sets reference images. In the following sections, we describe the dataset in detail, including proposed methodology, all phases of the proposed system, proposed algorithm, and a complete flowchart.

3.1 PROPOSED METHODOLOGY

It is the primary goal of the proposed systems to improve underwater object quality by using filters, neural network image classification and underwater image detection. Deep learning approaches have arisen as a supplement to classic model-based techniques and therefore have contributed to the definition of a new state of the art in terms of the quality of dehazed images that may be achieved.

In this research, we proposed a deep learning model. For this research starts with collecting dataset. Firstly, collect the EUVP (Enhancing Underwater Visual Perception). Next apply preprocessing technique that covert the data into various forms like navigating the directory containing the input image The images are read one by one, the BGR format is converted to RGB format, the images are normalized, and so on. Finally, EDA was used to plot histogram layouts of the raw as well as predicted images to better understand the difference. Then implement the VGGNet based UNet neural network model to enhance underwater images. In this model, split the dataset into training and testing, used activation function i.e., ReLU, and sigmoid, adam optimizer, number of epoch five and batchsize 1. Finally, calculate the performance evaluation matrix that is RMSE, SSIM, PSNR, UCIQE, and UIQM, the experimental results confirm the efficiency and robustness of the proposed method. All process described below briefly.

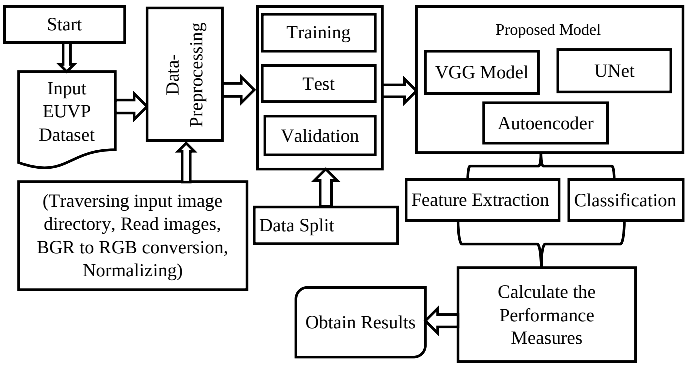


Fig.2. Block architecture of the proposed method

In above Fig.2 shows the proposed flowchart of research work. We can see in graph, the process starts with data collection, then preprocess, data splitting training and testing then implement proposed VGGNet model and lastly calculate the performance evaluation of proposed model.

3.1.1 Image Preprocessing:

Data pre-processing is a prominent and effective strategy in the process of deep learning. In part, this is owing to its ability to both expand the size of the original database and to enhance the information contained inside. Therefore, pre-processing has a considerable influence on the efficiency with which following operations may be completed. The basic purpose of image processing is to improve the visual data by eliminating noise and improving image pixels that have been corrupted. A variety of strategies are used to reach this goal. Here, in this project, we had to transform them into different shapes to create the neural network that would carry out all of the job. Both the raw and ground truth images, also known as CLAHE and GC, are included in the dataset. Even though the set of data we used has already been processed, there is still work to be done on it. After this, apply the data augmentation in this preprocessing phase. Data augmentation is a process of replicating the image in different forms for example, horizontal flips, vertical flips, scaling, and zooming. Then, apply the data pre-processing technique. In this technique, first, converting the images from BGR to RGB and normalizing the images one by one as they are read from the input image directory.

3.1.2 Data Splitting:

We carried out the evaluation of the model by first dividing the data set into two distinct parts: the training set as well as the testing set. The model is put through its paces with the help of the training set, while its performance is evaluated with the help of the testing set. Data split into 80% for training and 20% for testing.

3.1.3 Proposed Model (VGGNet based UNet Neural Network):

A neural network [16] [17] approaches that are used to find correlations in data using a procedure that is similar to how the human brain functions. In this sense, the phrase neural network systems are made up of neurons, both biological and artificial. Even if the output criteria are not altered, neural networks will deliver the greatest possible conclusion since they are able to adapt to changes in input. Trading systems are increasingly using neural networks, a concept that has its roots in artificial intelligence (AI).

In 2014, The University of Oxford's Karen Simonyan as well as Andrew Zisserman developed VGGNet, a Dcnn. The Visual Geometry Group is the full name of the organisation. From VGG16 through VGG19, it has produced a succession of convolutional network models for face recognition and image categorization. Convolutional network depth was first investigated by VGG to better understand how network depth influences accuracy and precision while classifying and recognising large-scale images. - CNN (Deep-16 p.m. ET) A 3x3 convolution kernel is utilised in all levels that more network layers can be added without introducing too many parameters.

3.1.4 VGGNet Architecture:

As a foundation, object recognition models have been developed using the VGG architecture. The VGGNet, designed The ImageNet baseline is outperformed on a wide variety of functions and datasets by a deep neural network. In addition, it is a widely used technique for recognizing images nowadays.

VGG accepts an RGB image with a resolution of 224x244 as input. All images in the training set are averaged out and used as a source for constructing VGG convolutional neural networks. This method uses a 3x3 or even a 1x1 filter for convolution. There are between 11 and 19 VGG layers in total, which are made up of convolutional and full-connected layers. Eight convolutional layers as well as 3 fully connected layers are required to produce a VGG11. There are up to 16 convolutional layers in the maximum VGG19. Three more layers of complexity. The VGG network has a convolutional layer for each of its neurons has no pooling layer beneath it, resulting in a total of five pooling layers in the VGG network. VGG's structure is depicted in this diagram:

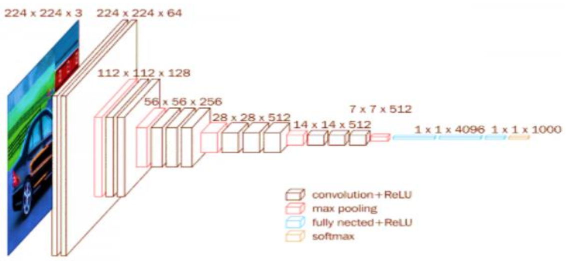


Fig.2. A visualization of VGG architecture [25]

Weight layers in VGGNet consist of A total of 16 convolution layers, three fully connected layers, and five pooling layers. For each of the 2 VGGNet variations, there are 2 4096-channel layers that are fully connected, followed by such a 1000-channel layer that is fully connected and predicts 1000 labels. Lastly, the softmax layer connects all of the other layers together, is being used to classify data.

3.1.5 Architecture:

- There are 64 filters in the first two convolutional layers, which resulting in 224×224×64 volume since the identical convolutions are employed in the first two layers. In every case, the filters have a stride of one and are always 3×3.
- As a result, the 224×224×64 volume's height and breadth were lowered to 112×112×64 using the pooling layer's max-pool 2×2 size and stride 2.
- Another 2 layers of convolution are used, each with 128 filtering options. The new dimension is 112×112×128 because of this.
- Two convolution layers, each one with 256 filter elements, are added to reduce the volume to 56×56×128 before the pooling layer and down sampling layer are added.
- Stacks with three convolutions are separated by a max-pool layer.
- It then flattens into a Connected (FC) layer to 4096 channels and 1000 softmax outputs.

3.2 UNET NEURAL NETWORK

The U-net algorithm was first developed and use for the purpose of segmenting biological images. Encoder networks are followed by decoder networks in their construction, which may be considered of in a more general sense. Semantic segmentation, unlike classification, needs discrimination at the pixel level as well to map the encoder's discriminative features onto the pixel space, as learned in subsequent stages.

- The encoder is the first half of the architectural diagram (Fig.4). When using a pre-trained classification system such as VGG or ResNet, convolution blocks are used to encode the input image and then a maxpool downsampling is used to produce feature representations at multiple distinct levels.
- The decoder is indeed the final piece of the puzzle. Semantically projecting the encoder's discriminative features (that have a lower resolution) onto the pixel space is the goal for a dense classification (which has a greater resolution). In the decoder, the first steps are upsampling and concatenation, followed by the standard convolution processes.

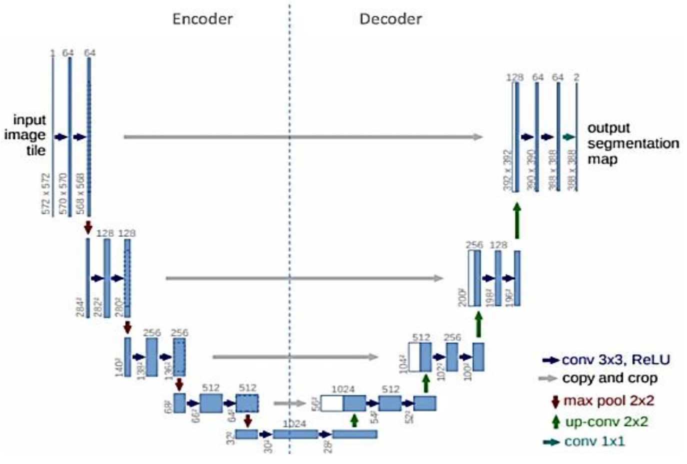


Fig.3. U-net architecture

Feature maps with blue boxes denote multi-channel maps, while those with boxes denote copied maps. Different coloured arrows denote different actions.

Upsampling in CNN [18] may be unfamiliar to those of you are so used to the architecture of classification and object recognition, but the concept behind it is very straightforward. Because our first instinct tells us that we need to try to get the compressed feature map back to the proportions it had when it was applied to the input image, we extend the feature dimensions. There are a few other names for upsampling, including transposed convolution, upconvolution, and deconvolution. From the most straightforward to the most involved method of upsampling is the nearest neighbour method, followed by bilinear interpolation, and then finally transposed convolution.

3.2.1 Auto Encoder:

Autoencoders are a kind of unsupervised learning that makes use of neural networks to perform the job of representation learning. To be more specific, we will create an architectural for a neural network in a way that we will induce a bottleneck in the network, which will need a condensed knowledge bases of the initial input. This compression as well as subsequent reconstruction would be an extremely tough job to do if each of the input characteristics could be considered independent of the others. Nevertheless, if there is some kind of organization in the data, such correlations between the input characteristics, then this structure may be learnt, and as a result, it can be used when driving the input through the bottleneck in the network.

The goal of this supervised learning problem is to reconstruct the original input x from an unlabeled dataset (called an unlabeled dataset in this case). For example, we can train this network by reducing our initial input discrepancies and the successive reconstruction error, $L(x,x)$. There are two values here: x as an input and x as a result. Our network would have a much easier time learning to merely remember the input values if there wasn't a bottleneck in the information flow; if there wasn't a bottleneck in the information flow, our network could easily learn to do so by simply sending these data along through the network [19].

3.2.2 Model Parameters:

- ***Sigmoid Activation function***

Every node's output value is determined by combining all of its input values with a preset function that is used by all other nodes in the network. o_j (sigmoid) functions are frequently utilized, and are described in the following manner [20]:

$$o_j = \frac{1}{1+e^{-i_j}} \tag{1}$$

where, ij is a sum of input nodes of j.

With its normalization to values between 0 to 1 as well as nonlinear nature, Erb considers that this function has two major benefits: it facilitates network learning as well as avoids overloading and dominance impacts. Domination happens an effect on the predicted target variable that is so significant that it renders all other qualities irrelevant and thus completely dominates the attribute to which it is being applied [20].

Relu: In deep learning models, the Rectified Linear Unit (ReLU) is by far the most frequently used activation function. The function returns 0 for any negative input and that value for any positive input. In comparison to the sigmoid function, ReLu is much quicker to calculate, as well as its derivative is also much faster to calculate. For neural networks, it has a major impact on the training and inference time required. As a result, it may be written as:

$$f(x)=\max(0,x) \tag{2}$$

- ***Loss Function***

The Neural Network is comprised of several crucial parts, one of which is the Loss Function. A prediction mistake made by the neural network is all that constitutes a loss. The process of determining the amount of the loss is referred to as the Loss Function. Absolute Loss (Custom made) is the name of the loss function that I used for this job. The real values and the forecasted values may be used as needed parameters in the creation of a custom loss function by constructing a function that accepts these values as inputs. It is expected that the method will return an array of the losses. After that, the function may be sent along throughout the compilation process.

- ***Optimizer***

An approach for optimizing the gradient descent optimization technique, adaptive moment estimation, was developed. When dealing with huge problems that include a significant amount of data or parameters, the approach is quite effective. It makes less of an impact on memory and is effective.

Momentum: Speed up gradient descent by taking into account a exponentially weighted average, as this approach does. This is made possible with considering the gradients. When averages are used, the algorithm moves more quickly in the direction of reaching the minimum value.

$$\omega_{t+1} = \omega_t - \alpha m_t \tag{3}$$

$$m_t = \beta m_{t-1} + (1-\beta)[\delta L/(\delta\omega_t)] \tag{4}$$

3.3 PROPOSED ALGORITHM

Input: EUVP Dataset

Output: Predicated Results

Step 1: Dataset gathering and information

Step 2: Data preprocessing

Searching for images in the input directory

Reading each image one at a time

Converting from BGR to RGB format for the images;

Adjusting the images to a more consistent standard.

Step 3: Exploratory Data Analysis (EDA)

Plotting histograms of the original and predicted images will help to show the disparity between the two.

Step 4: Neural Network Modal to enhance underwater images.

Training and testing samples for underwater and ground truth images are required

Split into training and testing 80% and 20% of the data

Neural network parameters

Activation Function used (RELU, Sigmoid)

Hyper parameters of Training

VGGNET based UNET Neural Network that output an image

Step 5: Calculate Performance Evaluation Metrics: RMSE, SSIM, PSNR, UCIQE and UIQM.

Step 6: Predicated Outcome

4. RESULTS AND DISCUSSION

This section contains the proposed experimental results of implemented model. All experiments were performed on the EUVP dataset that is collected from the Minnesota Interactive Robotics and Vision Laboratory. The results of the tests conducted, which include the performance of the proposed model as well as their evaluation, are reported in this portion of the report.

4.1 DATASET DESCRIPTION

EUVP (Enhancing Underwater Visible Perception) datasets contain separate sets of paired as well as monozygotic images with poor and good perceptual quality for training image enhancement models under supervision. Please visit http://irvlab.cs.umn.edu for more information on how to access this dataset.

Table.1. Description of the Dataset

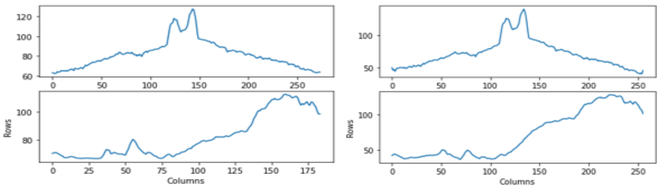
Dataset Name	Training Pairs	Validation	Total Images
Underwater Dark	5550	570	11670
Underwater ImageNet	3700	1270	8670
Underwater Scenes	2185	130	4500

The Table.1 shows the dataset information. In table first column shows the dataset name, second is training pair and third is validation and last is total images. The underwater dark images are total 11670 and its training plot is 5550 and validation is 570. For underwater ImageNet total images is 8670, training pair is 3700 and validation is 1270. And last underwater scenes dataset that contain total images is 400, validation is 130 or training pair is 2185 respectively.

4.2 EXPLORATORY DATA ANALYSIS (EDA)

EDA is a method that may be conceptualized as the practice of Analysing one or more datasets with the intention of gaining an understanding of the fundamental structure of the data that is included within those datasets. This can be seen as the art of studying datasets. EDA is sometimes described as both a frame of mind and a method for doing flexible data analysis without presuming anything about the method by which the data was generated. Enterprises are producing data at an ever-increasing rate, which results in an ever-increasing amount of data as well as an ever-increasing degree of complexity. EDA stands for exploratory data analysis and is a method for doing statistical research.

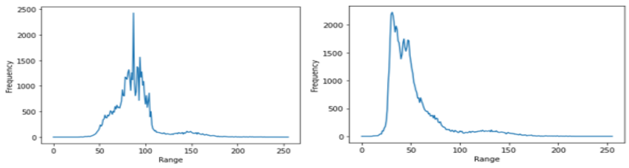
Using the histogram equalization concept, this study suggests a strategy for improving images captured underwater. The vividness and contrast of underwater images are sometimes diminished due to the pervasive presence of a single, dominating color over the whole image. The image is first changed from a colored image to a grayscale image so that further operations may be performed on it. After that, the histogram equalization method is used on the image. Histogram equalization is a method that may be used to alter the intensities of an image in order to improve the contrast. A convolutional neural network model that is trained by the datasets of underwater images is used to preserve the colors in the image so that the model can provide more accurate results.



(a) Raw Image (b) Predicted Image

Fig 6. Average columns and rows of every pixel of raw image and Predicted image

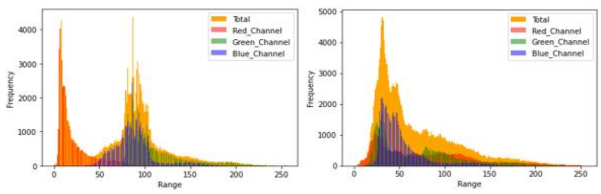
The Fig.6 shows the average columns and rows of every pixel of raw and predicted image. In Fig.6, x-axis shows the columns and y-axis shows the rows. Each image is comprised of a grid of pixels, and that each grid has its own width and height. The number of columns determines the width, while the number of rows determines the height.



(a) Raw Image (b) Predicted Image

Fig.7. Raw and Predicted Image Pixel Frequency in the 0-255 Range

It is shown in Fig.7 how often pixels in the predicted images appear in the raw image. Images Fig.7(a) and Fig.7(b) show the raw image as well as predicted image frequency in the form of graphs. The frequency of image shows on y-axis and range shows the y-axis. The frequency range is 0 to 2500. The numbers that are closer to zero indicate shades that are deeper, while the numbers that are closer to 255 describe shades that are lighter or whiter.

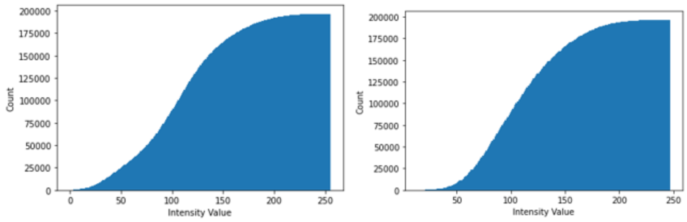


(a) Raw Image (b) Predicted Image

Fig.8. Color channel intensity in both the raw image and the predicted image

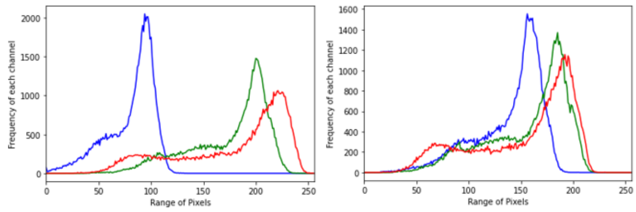
As depicted in Fig.8, raw and predicted images are shown to have different colour intensities. The graph Fig.8(a) shows the intensity color of raw image and graph Fig.8(b) shows the intensity color of dehaze images. In graph x-axis and y-axis shows the range and frequencies of both type of data. The graph shows the RGB color performance.

The histogram shown in Fig.9 is the sum of all raw and premised images. The counts for a single bin and for the total number of bins are shown on a cumulative histogram's vertical scale. totals from smaller values of a given response variable. There are no intervals or bins in a cumulative frequency plot, which is unlike a histogram. Because of this, it is less likely to introduce any sort of bias or error into the analysis than using a wider bin width.



(a) Raw Image (b) Predicted Image

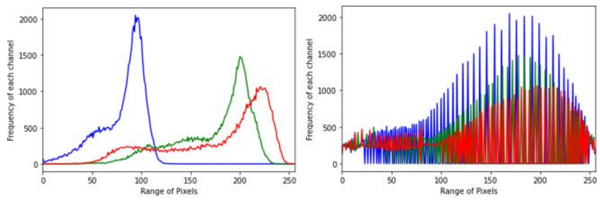
Fig.9. Raw image as well as predicted image cumulative histogram



(a) Raw Image (b) Predicted Image

Fig.10. Color Histogram of every channel of Raw Image and Predicted Image

Raw and predicated images of the channel colour histograms are shown in the images above. Color histograms are used in image processing as well as image to show the distribution of colours in an image. The original image is shown in Fig.10(a). highest frequency shows start between 0 to 90 approx. while in Fig.10(b) the frequency is highly increase between in 100 to 160 range of image pixels.



(a) Normal Histogram (b) Equalized Histogram

Fig.11. Normal Histogram and Equalized histogram of Raw Image

The Fig.11 shows the histogram visualization of input dataset. The Fig.11(a) graph shows the normal histogram, we can see the between 50 to 100 the frequency is very High that shows in blue line. The Fig.11(b) graph shows the equalization histogram, we can see in graph these are various colors that continuously increase and decrease between 0 to 250 range of pixels.

4.3 PARAMETER PERFORMANCE

Utilizing the performance evaluation metrics is necessary to accurately evaluate the performance of the trained DL models.

[21] This helps to determine how much higher a DL model can operate on a dataset that it has never seen before, therefore it's useful in a few situations. In this section, we will present an overview of some of the most helpful performance evaluation tools that may be used in DL [22] [23] [24].

4.3.1 RMSE (Root Mean Square Error):

Another way to measuring error is the root mean square error (RMSE), which is often used to examine the disparities between an estimator's anticipated value and the actual value. The term root mean square error may be used to describe this technique of error measurement. As a result, the error's importance may be calculated. Measures of accuracy are used to compare different estimators for a given variable and their predicted errors. A flawless measure of precision, in other terms.

Suppose we have an estimator for a given estimated parameter; the RMSE is basically the square root of the MSE:

$$RMSE\left(\hat{\theta}\right)=\sqrt{MSE\left(\hat{\theta}\right)} \tag{5}$$

4.3.2 PSNR (Peak Signal to Noise Ratio):

The PSNR is a metric for evaluating the accuracy with which a signal's greatest possible signal power compares to the distorting noise power. The decibel ratio is used to determine the difference between two images when comparing them. Decibel scale logarithm is often employed to calculate the PSNR because of the wide dynamic range that is measured. The dynamic range between the highest and lowest possible values may be altered by their quality. In terms of the PSNR:

$$PSNR=10\log_{10}\frac{\left(peakval^{2}\right)}{MSE} \tag{6}$$

4.3.3 Structure Similarity Index Method (SSIM):

SSIM is a method that depends on people's erroneous assumptions about how similar two things are. When the structural information of an image is changed, it is assumed to degrade the image. Other essential perception-based facts like luminance masking or contrast masking are also involved in this process. Structural information refers to pixels that are interdependent or placed near to one another. Details in the image may be deduced from the densely packed pixels. When the image's distortion is minimized near the image's borders, the technique is known as luminance masking. An image's textural distortions are less noticeable when using contrast masking. SSIM is used to evaluate video and image quality. Both the original and the one that was retrieved are compared.

4.4 COMPARATIVE OUTCOME

Here in this section, the research provides a comparative result of base and propose model with the help of multiple parameters that described below. This parameter calculates in Jupyter notebook as simulation platform.

The Table.2 shows the experimented results of existing and proposed model in the form of RMSE, PSNR, SSIM, UCIQE and QIQM. The proposed model gets RMSE (6.02), SSIM (0.91), PSNR (35.1), UCIQE (0.63) and UIQM (0.67) while base model RMSE is 27.45, SSIM 75%, and UCIQE is 63%, respectively. The experimental findings reveal that the suggested model not only surpasses the older techniques in terms of SSIM and PSNR scores for practically all different kinds of underwater but also generalizes very well to datasets taken from the real world. Using improved images together with our model also demonstrates an increase in the performance of a high-level vision test known as object detection. The proposed approaches are an improvement above those of the existing deep learning models.

Table 2. Comparison between base and purpose model

Results	**RMSE**	**SSIM**	**PSNR**	**UCIQE**	**UIQM**
Base	27.45	0.75	-	0.63	-
Proposed	6.02	0.91	35.5	0.39	0.67

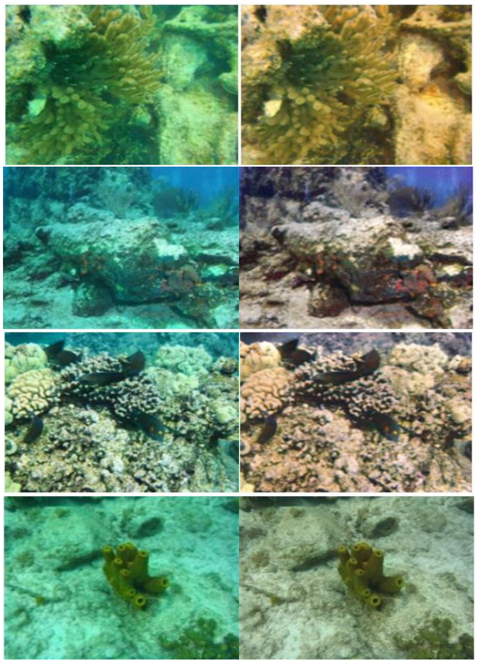


Fig.13. Output Images

The Fig.13 shows the output images of comparison between before implementation and after implementation. The restored archaeology images retain a distinct green tint. There has been a noticeable reduction in greenish after the colour corrections. Although the images have been enhanced, they are over-saturated, making the image look unnatural and losing valuable details.

5. CONCLUSION AND FUTURE WORK

Underwater images have recently emerged as the most efficient means of researching aquatic habitats. In order to counteract refraction (similar to haze) as well as correct colour casts, submerged image restoration is indeed a traditional technique for underwater image processing. Because of the success of deep learning in other computer vision applications, a machine learning approach is proposed here tasks such as colorizing images and identifying objects. To dehaze single images, a convolutional neural network (VGG and UNET) is trained using underwater image improving approaches, surpassing previous image improvement techniques. With a single image as input, the suggested method may generate image restoration quality images. To demonstrate the neural network's generalization capabilities, images from various places and attributes are used to verify it. Finally, the consequences of batch normalization are examined. Our suggested model outperforms previous work in terms of RMSE (6.02), SSIM (0.91), PSNR (35.1), UCIQE (0.63), and UIQM (0.67). According to underwater image improvement studies and a comparison study, the suggested approaches outperform earlier deep learning models and conventional methods in colour correction and detail enhancement.

For underwater activities like pattern recognition, object identification, and more, we will continue to improve our method to achieve a faster runtime and a lower time complexity, both of which will be beneficial.

REFERENCES

[1] J.S. Jaffe, "Underwater Optical Imaging: The Past, the Present, and the Prospects", *IEEE Journal of Oceanic Engineering*, Vol. 67, pp. 1-14, 2015.

[2] M. Sheinin and Y.Y. Schechner, "The Next Best Underwater View", Available at https://openaccess.thecvf.com/content_cvpr_2016/papers/Sheinin_The_Next_Best_CVPR_2016_paper.pdf, Accessed in 2016.

[3] D. Akkaynak, T. Treibitz, T. Shlesinger, R. Tamir, Y. Loya and D. Iluz, "What is the Space of Attenuation Coefficients in Underwater Computer Vision?", *Proceedings of International Conference on Computer Vision and Pattern Recognition*, pp. 21-26, 2017.

[4] C. Li, C. Guo, J. Guo, P. Han, H. Fu and R. Cong, "PDR-Net: Perception-Inspired Single Image Dehazing Network with Refinement", *IEEE Transactions on Multimedia*, Vol. 78, pp. 1-14, 2020.

[5] W. Wang, Q. Lai, H. Fu, J. Shen, H. Ling and R. Yang, "Salient Object Detection in the Deep Learning Era: An In-depth Survey", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, pp. 3239-3259, 2021.

[6] C. Li, R. Cong, J. Hou, S. Zhang, Y. Qian and S. Kwong, "Nested Network with Two-Stream Pyramid for Salient Object Detection in Optical Remote Sensing Images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 99, pp. 1-11, 2019.

[7] R. Cong, "An Iterative Co-Saliency Framework for RGBD Images", *IEEE Transactions on Cybernetics*, Vol. 49, No. 1, pp. 233-246, 2019.

[8] C. Guo, C. Li, J. Guo, R. Cong, H. Fu and P. Han, "Hierarchical Features Driven Residual Learning for Depth Map Super-Resolution", *IEEE Transactions on Image Processing*, Vol. 28, No. 5, pp. 2545-2557, 2019.

[9] J. Zhang, L. Zhu, L. Xu and Q. Xie, "Research on the Correlation between Image Enhancement and Underwater Object Detection", *IEEE Transactions on Image Processing*, Vol. 29, No. 5, pp. 1525-1543, 2020.

[10] C. Li, "An Underwater Image Enhancement Benchmark Dataset and Beyond", *IEEE Transactions on Image Processing*, Vol. 29, No. 3, pp. 1235-1241, 2020.

[11] Y.C. Wu, P.Y. Shih, L.P. Chen, C.C. Wang and H. Samani, "Towards Underwater Sustainability using ROV Equipped with Deep Learning System", *Proceedings of International Conference on Automation Control*, pp. 1-5, 2020.

[12] R. Thomas, L. Thampi, S. Kamal, A.A. Balakrishnan, T.P. Mithun Haridas and M.H. Supriya, "Dehazing Underwater Images using Encoder Decoder Based Generic Model-Agnostic Convolutional Neural Network", *Proceedings of International Conference on Ocean Technology*, pp. 1-6, 2021.

[13] L. Chen, "Perceptual Underwater Image Enhancement with Deep Learning and Physical Priors", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 99, pp. 1-16, 2021.

[14] P. Hambarde, S. Murala and A. Dhall, "UW-GAN: Single Image Depth Estimation and Image Enhancement for Underwater Images", *IEEE Transactions on Instrumentation and Measurement*, Vol. 70, pp. 1-15, 2021.

[15] Q. Qi, "Underwater Image Co-Enhancement with Correlation Feature Matching and Joint Learning", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 32, No. 3, pp. 1133-1147, 2022.

[16] S. Kollmannsberger, D. D'Angella, M. Jokeit and L. Herrmann, "Neural Networks", *Proceedings of International Conference on Studies in Computational Intelligence*, pp. 1-12, 2021.

[17] L Haripriya and M.A. Jabbar, "A Survey on Neural Networks and Its Applications", *International Journal of Engineering Research in Computer Science and Engineering*, Vol. 13, No. 1, pp. 1-8, 2018.

[18] O. Ronneberger, P. Fischer and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation", *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 234-241, 2015.

[19] J. Jordan, "Introduction to Autoencoders", Available at https://www.jeremyjordan.me/autoencoders/, Accessed at 2018.

[20] R. J. Erb, "Introduction to Backpropagation Neural Network Computation", *Pharmaceutical Research: An Official Journal of the American Association of Pharmaceutical Scientists*, Vol. 34, pp. 1-9, 1993.

[21] C. Hodges, M. Bennamoun and H. Rahmani, "Single Image Dehazing using Deep Neural Networks", *Pattern Recognition Letters*, Vol. 79, pp. 1-17, 2019.

[22] B. Sankur, "Statistical Evaluation of Image Quality Measures", *Journal of Electronic Imaging*, Vol. 11, No. 2, pp. 206-223, 2002.

[23] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity", *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600-612, 2004.

[24] A. Mittal, A.K. Moorthy and A.C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain", *IEEE Transactions on Image Processing*, Vol. 21, No. 12, pp. 4695-4708, 2012.

[25] Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", *Proceedings of International Conference on Learning Representation*, pp. 1-14, 2014.