UNDER WATER IMAGE ENHANCEMENT

Amit Mohapatra
202218033
MSc Data Science

K. Sethu Srivatsa
202218021
MSc Data Science

Project Supervisor: - Prof. Srimanta Mandal

Abstract— The realm of underwater imaging holds significant importance in various research domains, including marine biology, oceanography, and underwater exploration. Unfortunately, underwater images commonly encounter challenges such as diminished colour vibrancy and contrast, attributed to light scattering and absorption in water. This project introduces a novel approach aimed at augmenting the colour and contrast of underwater images. The effectiveness of the proposed method is assessed using a dataset of underwater images, revealing notable enhancements in the visual quality of the images.

I. INTRODUCTION

Underwater imaging poses formidable challenges arising from the intricate interplay of light scattering and absorption in aquatic environments. This often results in underwater images characterized by lackluster colour rendition and diminished contrast, posing a hindrance to the extraction of meaningful information. The conundrum is particularly pronounced in pivotal disciplines like marine biology, oceanography, and underwater exploration, where precise and detailed imaging is imperative. Existing strategies, encompassing image restoration, colour correction, and fusion, image enhancement has sought to address these issues. Unlike the clarity associated with an ideal transmission medium, underwater conditions introduce a myriad of complexities. The available light underwater is contingent upon various factors, including sunlight interaction with the sea surface influenced by time of day and sea conditions. Diving locations further compound the challenge, introducing location-specific colour casts that imbue underwater scenes with distinct hues. Moreover, the density of particles in seawater, significantly denser than the atmosphere, results in the gradual absorption of different light wavelengths. Red wavelengths, the longest, are absorbed first, followed by orange and yellow, manifesting as a discernible loss of red in images taken at shallower depths.

The underwater milieu introduces additional complexities such as blurring, distortion due to light scattering, underwater noise, absorption, colour-cast, nonuniform illumination, and shading. To surmount these hurdles, we propose an algorithm designed to elevate the quality of underwater images by rectifying distortions induced by underwater haze, mitigating colour-cast, and identifying objects within distorted underwater frames. This undertaking represents a significant contribution to the realm of underwater imaging, offering a pragmatic and efficient solution to enhance the quality of submerged visual data. [1]

II. IMAGE RECONSTRUCTION USING PRINCIPAL COMPONENT ANALYSIS:

Principal Component Analysis (PCA) is a fundamental technique in image enhancement, commencing with the collection of a set of training images represented as high-dimensional vectors. These vectors are standardized by subtracting the mean image vector, thereby aligning them with the matrix A, derived from the eigenvectors of the covariance matrix. This transformation is expressed as

$$\mathbf{y} = \mathbf{A}\,(x - \mathbf{M}x) \tag{1}$$

where x is the matrix of standardized image vectors. The covariance matrix of the transformed vectors, denoted as Cy, is then computed as

$$\mathbf{C}_y = \mathbf{A}C_x\mathbf{A}^{\mathbf{T}} \tag{2}$$

The subsequent step involves calculating the eigenvectors and eigenvalues of Cy to establish the eigen image basis. Notably, the matrix A can be truncated to retain only the eigenvectors with the highest eigenvalues, signifying the most significant modes of variation in the image data—an approach known as the Hotelling transform. For image reconstruction using the eigen image basis, the inverse transformation is applied through the formula

$$\mathbf{x} = \mathbf{A}^{\mathbf{T}}\mathbf{y} + Mx \tag{3}$$

The coefficients necessary for this process can be computed using the formula

$$\mathbf{c} = \mathbf{A}^{T}\,(\mathbf{y} - \mathbf{M}x) \tag{4}$$

where y represents the matrix of transformed image vectors. This encapsulates the essence of PCA in image enhancement, forming the foundation for subsequent techniques employed in our project. [2]

The image reconstruction process utilizing Principal Component Analysis (PCA) involved a detailed sequence of operations aimed at addressing the challenges posed by underwater image distortion. Beginning with the division of the image into its Red, Green, and Blue (RGB) channels, we implemented a patch-based approach. This involved breaking down each channel into smaller patches of a defined size (10x10 pixels). These patches were pivotal in capturing localized variations within the image. Subsequently, we computed the mean of these patches, forming the basis for cantering the patches around this mean. The centering process ensured that the subsequent analysis was focused on the variations within the patches rather than absolute pixel values. Following this, the covariance matrix of the centred patches was computed. This matrix encapsulates the relationships and interactions between the different pixel values, providing insights into the variance distribution within the image. Eigen decomposition of the covariance matrix yielded eigenvalues and eigenvectors. Sorting these eigenvalues in descending order allowed us to prioritize the most significant modes of variation in the image data. Intriguingly, the mean eigenvalues of each channel were calculated, facilitating the introduction of a channel weight factor. This weight, calculated as the ratio of the global mean eigenvalues to the mean eigenvalues of the specific channel, was then used to update the eigenvectors. The channel weight factor, denoted as channel weight, played a crucial role in adjusting the eigenvectors to account for variations specific to each channel. The formula for this adjustment was articulated as follows:

$$C_w = \frac{\lambda_{mg}}{\lambda_{mc}} \tag{5}$$

Where C_w is the channel weight, λ_{mg} is the mean of eigen values of all the channel and λ_{mc} is the mean of the eigen values of the particular channel. Multiplying the eigenvectors by this channel weight factor ensured that the eigenvectors were appropriately scaled, aligning with the specific characteristics of each channel. Moving forward, the reconstructed image was formed through a series of dot product operations. The standardized patches, after centering, were transformed using the updated eigenvectors. This transformation involved subtracting the mean of the patches and then applying the dot product with the eigenvectors. The resulting transformed values were then used in the reverse transformation, where dot products with the transposed eigenvectors and the addition of the mean patches reconstructed the image. The utilization of patch-based analysis, mean-centering, covariance matrix computation, and eigen decomposition, along with the incorporation of channel-specific adjustments, collectively contributed to the success of the image reconstruction process. The amalgamation of these steps showcased the depth of our approach in tackling the challenges associated with underwater imaging.

(a)

(b)

Fig.1 PCA (a) is the Input image (b) is the Reconstructed image of the input image using the principal components.

Transitioning from PCA to address limitations encountered, our project explored the efficacy of autoencoders. This shift was prompted by the inadequacies of PCA in achieving the desired results. Autoencoders, a type of neural network, were employed to surpass the shortcomings of traditional methods. The adoption of autoencoders allowed for a more nuanced exploration of image reconstruction, leveraging the network's ability to learn complex non-linear patterns and representations within the data. The ensuing sections detail our methodological approach, specifically focusing on the implementation of autoencoders to enhance underwater image quality.

III. IMAGE ENHANCEMENT USING AUTOENCODERS

Autoencoders, a class of neural networks, have gained prominence in image enhancement tasks due to their ability to capture and reproduce intricate patterns within data. Autoencoders learn hierarchical representations of input data, making them adept at grasping complex relationships within images. [3]

A. Autoencoder Architecture:

The autoencoder architecture adopted in our project consists of an encoder and decoder. The encoder compresses the input image into a latent representation, and the decoder reconstructs the image from this compressed representation. The layers in both the encoder and decoder involve convolution operations, batch normalization, and activation functions. The loss function used for training combines L1 loss, L2 loss, and perceptual loss, enhancing the network's ability to capture both pixel-level and perceptual details :

(a)

(b)

Fig. 2 Autoencoder (a) is the input image (b) is the enhanced image using Autoencoder

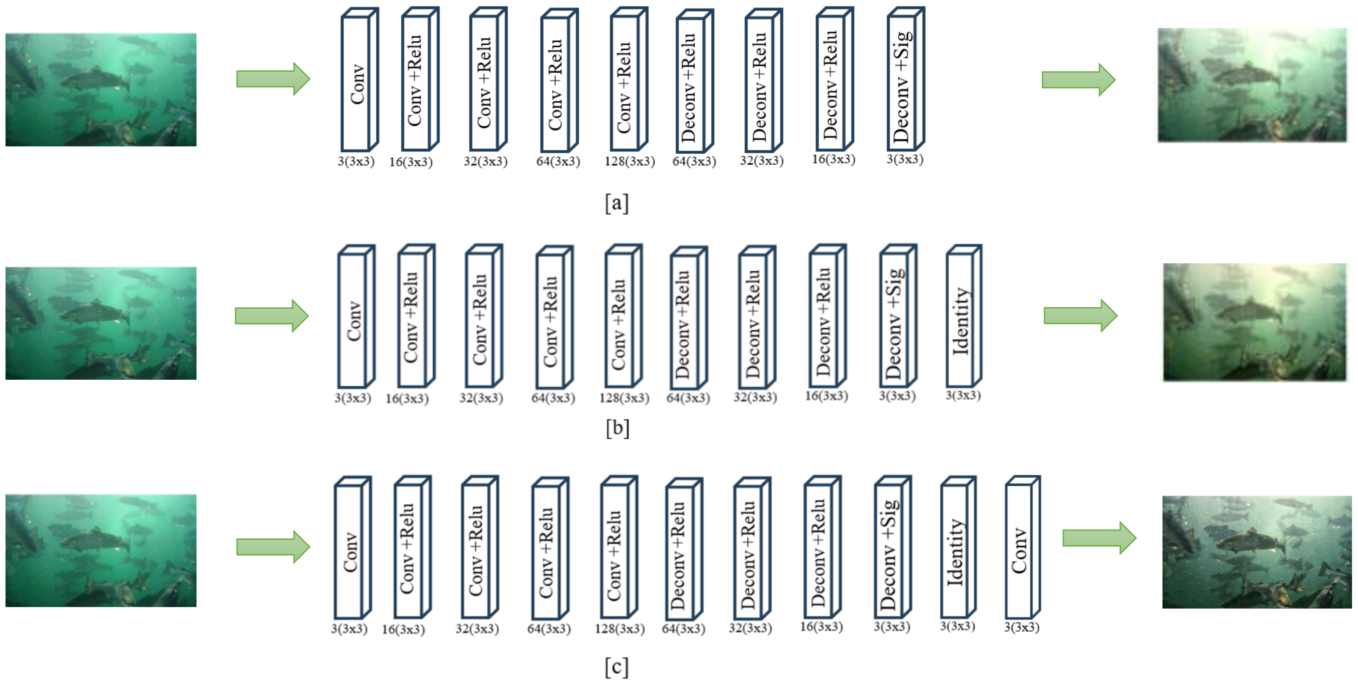


Fig.3 [a] Shows the Batch Normalization Architecture, [b] shows identity connection after decoding layers and [c] shows convolution block after skip connection.

B. *Architectural Variation*

Three architectural variations were explored in our experimentation. The first follows a straightforward encoder-decoder structure. The second incorporates a skip connection (identity connection) after the decoding layers, allowing the model to retain some of the original input information. The third variation introduces a small convolutional block after the skip connection, enhancing the flexibility of the model in capturing complex features. [4]

C. *Loss Function*

The adopted loss function, termed custom combined loss, is a composite of L1 loss, L2 loss, and perceptual loss. L1 and L2 losses measure pixel-wise differences between predicted and ground truth images. The perceptual loss leverages a pre-trained VGG19 model to compare feature representations, aligning with human perceptual judgments. [5]

$$CL = \lambda_1 \cdot \text{L1 Loss} + \lambda_2 \cdot \text{L2 Loss} + \lambda_3 \cdot \text{P. Loss} \quad (6)$$

$$\text{L1 Loss} = \frac{1}{N}\sum_{i=1}^{N}\left|y_{true}^{i} - y_{pred}^{i}\right| \quad (7)$$

$$\text{L2 Loss} = \frac{1}{N}\sum_{i=1}^{N}\left(y_{true}^{i} - y_{pred}^{i}\right)^2 \quad (8)$$

$$\text{P. Loss} = \frac{1}{N}\sum_{i=1}^{N}\left(VGG(y_{true}^{i}) - VGG(y_{pred}^{i})\right)^2 \quad (9)$$

Here λ1 , λ2 and λ3 are non-negative weighting coefficients controlling the relative importance of each loss term. y_{true} is the ground truth image, y_{pred} is the predicted image

This custom combined loss function offers several advantages over relying on a single loss term. By combining L1, L2, and perceptual losses, it addresses both low-level pixel-level details and high-level perceptual qualities, leading to more realistic and visually pleasing outputs. Additionally, the flexibility of weighting each loss term allows for fine-tuning the balance between these factors, tailoring the loss function to specific tasks and desired image characteristics. This combination of robustness and adaptability makes the custom combined loss a valuable tool for image generation and restoration tasks.

IV. Results

A. *Evaluation Metrics:*

To quantitatively assess the image enhancement, we computed UCIQE (Underwater Colour Image Quality Evaluation) and UIQM (Underwater Image Quality Metrics) scores. These metrics provide objective measures of image quality, considering factors like colourfulness, brightness, and sharpness. The computed scores serve as indicators of the effectiveness of our image enhancement approach. In summary, our work introduces and explores three variations of autoencoder architectures for underwater image enhancement. The use of perceptual loss, skip connections, and additional convolution blocks aims to enhance the network's ability to preserve important details and improve visual quality.[6]

UCIQUE/UIQM 0.83/0.89 0.76/0.82 0.65/0.79

UCIQUE/UIQM 0.74/0.88 0.75/0.78 0.71/0.89

UCIQUE/UIQM 1.00/0.79 1.07/0.78 0.79/0.96

UCIQUE/UIQM 0.86/0.91 0.88/0.87 0.76/0.93

UCIQUE/UIQM 0.80/0.92 0.73/0.87 0.60/0.76

UCIQUE/UIQM 0.75/1.03 0.63/0.94 0.51/0.87

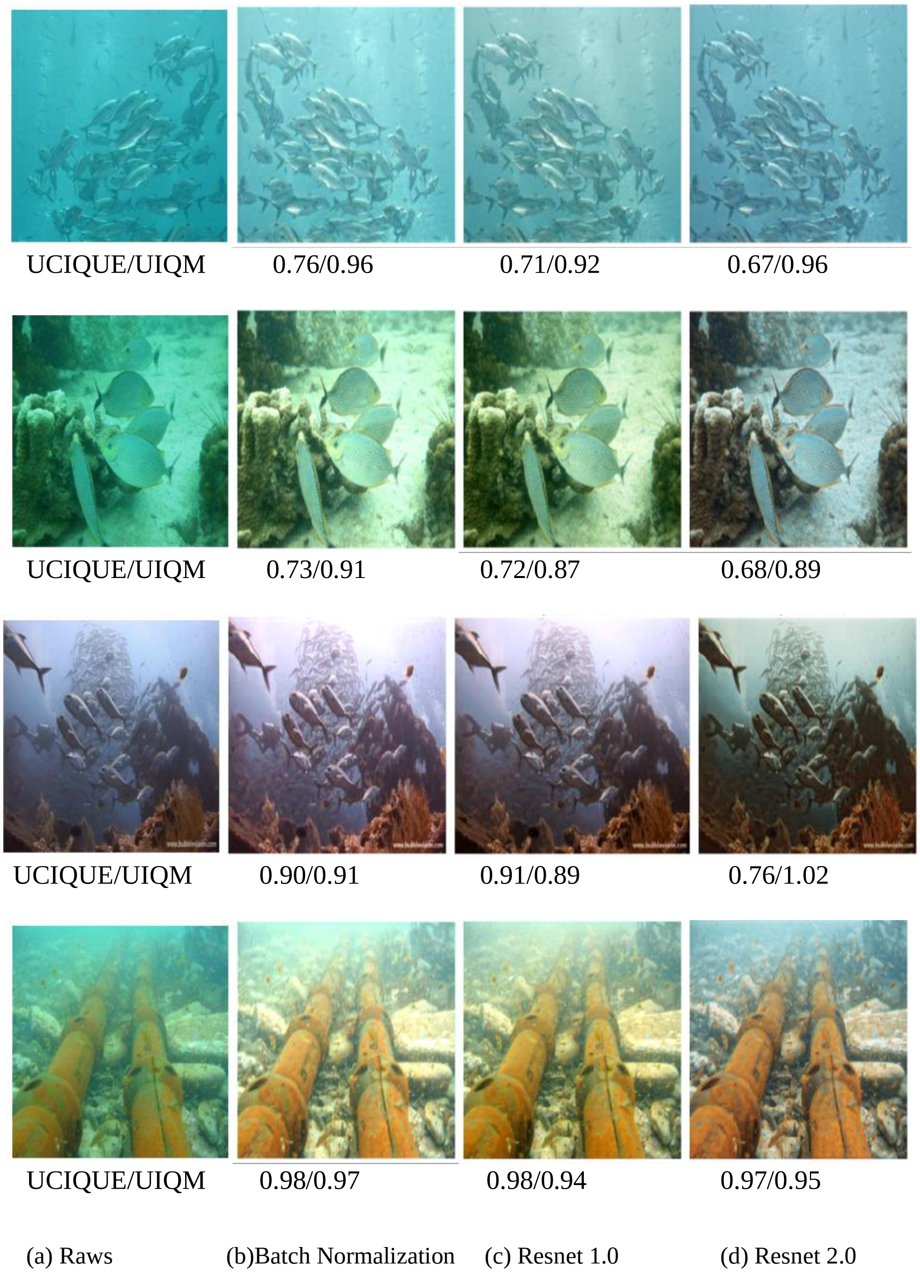


Fig. 5 Visual comparisons in terms of UCIQE and UIQM metrics for all the 3 architectures.

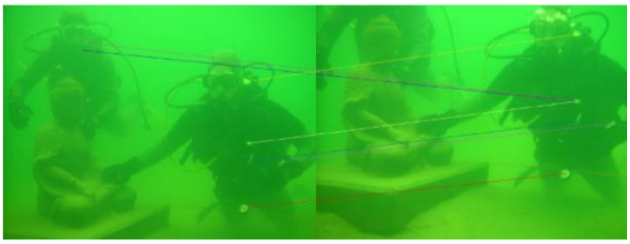
ARCHITECTURE	AVG (UCIQE)	AVG (UIQM)
BATCH NORMALIZATION	0.832857	0.865476
RESNET 1.0	0.878810	0.804286
RESNET 2.0	0.746905	0.907619

Table 1. Visual comparisons in terms of UCIQE and UIQM metrics for all the 3 architectures.

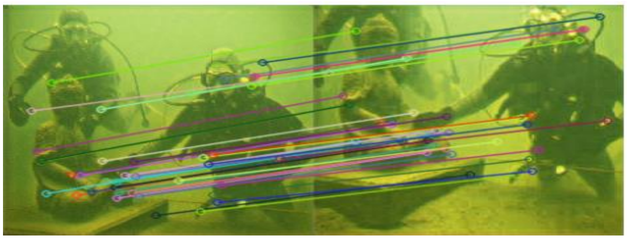
The choice between these methods depends on the specific priorities of application. If colour accuracy is crucial, ResNet 2.0 might be a favorable choice. If a balance between colour quality and overall image quality is preferred, ResNet 1.0 seems to be a reliable option. On the other hand, if superior overall image quality is the primary concern, Batch Normalization stands out.

B. Scale-Invariant Feature Transform Application

It is a powerful computer vision algorithm designed to identify and describe distinctive features within images, irrespective of their scale, rotation, or illumination changes. It excels at detecting key points, which are unique and invariant points in an image, and generating descriptors that encapsulate information about the local image structure. We used SIFT to compare between pairs of images, once non enhanced and then enhanced, and identify corresponding key points between them. The algorithm calculates descriptors for these key points and matches them between the images, emphasizing those pairs with a significant similarity. The final result is a visualization of these matches, where lines connect the corresponding key points, providing insights into the similarities and alignments between the two images. It also describes how good our enhanced images are. This SIFT-based matching approach is particularly useful in scenarios where images may undergo variations due to factors such as underwater conditions. By focusing on distinctive features and their descriptors, SIFT enables robust comparisons and matching, contributing to tasks like image alignment and localized feature preservation. [7]



[a]



[b]

Fig. 6 SIFT (a) Sift on Normal Images (b) Sift on Enhanced Images

We can notice 12 connections between enhanced images, compared to 6 between normal images which proves our enhancement works better to identify similar objects.

V. CONCLUSION

Initially, we explored the limitations of traditional image enhancement methods, exemplified by Principal Component Analysis (PCA), which struggled to effectively address the complexities introduced by underwater images. In response to these challenges, we innovatively employed three distinct architectures of autoencoders for underwater image enhancement. This approach showcased the versatility of neural network-based methods in learning and reproducing intricate patterns within underwater images, leading to significant improvements in visual quality.

Moreover, our construction of an extensive underwater image enhancement benchmark dataset, featuring large-scale real underwater images alongside corresponding reference images, provides a robust foundation for in-depth analysis. Through qualitative and quantitative evaluations, we observed that no single method consistently outperforms others across a range of full- and no-reference metrics. This underscores the complexity of underwater image enhancement and the need for adaptive, multifaceted approaches. Looking ahead, our project's future trajectory involves extending our investigation into object detection and tracking, comparing performance on both normal and enhanced images. This next phase aims to provide a holistic understanding of the impact of image enhancement methods on downstream computer vision tasks in the underwater domain. [8]

VI. REFERENCES

[1] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer and P. Bekaert, "Colour Balance and Fusion for Underwater Image Enhancement," in IEEE Transactions on Image Processing, vol. 27, no. 1, pp. 379-393, Jan. 2018, doi: 10.1109/TIP.2017.2759252.

[2] Rafael C. Gonzalez and Richard E. Woods , "Digital Image Processing ", pp. 842-852, 2008.

[3] J. Zhai, S. Zhang, J. Chen and Q. He, "Autoencoder and Its Various Variants," 2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Miyazaki, Japan, 2018, pp. 415-419, doi: 10.1109/SMC.2018.00080.

[4] P. Liu, G. Wang, H. Qi, C. Zhang, H. Zheng and Z. Yu, "Underwater Image Enhancement With a Deep Residual Framework," in IEEE Access, vol. 7, pp. 94614-94629, 2019, doi: 10.1109/ACCESS.2019.2928976.

[5] A. Arezoomand, P. Cheraaqee and A. Mansouri, "Perceptually Optimized Loss Function for Image Super-Resolution," 2021 7th International Conference on Signal Processing and Intelligent Systems (ICSPIS), Tehran, Iran, Islamic Republic of, 2021, pp. 01-05, doi: 10.1109/ICSPIS54653.2021.9729334.

[6] Q. Chen, Z. Zhang and G. Li, "Underwater Image Enhancement Based on Colour Balance and Multi-Scale Fusion," in IEEE Photonics Journal, vol. 14, no. 6, pp. 1-10, Dec. 2022, Art no. 3963010, doi: 10.1109/JPHOT.2022.3227159.

[7] F. Guo, J. Yang, Y. Chen and B. Yao, "Research on image detection and matching based on SIFT features," 2018 3rd International Conference on Control and Robotics Engineering (ICCRE), Nagoya, Japan, 2018, pp. 130-134, doi: 10.1109/ICCRE.2018.8376448.
2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-6, doi: 10.1109/MysuruCon55714.2022.9972512.

[8]H. Balta, C. O. Ancuti and C. Ancuti, "Effective underwater image restoration based on colour channel compensation," OCEANS 2021: San Diego – Porto, San Diego,CA,USA,2021,pp.1-4,doi: 10.23919/OCEANS44145.2021.9705902.

[8] K. Raza, S. K. Selvaperumal, C. Natarj, R. Lakshmanan and A. Seeralan, "Enhanced Object Detection and Tracking Algorithm in Complex Videos," 2018 Fourth International Conference on Advances in Electrical, Electronics, Information, Communication and Bio-Informatics (AEEICB).