An End-to-End CNN Approach for Enhancing Underwater Images Using Spatial and Frequency Domain Techniques

Ayah Abo El Rejal[*] , Khaled Nagaty , Andreas Pester

Faculty of Informatics and Computer Science, The British University in Egypt, 11837 Cairo, Egypt

[*] Correspondence: Ayah Abo El Rejal (Ayah.hesham@bue.edu.eg)

Abstract: Underwater image processing area has been a central point of interest to many people in many fields such as control of underwater vehicles, archaeology, marine biology research, etc. Underwater exploration is becoming a big part of our life such as underwater marine and creatures research, pipeline and communication logistics, military use, touristic and entertainment use. Underwater images are subject to poor visibility, distortion, poor quality, etc., due to several reasons such as light propagation. The real problem occurs when these images have to be taken at a depth which is more than 500 feet where artificial light needs to be introduced. This work tackles the underwater environment challenges such as as colour casts, lack of image sharpness, low contrast, low visibility, and blurry appearance in deep ocean images by proposing an end-to-end deep underwater image enhancement network (WGH-net) based on convolutional neural network (CNN) algorithm. Quantitative and qualitative metrics results proved that our method achieved competitive results with the previous work methods as it was experimentally tested on different images from several datasets.

Keywords: Underwater; Poor visibility; Distortion; Convolutional; Qualitative; Quantitative

1. Introduction

In the last decade, various techniques and algorithms were proposed to solve the problem of distorted underwater images which is caused due to several reasons such as light attenuation, water type, depth of water level and different wavelengths of light rays. Light attenuation limits the visibility distance, at about five meters or less in turbid water. It is not only caused by absorption that removes light energy, but also by scattering which changes the direction of light path. Backscattering is one of the main challenges that face underwater images. It happens due to flash lighting up the small dust-like particles present between the lens of the camera and the object that has been photographed. It affects the images negatively and they appear as if they are taken in a dust storm.

Several approaches were used to enhance the images taken underwater; however, none of these approaches were suitable. Some approaches did not yield accurate results and others were expensive to implement depending on the use of hardware such as professional cameras. Under water images are taken by different types of compact digital underwater cameras; however, to achieve a high-quality image very expensive cameras such as Sony RX100 V1, Olympus Tough TG-6 and many others should be used which are not a practical solution. Less expensive cameras can be used to take the images that have the enhancing algorithm integrated inside it. This provides an affordable solution; however, a very effective model should be designed to get high satisfying results. Improving the software side will decrease the dependencies on the hardware as a result expenses will be decreased.

Images taken underwater participate in many fields such as oceanic engineering, discovering new creatures, etc. Marine biologists use underwater images for many uses such as fish classification, monitoring the health level of the marine system and identifying new species without the need to remove them. It also has a great economic impact in projects such as inspection for underwater pipelines and cables for gas and oil industry. Moreover, Marine snow is the continuous flow of particles that starts from the top levels of the ocean and accumulates at the seafloor. As a result, it introduces some sort of noise to the images taken under water and increases the negative effects of scattering. Underwater vehicle navigation is used in archaeological and geological needs. Moreover, it

is also useful for international telecommunications traffic.

As illustrated in Figure 1, blue and green light have short wavelengths and as a result they have higher energy and can penetrate much more deeply than red light. The long wavelengths such as red can only penetrate to a very shallow depth that is not more than approximately 50 meters. This is the reason why underwater images appear bluish green unlike normal images taken above water [1]. It is known that different salinity levels of water, affects the amount of different wavelengths absorption.
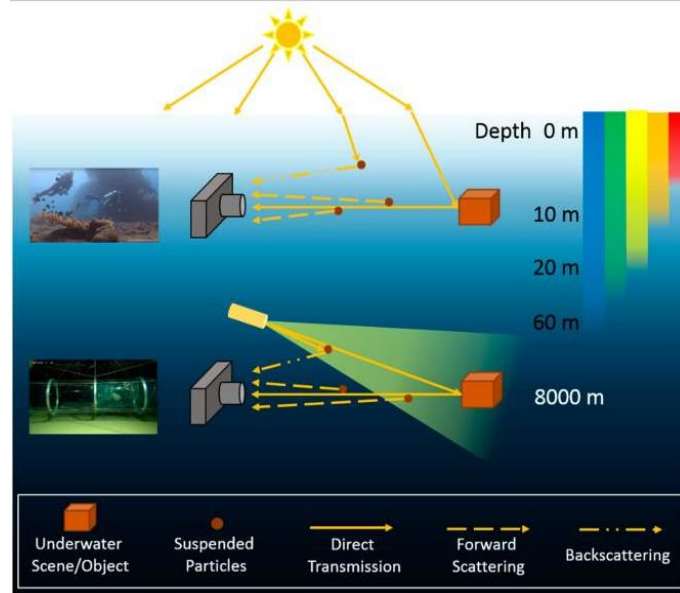

Figure 1. Underwater optical imaging in shallow water and deep sea [1]

Several models with different techniques were introduced in the past decade to improve the appearance and quality of underwater images, trying to restore colors and details so that they are more useful. These methods can be categorized into three main categories, model-free, model-based, and data-driven network models.

2. State of the Art Techniques

Various methods were proposed by other researchers and these methods can be categorized into 3 categories: Model -free methods, Model-based methods (also known as Model-agnostic Methods) and Data-Driven methods.

2.1 Model-Free Methods

At the beginning, researchers used model-free methods in enhancing underwater images. Model free methods are the simple enhancement techniques used in image processing. They can be further categorized into two subcategories that use either the spatial domain or the frequency domain. Spatial domain methods use simple algorithms to modify pixel values of an image without modelling the process of image formation. Some examples of this type of methods are histogram equalization, along with its different versions such as contrast limited adaptive histogram equalization (CLAHE) [2], white balancing and its versions for example automatic white balancing [3]. Moreover, in 2007, new color constancy method was proposed namely, Grey-Edge hypothesis based on using grey edge algorithms that assume that the average edge difference is achromatic. Color constancy is mainly the ability to measure object's colors independent of the value of light source present in the image [4]. These methods improved the visual quality to some extent, but accentuated noise, introduced artifacts and caused some color distortions.

Furthermore, transform (frequency) domain methods map image pixels into a specific domain where the physical properties are exploited to perform adjustments. Most used transformers include Fourier and wavelets [5]. It improved the quality of the images by amplifying the high frequency component and suppressing the low frequency ones. Underwater images suffer from the problem of having a small difference between the high frequency component of the edge and the low frequency component of the background. As a result, approaches such as holomorphic filter [6] were used.

Later on, in 2016 and 2017, Khan et al. [7] and Vasamsetti et al. [8] proposed two methods that were wavelet based that can be used as a pre-processing step to increase the accuracy of high-level underwater computer vision tasks. Given the fact that these methods performed well in smearing noise, they introduced artefacts and made

noise more visible. Moreover, these methods suffered from low contrast, color deviations and loss of details [9]. Underwater environment conditions made these methods insufficient for enhancing underwater images as they cannot recover high quality underwater images.

2.2 Model-Based Methods

To overcome the drawbacks of model-free methods, model-based methods were used and proved to yield better outcomes. Model-based methods combines several model-free methods into a single image aiming to modify the image pixel values and as a result, improve the quality of the image. They can be divided into two subcategories which are physical and non-physical methods. One example of a non-physical model-based is two-step approach that was proposed in the research [10]. It contained both contrast enhancement and colour correction algorithms which generated promising outcomes and can be used in real-time applications. According to 8-bit images, the mean value is equal to 128. Based on this hypothesis, the colour correction technique used is based on piecewise linear transformation to spread the mean of the image until it reaches 128. A positive coefficient is used to ensure that the shifting range is logical to avoid overcorrection. Later, the objects and important details in the images should be highlighted and this is achieved by enhancing the contrast. This method proved to be suitable for real time applications.

In Jan. 2018, a method was proposed that did not depend on information about the underwater conditions for the image captured [11]. The original degraded image passes by a sequence of steps, starting with white balancing which has a main purpose of removing the colour casts that were introduced by underwater light scattering. Once a white-balanced version is produced, the image is passed in parallel to two other techniques, gamma correction and sharpening. Sharpening technique uses un-sharp masking principle in which a version blurred of the original image is added to the original image itself. This results in an image that is less blurred. These two output images are merged based on a weight map of multistate fusion algorithm and finally an enhanced image is produced. The main drawbacks of this method are that some haze is still present and could not be removed particularly in images that are taken in regions very far from the camera and that the colours of the images could not always be fully restored [11]. Another line of research modified existing algorithms such as Dark Channel Prior (DCP) [12] were combined with algorithms such as wavelength dependent compensation [13] to restore underwater images. Underwater Dark Channel Prior (UDCP) was proposed given the fact that information of the red channel is undependable [14].

Retinex based models were also made good use of in the research proposed by Zhang et al. [15]. The enhancement method consisted of colour correction, layer decomposition and enhancement. Fu et al. [16] proposed an extended multi-scale retinex-based method that was used also for general reasons such as enhancing sandstorm images. Physical model-based methods usually follow same, specific procedure. It treats the enhancing problem as an inverse problem. The procedure starts with building the model and then estimating the unknown parameters and finally handling the inversed problem. The current techniques implemented using physical model-based methods suffer from unstable and visually unsatisfying outcomes because they are built on the assumption that the attenuation coefficients are uniform across as they are only properties of the water.

2.3 Data-Driven Methods

Over the last ten years, deep learning was made good use of in low-level vision problems. However, the performance and the amount of deep learning-based enhancement techniques for underwater images does not match the success of recent deep learning based low level vision problems. This is because trained proposed CNN models on synthetic underwater images does not always generalize to real world cases. In 2017, Perez et al. [17] proposed a method based on CNN that trains an end-to-end transformation model between the distorted images and its corresponding clear images. Another method was proposed by Wang et al. [18] named UIEnet (Underwater Image Enhancement-net) that aimed to correct colours and remove the haze. It adopted a pixel disrupting strategy to extract the inherent features of local patches available in the images. This helped to fasten the model convergence and improved its accuracy. Later, in 2018, Anwar et al. [19] proposed a model named Underwater Convolutional Network (UWCNN) and used it to reconstruct the clear underwater latent images. It was trained on images from different databases covering images from different scenes and conditions. This method was able to tackle the problem of colour casts; however, due to the limitation of training images, the output images produced suffered from low dynamic range and appeared too hazy. To solve these problems extra post processing was required. However, this method did not yield best results with all testing images. A model named Underwater Resnet (UResnet) is introduced in the research [20]. This model improved the visual appearance of the images; however, one of the main drawbacks of using residual blocks in models is the network gets deeper, it takes a lot of time to train the model that can reach to several weeks and months. In this case, a special process GPU is required to speed up the training.

Moreover, Generative Adversarial Network (GANs) are recently made use of in many fields such as image

generation, video generation and voice recognition. Some researchers used GANs in their proposed solution to enhance underwater images. Fabbri et al. [21] used CycleGAN to reconstruct distorted images based on the undistorted images. These pairs were fed to train an underwater-GAN which can transform hazed underwater images to clear enhanced images. Li et al. [22] proposed a weakly supervised underwater colour correction model and weak supervision means that the model relaxes the need of paired underwater images for training. As discussed earlier, many researchers started using GAN-based models and some of them yielded good results; however, GANs are usually prone to training instability and are time consuming [23]. Moreover, the generated images tend to contain inconsistent stylizations with undesirable artifacts. As a result, it is found that using end-to-end networks yields much better results and is selected for the task.

3. Methodology

Neural networks have proven to be successful in solving many problems from different fields such as classification, clustering, compression and many more. A simple architecture of a neural network consists of neurons or nodes that make up the layers of the network. The proposed solution is based on an end-to-end CNN model which is based on machine learning as illustrated in Figure 2. It is assumed that our proposed methodology would outperform the existing state of the art techniques.
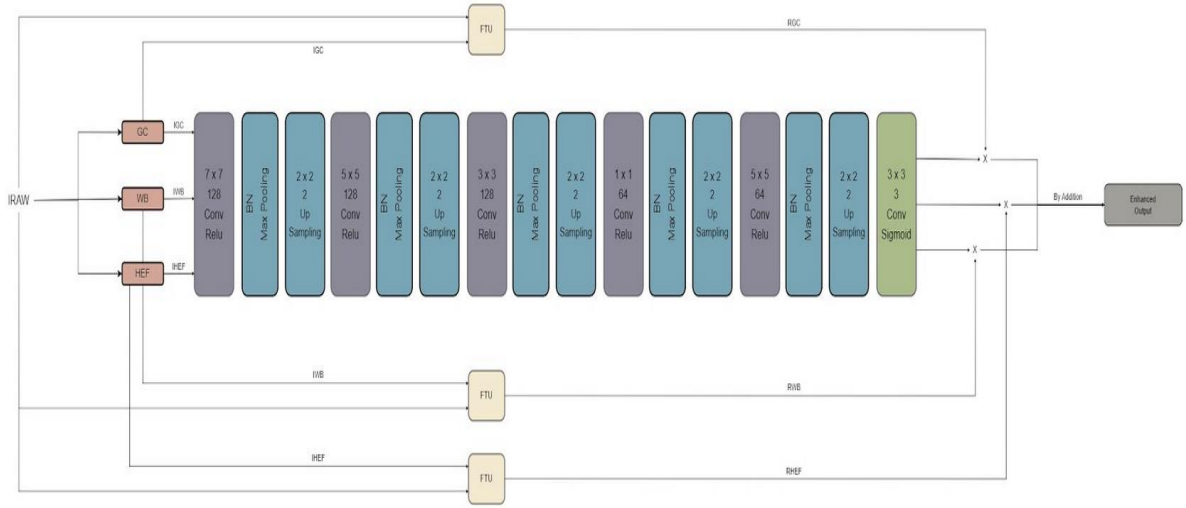

Figure 2. Proposed model architecture

End-to-end means that the network takes the input image from one end and produces the output image at the other end of the model. The model uses a gated fusion network to learn three confidence maps. As we discussed earlier, there are two types of algorithms, spatial domain, and frequency domain techniques. The model takes a single input image and performs three enhancing algorithms. Two of these algorithms work on the spatial domain whereas the third algorithm on the frequency domain. The methods are Gamma Correction (GC), White Balancing (WB) and High Frequency Emphasis Filtering (HEF). Techniques from both categories are used to make the most benefit of the enhancing algorithms. Spatial domain techniques deal with the image as it is and enhances the overall contrast of the image. On the other side, frequency domain techniques give us the control over the whole image and allows us to observe various characteristics that were not visible in the spatial domain. Each algorithm tackles a specific problem from the issues mentioned earlier. These three enhanced versions along with the raw image are fed into the model. This step produces three derived images (I_{GC}, I_{WB}, I_{HEF}). The derived images along with the raw image are fed to the network. As it is not possible to feed multiple input images to the network in parallel, the 4 input images were concatenated into one single numpy array. The output of the layers are three confidence maps namely, (C_{GC}, C_{HEF}, C_{WB}). Multiplication of the refined inputs with their corresponding confidence maps produces the final enhanced output as seen in Eq. (1).

$$I_{en} = I_{GC}*C_{GC} + I_{WB}*C_{WB} + I_{HEF}*C_{HEF} \qquad (1)$$

The model architecture includes three Feature Transformation Units (FTU) as shown in Figure 3 which has a main purpose of reducing the colour casts and artefacts introduced by the enhancing techniques mentioned earlier. Filter sizes of the layers are chosen to be (7*7) then decreases to (5*5) and finally to (3*3) to be more specific to

features in the images fed. Batch normalization layers were added to the main model and the FTUs as it makes training faster and more stable. It helps to solve Internal Covariate Shift problem that occurs due to the change of parameters in each layer which leads to change in the distribution of the inputs to subsequent layers. As a result, the learning process is faster, stabilized and the number of training epochs are reduced. Moreover, to make the model even more efficient max pooling layers are added to reduce the computational costs. This is done by reducing the number of parameters that the model has to learn. It is a standard benchmarking technique that is implemented to make benefit from the feature as it selects the brighter pixels from the image, and this is useful when the background is dark, and this is usually the case in underwater images.
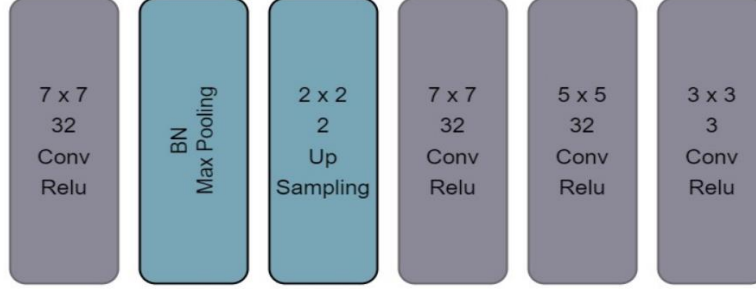

Figure 3. FTU model

The perceptual loss function is a fucntion for comparing two images that are very similar to one another in terms of content and style discrepancies. In this work it is based on the implementation of the Manhathan distance. It is the distance between the feature represntations of both the enhanced (which is represented by (I_{en})) and the reference (which is represented by (I_{RAW})) image. The Manhathan distance is a method of calculating the absolute difference distance between two points and is normally used in the case of high dimensional data. It is better than Euclidean distance which takes the square root of the sum of square values of differences between two points because it gives more robust results.

$$d(p,q) = \sqrt{\sum_{i=1}^{n}(q_i - p_i)^2} \tag{2}$$

Manhathan distance shown in Eq. (3) has ($C_j\,H_j\,W_j$) which is the dimensions of the feature map; number, height and width respectively. Additionally, N is the number of each batch in the training process whereas $\emptyset$ is the number of layers in the network and j are iterations over the layers. The equation is performed at each layer and the sum of all results is calculated. Only 8 layers were used to ensure that no overfitting occurs. In other words, if many layers are used, the model will learn the training data more than it should be and will negatively affect the performance of the model. The model will not be able to generalize when new data are introduced during the testing phase.

$$L_j^{\emptyset} = \frac{1}{C_i H_j W_j} \sum_{i=1}^{N} \left\| \emptyset_{j(I_{en}^i)} - \emptyset_{j(I_{RAW}^i)} \right\| \tag{3}$$

3.1 Gamma Correction

The GC technique is used to lighten up the dark areas in the images and its gamma value varies according to the purpose of using the technique. For example, in this work the gamma value γ is chosen to be 0.7. When the gamma value is less than 1, the process is named encoding gamma correction. Eq. (4) is used to calculate the corrected output represented by $g(x)$ [24], x is every single pixel value and γ represents gamma value.

$$g(x) = 255\left(\frac{x}{255}\right)^{\left(\frac{1}{\gamma}\right)} \tag{4}$$

3.2 White Balancing

The WB technique is used to correct the color casts which is unwanted tint of a particular color in the image by discarding the unwanted ones as a result of various illuminations. Due to very poor light propagation in underwater

scenes, there seems to be a lack of contrast. However, at water levels deeper than 30 feet, white balancing technique is not very efficient as it is very difficult to restore colors that were absorbed [25].

3.3 High Frequency Emphasis Filtering (HEF)

One of the mentioned problems in underwater images was the lack of image sharpness and images seem to have blurry appearance. High frequency emphasis filtering (HEF) technique [26] is used with a modification of using contrast limited adaptive histogram equalization (CLAHE) instead of histogram equalization. It consists of a sequence of steps and are represented in Figure 4.
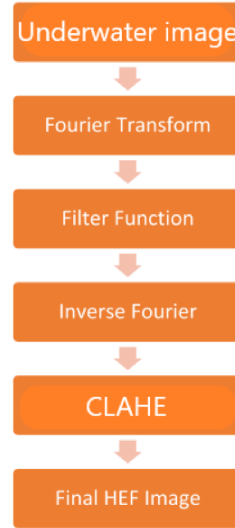

Figure 4. HEF algorithm

The first step in HEF is to convert the image into its frequency domain representation. Then the filter function is applied and in this case Gaussian high-pass filter is used to accentuate and emphasize the edges. Eq. (5) represents the filter function where Do is the cut off distance.

$$Gaussian\ Filter\ =\ 1 - e^{-\frac{D^2(i,j)}{2D_0^2}} \tag{5}$$

The 2D Fourier transform of *F(x, y)* and inverse Fourier transform of *F(i, j)* are denoted by Eqns. (6) and (7) respectively where *x* and *i* takes value starting from 0, 1, 2 till *M*−1 and *y* and *j*=0, 1, 2 till *N*−1.

$$F(i,j) = \sum_{X=0}^{M-1}\sum_{y=0}^{N-1} f(x,y)\, e^{-j2\pi\left(\frac{ix}{M}+\frac{jy}{N}\right)} \tag{6}$$

$$F(x,y) = \frac{1}{M}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1} f(i,j)\, e^{-j2\pi\left(\frac{ix}{M}+\frac{jy}{N}\right)} \tag{7}$$

In the high frequency spectrum, the expressed edges have more significant changes in frequencies. As a result, a low contrast image is produced and that is the reason why (CLAHE) is used in the last step, to increase sharpness and contrast. The HE technique is used to adjust image intensities to enhance the contrast of the image. In this work a special type of histogram equalization is used named contrast limited adaptive histogram equalization (CLAHE) [27] and it is illustrated in Figure 5. Subgraph (a) of Figure 5 shows the original histogram of the image whereas Subgraph (b) of Figure 5 shows the pixels distribution after the histogram has been clipped. It performs histogram equalization technique in small patches achieving high accuracy and contrast limited. A threshold is set, and any pixel value exceeds this threshold is cut by a clipper before computing the cumulative distribution function. The part clipped is then equally distributed upon all histogram bins. The selected threshold is 2.0 in CLAHE which helps limit the height of the histogram [28]. This means that the slope of the cumulative distribution function curve will be reduced. In other words, contrast enhancement is reduced to limit not only noise amplification, but also local over enhancement.
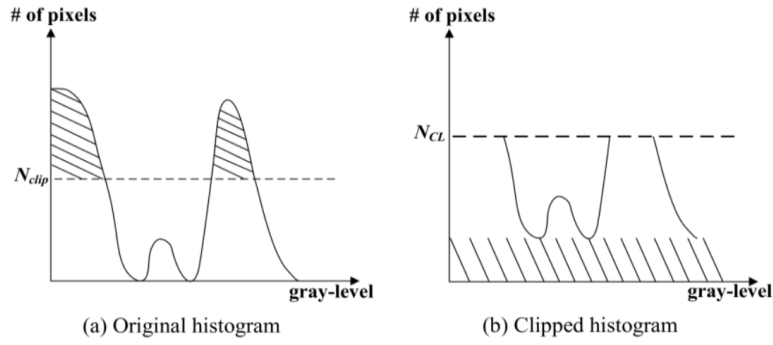
Figure 5. CLAHE algorithm

4. Results

The model was implemented using keras and trained on Google Pro colab. Adaptive moment estimation (ADAM); a first order gradient based was used for optimization. It is considered as one of the most optimal optimizers as it requires minimal memory requirements and is straightforward to implement [29]. It combines the heuristics of momentum and RMSProp optimizers which makes it more optimal, and able to handle sparse gradients on noisy problems. It is used in models that has more than one hidden layer by using the squared gradients to scale the learning rate which is one of the four hyperparameters it has. Assigning the learning rate to a small number is good approach as it would allow the weights to reach a minimum during training which leads to better accuracy. To achieve good reliable results and improve the data model prediction accuracy, data augmentation was applied as a pre-processing step on the dataset. It helps expose our model to different versions of the data which increases its generalizing ability and decreases the chance of overfitting occurring.

4.1 Experiments

There are several underwater images datasets available for different research purposes such as image enhancement, object detection, etc. Enhancement of Underwater Visual Perception dataset (EUVP) has a collection of 890 underwater images that were collected at various conditions such as oceanic explorations and human-robot collaborative experiments [30]. HICRD dataset consists of 6665 unpaired images collected from several sample locations around Heron Reef that is located in the Southern Great Barrier Reef [31]. For training our model we used the dataset called UIEB (Underwater Image Enhancement Benchmark) [32]. This dataset was also used by other state-of-the-art techniques, so we decided to stick to it to have a fair comparison. Another reason for choosing the UIEB dataset is that it has corresponding ground truth images which are needed while training the model. 12 filtering methods were performed on the raw images to produce 12 reference images and only one is selected based on the most choice selected by 50 participants. The dataset contains images taken at various diverse scenes that has different objects such as corals, rocks, sculptures, etc as seen in Figure 6. The images have various image quality degradation characteristics. Figure 7 has subjective comparisons on two images from the dataset.
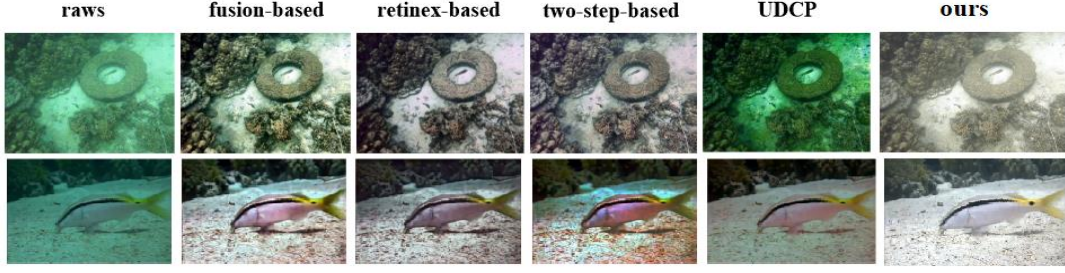

Figure 6. UIEB dataset sample

Figure 7. Results comparisons

4.2 Metrics

There are several image quality metrics that can be used to measure and determine the quality of images after they have been enhanced using various enhancing techniques. These metrics could be divided into categories namely, full reference metrics and non-reference metrics that will be explained in the next sub section [33].

4.2.1 Full reference metrics

Full reference metrics techniques perform a direct comparison between the enhanced image and the raw image fed to the model. In this work, two full reference techniques were namely Structural Similarity Index (SSIM) [34] and Peak Signal to Noise Ratio (PSNR) [35]. In SSIM the two images are represented by windows x and y. Eq. (8) is the formulae of SSIM where μ_x is the average of values in x, μ_y is the average of values in y, σ_x^2 is the variance in window x, σ_y^2 is the variance in window y, σ_{xy} represents the covariance of both x and y, c_1 and c_2 are variables used to stabilize the division in the equation using a weak denominator. The value ranges between: [-1, +1] and will be equal to 1 if the two images are identical.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{8}$$

PSNR is calculated as shown in Eq. (9) and Eq. (10). MSE stands for mean squared error which simply represents the mean of the squares of errors between the enhanced image and the distorted one and The error is the diferrence in pixel values of both images. Y is the observed vector, whereas $\hat{Y}$ is the predicted vector produced as an output from the model. The rule for MAX_f is $2^{(n)}-1$ which is 255 because the images are 8 bits where n is the number of bits. If we assume that the value for MSE is 1 then this means that $20 \log_{10} (255)$ is 48. In other words, the ideal highest value is 48 for a 8 bit image.

$$PSNR = 20 \log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right) \tag{9}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 \tag{10}$$

Table 1 shows the results of the previous state of the art techniques along with the results of our proposed model using the testing dataset UIEB mentioned earlier. Given the fact that SSIM and PSNR are quantitative metrics, they have a major drawback when it comes to the case of evaluating underwater image enhancement techniques which they do not consider any of the essential biological factors concerned with the human vision system. As a result, qualitative metrics should also be used which will be discussed in the next sub-section.

Table 1. Full-reference evaluation results

Method	SSIM	PSNR
Fusion-based [11]	0.8162	18.7461
UDCP [14]	0.4999	11.0296
Two-step based [10]	0.7199	18.7461
Retinex-based [15]	0.6233	16.8757
Water-Net [32]	0.7971	19.1130
Ours	**0.8995**	**22.9865**

4.2.2 Non reference metrics

Usually, it is easy to obtain ground truth images for image processing problems such as image super resolution. However, in underwater image enhancement problem it is challenging to achieve large number of paired images. Due to this reason non-reference metrics were proposed. In this work we evaluate our results using two of these techniques named, Underwater Colour Image Quality Evaluation (UCIQE) [36] and Underwater Image Quality Measure (UIQM) [37]. In UCIQE metric the blurring effect, low contrast and non-uniform color casts are linearly combined. In this technique the deviation of saturation is not used unlike other previous proposed method because it emphasizes the dark areas which are simply the result of images taken in limited lighting. In Eq. (11) the σ_c represents the standard deviation value of Chroma and ranges from (-∞, +∞), con_1 represents the contrast of luminance which ranges from [0, +∞), and represents the global grey scale distribution of the given image. The μ_S represents the average saturation and ranges from (-∞, +∞). Chroma is one of the properties in an image and represents the degree of color clarity and purity. The variables c_1, c_2, c_3 represent weighted coefficients. Several coefficient values were experimented according to several experiments made in the research [36] and the following values proved to yield best results: $c_1 = 0.4680$, $c_2 = 0.2745$, $c_3 = 0.2576$.

Minimum value for UCIQE is 0 and as the value increases, this means that the model has better performance.

$$UCIQE = c_1 * \sigma_c + c_2 * con_1 + c_3 * \mu_S \qquad (11)$$

The underwater image quality measure (UIQM) metric uses three measures: which are the underwater image colorfulness measure (UICM), the underwater image contrast measure (UIConM) and the underwater image sharpness measure (UISM). Sharpness is a property of images that describes the clarity of edges and important fine details. In UISM, the Sobel edge detector algorithm is used on every channel of RGB. This yields three edge maps which are then multiplied with the original channel values to produce the grey scale edge maps. This preserves only the pixels representing the edges in an underwater image. UICM is used to remove the effect of bright regions due to heavy noise in underwater images. The three measures are linearly combined as shown in Eq. (12). The values given to the weights are dependent on the type of application that the underwater image is used into. Several coefficients values were experimented according to several experiments made in the research [37] and the following values proved to yield best results: $c_1 = 0.0282$, $c_2 = 0.2953$, $c_3 = 3.5753$.

$$UIQM = c_1 * UICM + c_2 * UISM + c_3 * UIConM \qquad (12)$$

As both metrics UCIQE and UIQM use Chroma of the image as one of their variables, this makes it essential to convert colour space from red, green and blue Red-Green-Blue (RGB) to LAB. Representation of a LAB image is $l*a*b$ where l stands for lightness of the image and ranges from[0, 100] that is from black to white, a stands for red/green component and ranges [-120, +120] and b stands for yellow/blue component and has the same range of component a [-120, +120]. In LAB colour space, a change in numerical values usually corresponds to approximately the same amount of change in how it is visually perceived [38]. As shown in Figure 8, neither a cannot be both red and green at the same time, nor b can be blue and yellow at the same time. This is made clear by the coordinate axes which represents valu es running from positive to neagtive values. For example, positive a values indicate red colour, whereas a negative value repesents green colour. The same concept implies for b where a negative value indicates blue colour whereas a positive value indicates yellow colour. Table 2 represents the results when the previous enhancing approaches namely fusion-based and two-step-based were tested with evaluation metrics; UCIQE and UIQM. Both methods are used to rank under water image enhancing algorithms performance. A higher score achieved by the UCIQE [36] metric indicates better balance among saturation, contrast, and Chroma whereas a higher result of UIQM [37] indicates that the output is much more consistent with the human visual perception.
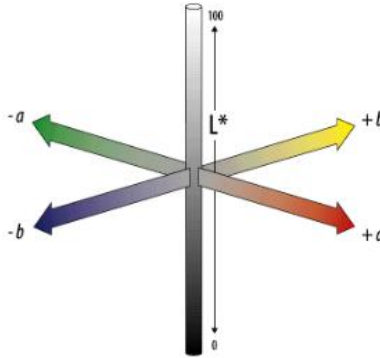

Figure 8. CIELAB color space representation [38]

Table 2. Non-reference evaluation results

Method	UCIQE (↑)	UIQM (↑)
Fusion-based [11]	0.6414	1.5310
UDCP [14]	0.5852	1.6297
Two-step based [10]	0.5776	1.4002
Retinex-based [15]	0.6062	1.4338
Water-Net [32]	0.6983	1.7216
Ours	**0.7851**	**1.9867**

5. Conclusions

In this work, we have discussed the importance of underwater images in different aspects of life such as marine engineering, control of underwater vehicles, etc. Moreover, main challenges that face underwater images and their corresponding effects on the images were also discussed. The main effects were low contrast, colour casts and blurry appearance. We presented a simple and efficient model that is based on the use of CNN to enhance underwater images that are taken at different environmental conditions. Results of our model were presented and compared to previous state of the art methods using both full-reference and non-reference metrics. Our model scored better results in both types of metrics that were used for evaluation. In the future work, we aim to extend investigating our model's feasibility on real-time videos.

Data Availability

The data used to support the research findings are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] Y. D. Liu, H. P. Xu, D. H. Shang, C. Li, and X. Q. Quan, "An underwater image enhancement method for different illumination conditions based on color tone correction and fusion-based descattering," *Sensors*, vol. 19, no. 24, Article ID: 5567, 2019. https://doi.org/10.3390/s19245567.

[2] 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 Proceedings of the First Conference on Visualization in Biomedical Computing, Atlanta, GA, USA, pp. 337-345, 1990. https://doi.org/10.1109/VBC.1990.109340.

[3] 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, 1995. https://doi.org/10.1109/30.468045.

[4] J. van de Weijer, T. Gevers, and A. Gijsenij, "Edge-Based Color Constancy," *IEEE T. Image Process*, vol. 16, no. 9, pp. 2207-2214, 2007. https://doi.org/10.1109/TIP.2007.901808.

[5] G. Singh, N. Jaggi, S. Vasamsetti, H. K. Sardana, S. Kumar, and N. Mittal, "Underwater image/video enhancement using wavelet based color correction (WBCC) method," In 2015 IEEE Underwater Technology, (UT), Chennai, India, February 23-25, 2015, IEEE, pp. 1-5. https://doi.org/10.1109/UT.2015.7108303.

[6] A. M. Grigoryan and S. S. Agaian, "Color image enhancement via combine homomorphic ratio and histogram equalization approaches: Using underwater images as illustrative examples," *Int J. Future Revolution Comput. Sci. Commun. Eng.*, vol. 4, no. 5, pp. 36-47, 2018.

[7] A. Khan, S. S. A. Ali, A. S. Malik, A. Anwer, and F. Meriaudeau, "Underwater image enhancement by wavelet based fusion," In 2016 IEEE International Conference on Underwater System Technology: Theory and Applications, (USYS 2016), Penang, Malaysia, December 13-14, 2016, IEEE, pp. 83-88. https://doi.org/10.1109/USYS.2016.7893927.

[8] S. Vasamsetti, N. Mittal, B. C. Neelapu, and H. K. Sardana, "Wavelet based perspective on variational enhancement technique for underwater imagery," *Ocean Eng.*, vol. 141, pp. 88-100, 2017. https://doi.org/10.1016/j.oceaneng.2017.06.012.

[9] Y. Wang, W. Song, G. Fortino, L. Z. Qi, W. Q. Zhang, and A. Liotta, "An experimental-based review of image enhancement and image restoration methods for underwater imaging," *IEEE Access*, vol. 7, pp. 140233-140251, 2019. https://doi.org/10.1109/ACCESS.2019.2932130.

[10] X. Y. Fu, Z. W. Fan, M. Ling, Y. Huang, and X. H. Ding, "Two-step approach for single underwater image enhancement," In 2017 International Symposium on Intelligent Signal Processing and Communication Systems, (ISPACS), Xiamen, China, November 6-9, 2017, IEEE, pp. 789-794.

https://doi.org/10.1109/ISPACS.2017.8266583.

[11] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, and P. Bekaert, "Color balance and fusion for underwater image enhancement," *IEEE T. Image Process*, vol. 27, no. 1, pp. 379-393, 2018. https://doi.org/10.1109/TIP.2017.2759252.

[12] K. M. He, J. Sun, and X. O. Tang, "Single image haze removal using dark channel prior," In 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, June 20-25, 2009, IEEE, pp. 1956-1963. https://doi.org/10.1109/CVPR.2009.5206515.

[13] J. Y. Chiang and Y. C. Chen, "Underwater image enhancement by wavelength compensation and dehazing," *IEEE T. Image Process*, vol. 21, no. 4, pp. 1756-1769, 2012. https://doi.org/10.1109/TIP.2011.2179666.

[14] P. L. J. Drews, E. R. Nascimento, S. S. C. Botelho, and M. F. Montenegro Campos, "Underwater depth estimation and image restoration based on single images," *IEEE Comput. Graph.*, vol. 36, no. 2, pp. 24-35, 2016. https://doi.org/10.1109/MCG.2016.26.

[15] S. Zhang, T. Wang, J. Y. Dong, and H. Yu, "Underwater image enhancement via extended multi-scale Retinex," *Neurocomputing*, vol. 245, pp. 1-9, 2017. https://doi.org/10.1016/j.neucom.2017.03.029.

[16] X. Y. Fu, P. X. Zhuang, Y. Huang, Y. H. Liao, X. P. Zhang, and X. H. Ding, "A retinex-based enhancing approach for single underwater image," In 2014 IEEE International Conference on Image Processing, Paris, France, October 27-30, 2014. https://doi.org/10.1109/ICIP.2014.7025927.

[17] J. Perez, A. C. Attanasio, N. Nechyporenko, and P. J. Sanz, "A deep learning approach for underwater image enhancement," In Biomedical Applications Based on Natural and Artificial Computing, IWINAC 2017, Corunna, Spain, June 19-23, 2017, Springer, pp. 183-192. https://doi.org/10.1007/978-3-319-59773-7_19.

[18] Y. Wang, J. Zhang, Y. Cao, and Z. F. Wang, "A deep CNN method for underwater image enhancement," In 2017 IEEE International Conference on Image Processing, (ICIP), Beijing, China, September 17-20, 2017, IEEE, pp. 1382-1386. https://doi.org/10.1109/ICIP.2017.8296508.

[19] S. Anwar, C. Y. Li, and F. Porikli, "Deep underwater image enhancement," *arXiv*, vol. 2018, 2018. https://doi.org/10.48550/arXiv.1807.03528.

[20] P. Liu, G. Y. Wang, H. Qi, C. F. Zhang, H. Y. Zheng, and Z. B. Yu, "Underwater image enhancement with a deep residual framework," *IEEE Access*, vol. 7, pp. 94614-94629, 2019. https://doi.org/10.1109/ACCESS.2019.2928976.

[21] C. Fabbri, M. J. Islam, and J. Sattar, "Enhancing Underwater imagery using generative adversarial networks," In 2018 IEEE International Conference on Robotics and Automation, (ICRA), Brisbane, QLD, Australia, May 21-25, 2018, IEEE, pp. 7159-7165. https://doi.org/10.1109/ICRA.2018.8460552.

[22] C. Y. Li, J. C. Guo, and C. L. Guo, "Emerging from water: underwater image color correction based on weakly supervised color transfer," *IEEE Signal Proc Let.*, vol. 25, no. 3, pp. 323-327, 2018. https://doi.org/10.1109/LSP.2018.2792050.

[23] H. H. Yang, K. C. Huang, and W. T. Chen, "LAFFNet: A lightweight adaptive feature fusion network for underwater image enhancement," In 2021 IEEE International Conference on Robotics and Automation, (ICRA), Xi'an, China, May 30-June 5, 2021, IEEE, pp. 685-692.

[24] G. Xu, S. Jian, H. D. Pan, Z. G. Zhang, and H. B. Gong, "An image enhancement method based on gamma correction," In IEEE 2009 Second International Symposium on Computational Intelligence and Design, Changsha, China, December 12-14, 2009, IEEE, pp. 60-63. https://doi.org/10.1109/ISCID.2009.22.

[25] C. Ancuti, C. O. Ancuti, T. Haber, and P. Bekaert, "Enhancing underwater images and videos by fusion," In 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence, RI, USA, June 16-21, 2012, IEEE, pp. 81-88. https://doi.org/10.1109/CVPR.2012.6247661.

[26] K. Munadi, K. Muchtar, N. Maulina, and B. Pradhan, "Image enhancement for tuberculosis detection using deep learning," *IEEE Access*, vol. 8, pp. 217897-217907, 2020. https://doi.org/10.1109/ACCESS.2020.3041867.

[27] Z. Xu, X. Liu, and N. Ji, "Fog removal from color images using contrast limited adaptive histogram equalization," In 2009 2nd International Congress on Image and Signal Processing, Tianjin, China, October 17-19, 2009. https://doi.org/10.1109/CISP.2009.5301485.

[28] Q. Q. Fu, Z. B. Zhang, M. Celenk, and A. P. Wu, "A Poshe-based optimum clip-limit contrast enhancement method for ultrasonic logging images," *Sensors,* vol. 18, no. 11, Article ID: 3954, 2018. https://doi.org/10.3390/s18113954.

[29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv, vol. 2017, 2017. https://doi.org/10.48550/arXiv.1412.6980.

[30] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robot. Autom. Let.*, vol. 5, no. 2, pp. 3227-3234, 2020. https://doi.org/10.1109/LRA.2020.2974710.

[31] J. L. Han, M. Shoeiby, T. Malthus, E. Botha, J. Anstee, S. Anwar, R. Wei, M. A. Armin, H. D. Li, and L. Petersson, "Underwater image restoration via contrastive learning and a real-world dataset," arXiv, vol. 2021, 2021. https://doi.org/10.48550/arXiv.2106.10718.

[32] C. Y. Li, C. L. Guo, W. Q. Ren, R. M. Cong, J. H. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," arXiv, vol. 2019, 2019. https://doi.org/10.48550/arXiv.1901.05495.

[33] S. Raveendran, M. D. Patil, and G. K. Birajdar, "Underwater image enhancement: A comprehensive review, recent trends, challenges and applications," *Artif Intell Rev.,* vol. 54, pp. 5413-5467, 2021. https://doi.org/10.1007/s10462-021-10025-z.

[34] A. C. Brooks, X. Zhao, and T. N. Pappas, "Structural similarity quality metrics in a coding context: Exploring the space of realistic distortions," *IEEE T. Image Process*, vol. 17, no. 8, pp. 1261-1273, 2008. https://doi.org/10.1109/TIP.2008.926161.

[35] D. Salomon, *Data Compression*, UK, London: SpringerLink, Springer, pp. 281-282, 2007.

[36] M. Yang and A. Sowmya, "New image quality evaluation metric for underwater video," IEEE Signal Proc Let., vol. 21, no. 10, pp. 1215-1219, 2014. https://doi.org/10.1109/LSP.2014.2330848.

[37] K. Panetta, C. Gao, and S. Agaian, "Human-visual-system-inspired underwater image quality measures*,*" *IEEE J. Magazine*, vol. 41, no. 3, pp. 541-551, 2016. https://doi.org/10.1109/JOE.2015.2469915.

[38] D. J. Bora, A. K. Gupta, and F. A. Khan, "Comparing the performance of LAB and HSV color spaces with respect to color image segmentation," arXiv, vol. 2015, 2015. https://doi.org/10.48550/arXiv.1506.01472.