Optimizing Underwater Image Quality with Deep Reinforcement Learning

Marc Anthony B Reyes
MS Computer Science
University of the Philippines
mbreyes12@up.edu.ph

ABSTRACT

Underwater imaging faces challenges such as poor visibility, color distortion, and low contrast due to light scattering and absorption. This research presents a Deep Q-Network (DQN)-based model for enhancing underwater image quality by addressing brightness, contrast, and color saturation. Using a Markov Decision Process (MDP) framework, the agent interacts with the environment to improve image quality. The model is trained on the large-scale underwater image (LSUI) dataset and the Underwater Image Enhancement Benchmark (UIEB), applying techniques like brightness adjustment, contrast modification, and deblurring. The reward function combines pixelwise and perceptual losses to ensure numerical accuracy and visual appeal. Experimental results show significant improvements in Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) metrics, with the GPU implementation offering substantial speed advantages.

1. INTRODUCTION

1.1 Background

Underwater image enhancement has become an increasingly important area of research due to its critical applications in marine exploration, underwater robotics, and aquatic environmental monitoring. The degradation of underwater images is primarily caused by the absorption and scattering of light, which leads to issues such as low contrast, color distortion, and haziness [2]. These challenges hinder the effective analysis and interpretation of underwater imagery, which is vital for various scientific and industrial applications.

The physical properties of water significantly affect light propagation. As light travels through water, it is absorbed and scattered by water molecules and particles.

1.2 Literature Review

A comprehensive review of the literature reveals several important studies focused on underwater image enhancement. For instance, [8] demonstrated that convolutional neural networks (CNNs) could effectively enhance underwater images by learning complex mappings between degraded and clear images. However, these models often require large amounts of labeled data and may not generalize well to different underwater conditions.

One research utilized a U-shape Transformer network to tackle the challenges of underwater image enhancement [5]. The U-shape Transformer integrates a channel-wise multi-scale feature fusion transformer (CMSFFT) module and a spatial-wise global feature modeling transformer (SGFMT) module, specifically designed to address inconsistent attenuation in different color channels and spatial areas. This architecture enhances the network's attention to crucial image regions, significantly improving visual quality.

Reinforcement learning, particularly Deep Q-Networks (DQN) [7], has shown promise in various image processing tasks. [3] highlighted the potential of reinforcement learning in image processing, showing that adaptive models could improve image quality by optimizing a reward function designed for visual enhancement. Despite these advancements, traditional models often struggle with overfitting and lack robustness when applied to new environments [1].

Furthermore, image quality metrics such as Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) are commonly used to evaluate the performance of image enhancement techniques. These metrics can be defined as follows:

$$\text{MSE} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i,j) - K(i,j)]^2, \tag{1}$$

where I and K are the original and enhanced images, respectively, and m and n are the dimensions of the images.

$$\text{PSNR} = 10 \log_{10} \left(\frac{MAX_I^2}{\text{MSE}} \right), \tag{2}$$

where MAX_I is the maximum possible pixel value of the image.

$$\text{SSIM}(x,y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \tag{3}$$

where μ_x and μ_y are the mean values, σ_x and σ_y are the variances, and σ_{xy} is the covariance of x and y. Constants C_1 and C_2 are used to stabilize the division.

1.3 Image Improvement Techniques

The problem of underwater image enhancement concerns the improvement of underwater images and is attracting a lot of attention due to its importance in naval engineering and aquatic robotics. It is a challenging task due to problems caused by the environment in which such photos are taken.

The main problems are caused by two phenomena: scattering, which is the presence of particles such as dust and plankton, and absorption, the absorption due to the wavelength that alters the color of the image making it darker and changing its dominant color.

Existing approaches used to address the problem of underwater image enhancement present some limitations:

- Given the complexity of the underwater environment, a single image processing method is not able to adapt to different scenes (such as different turbidities, different lighting conditions, and shots from different angles), causing inconsistency in performance (imprecise or incorrect improvements).
- Models based on Deep Learning methods, which are usually structured with an end-to-end network, apply processing that is effectively a black box, reducing the model's explainability and preventing insights for improvement.
- Training a deep network requires a large number of training data pairs, consisting of a raw image and a reference image, considered the ground truth and obtained through selection made by expert users, since it is not possible to have a true ground truth image. The reference image for such models acts as an upper limit for the quality of the image obtained from the raw image, preventing the generation of better images.

1.4 Research Problem

Given the current state of research, there is a clear need to address the problem of generalizing underwater image enhancement models to diverse conditions. This study aims to fill this gap by developing a Deep Q-Network (DQN)-based reinforcement learning model [7] that iteratively improves the quality of underwater images by maximizing a reward function tailored for visual enhancement.

1.5 Objectives

The primary objectives of this research are:

- To investigate the application of DQN [7] for underwater image enhancement.
- To analyze the effectiveness of the DQN [7] model in improving image clarity, reducing color distortion, and enhancing visibility.
- To evaluate the performance of the proposed model against traditional image enhancement techniques.

1.6 General Framework

The model solves a Markov Decision Process (MDP) [6] that consists of the interaction at each step t between an agent and an environment. The MDP [6] is composed of four components:

State $s_t \in S$. This represents the set of information that the agent perceives from the environment. In this context, the state is composed of features extracted from the underwater images. These features include both color features (such as histograms in the CIELAB color space) and perceptual features (extracted using a pre-trained VGG-19 network). The state provides a comprehensive representation of the current condition of the image, enabling the agent to make informed decisions.

Action $a_t \in A$. Actions are the behaviors that the agent can perform in the environment. For image enhancement, actions include operations such as adjusting brightness, contrast, and color saturation; applying gamma correction; performing histogram equalization; and executing deblurring techniques. Each action is designed to incrementally improve the quality of the image.

Reward $r_t \in R$. The reward is the feedback the agent receives from the environment after performing an action. In this case, the reward measures how much the action improves the image quality, making it closer to a ground truth reference image. The reward function combines pixelwise loss (mean absolute error) and perceptual loss (mean squared error of VGG-19 feature differences), providing a balanced evaluation of both numerical and perceptual improvements.

Policy Ω_θ. The policy is the strategy that the agent applies to decide its next action based on the current state. In this model, an ϵ-greedy policy is used, where the agent selects the best-known action most of the time but occasionally explores random actions to discover potentially better strategies. This balance between exploitation and exploration helps the agent learn an optimal policy over time.

1.7 Contribution

This research makes several contributions to the field of underwater image processing. First, it introduces a novel approach to underwater image enhancement using deep reinforcement learning. Second, it provides new insights into the applicability of reinforcement learning for adaptive image enhancement tasks. Finally, the findings have practical implications for improving the quality of underwater imagery used in various applications, including marine biology, underwater robotics, and environmental monitoring.

Figure 1: Gamma Correction

2. METHODOLOGY

2.1 Data Collection

Figure 2: HE and CLAHE Correction

Figure 3: White Balance Correction

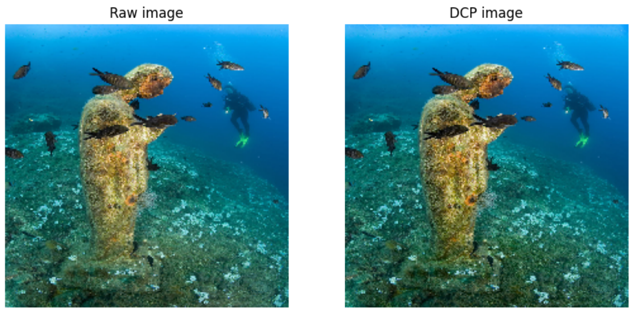

Figure 4: DCP Correction

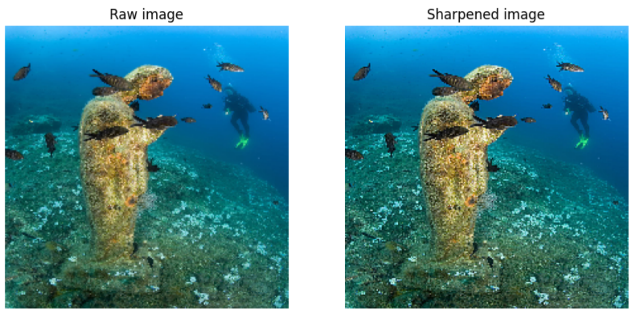

Figure 5: Sharpen Correction

The data used in this study were collected from two primary sources: the Large-scale Underwater Image (LSUI) dataset and the Underwater Image Enhancement Benchmark (UIEB) [4]. The LSUI dataset comprises 5,004 image pairs featuring diverse underwater scenes with various lighting conditions, water types, and target categories, offering high-quality reference images. The UIEB dataset [4] contains 890 raw images with corresponding high-quality ref-

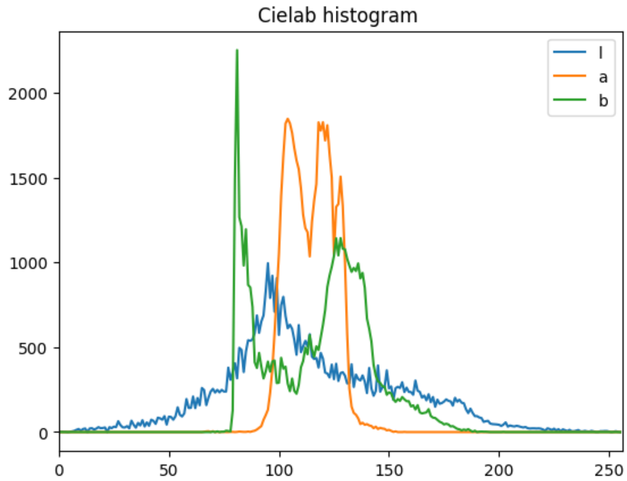

Figure 6: CIELAB Histogram

erence images and an additional 60 images without corresponding references, defined as challenge images. Prior to analysis, several preprocessing steps were undertaken to ensure data quality, including data cleaning to remove noise, normalization to standardize pixel values, and augmentation techniques such as rotation, flipping, and cropping to increase the diversity of the training set.

2.2 Experimental Setup

The experiments were conducted on a high-performance computing setup equipped with an NVIDIA RTX A4000 GPU. The computational resources were crucial for handling the large-scale data and complex models involved in this research.

2.3 Algorithms and Models

The primary algorithm employed in this research is the Deep Q-Network (DQN) [7] algorithm. This choice was motivated by its success in reinforcement learning tasks, particularly in scenarios requiring adaptive decision-making. The model architecture consists of several convolutional layers followed by fully connected layers. The network was trained to optimize a reward function designed to maximize visual quality metrics.

2.4 Training and Evaluation

The model was trained using a reinforcement learning approach with a reward function that combines color and perceptual rewards. The training process involved 20,000 episodes with a batch size of 32. The reward function R is defined as:

$$R = \alpha R_c + \beta R_p, \tag{4}$$

where R_c is the color reward and R_p is the perceptual reward, with balancing coefficients $\alpha = 1$ and $\beta = 0.05$. The color reward R_c is given by:

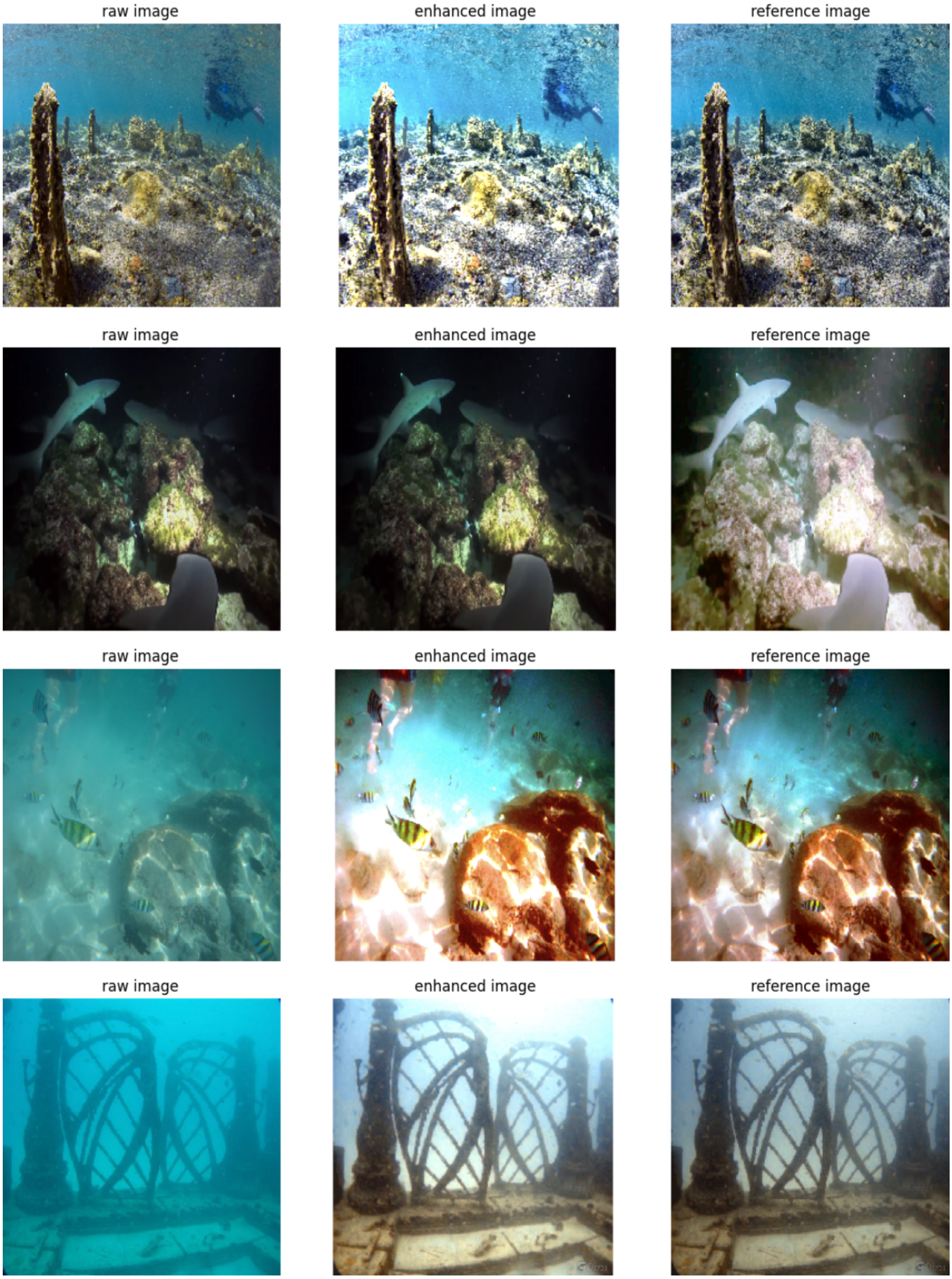

Figure 7: Enhanced images with respect to the raw images and reference images.

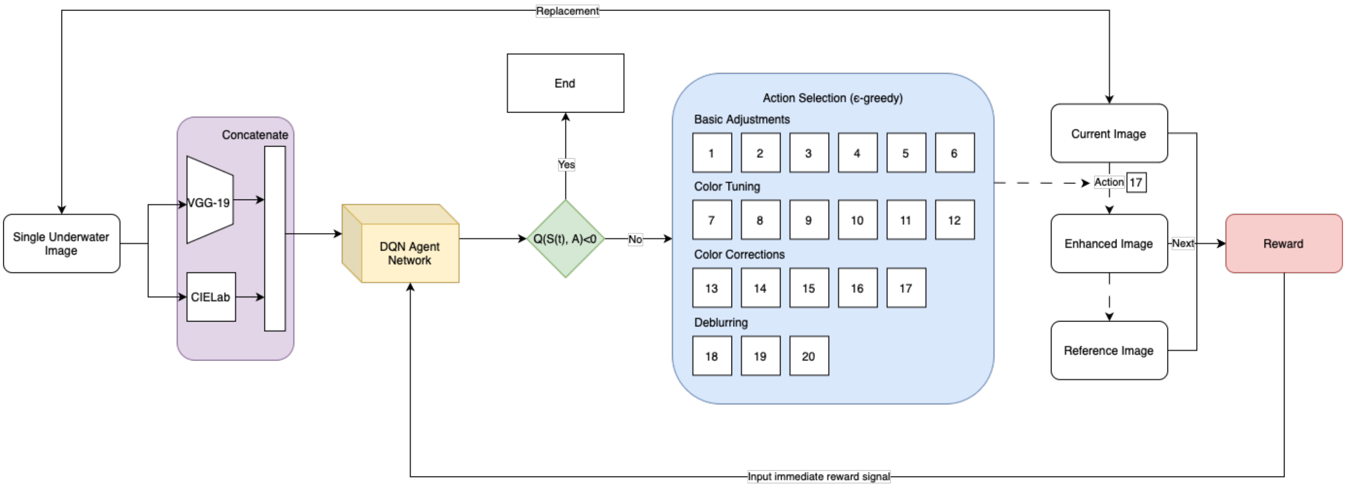

Figure 8: General Framework: 1) Feature extraction. Extracts features from an image. The extracted features will constitute the states. 2) Deep Q-Network. Takes a state as input and returns a $Q(s, a)$ value for each action. 3) Action selection. Chooses an action using an ϵ-greedy policy. 4) Reward calculation

$$R_c = -\sum_i \left| C_i(G(X(t))) - C_i(Y) \right|, \tag{5}$$

where C_i extracts color features, $G(X(t))$ is the improved image, and Y is the reference image. The perceptual reward R_p is calculated as:

$$R_p = -[L_p(t) - L_p(t-1)], \tag{6}$$

where L_p is the perceptual loss defined by the difference in VGG-19 extracted features between the improved and reference images. Evaluation metrics include Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM):

$$\text{MSE} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i,j) - K(i,j)]^2, \tag{7}$$

$$\text{PSNR} = 10 \log_{10} \left(\frac{MAX_I^2}{\text{MSE}} \right), \tag{8}$$

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{9}$$

The dataset was divided into training and validation sets with an 80/20 split.

2.5 Implementation Details

The implementation was carried out using TensorFlow and Keras for model development, with additional support from NumPy for data manipulation. The code was structured to ensure reproducibility and ease of experimentation. Key implementation details included setting up a robust data pipeline for efficient data loading and augmentation, hyperparameter tuning using grid search, and leveraging TensorFlow's callback functions for monitoring training progress. The training parameters were set as follows:

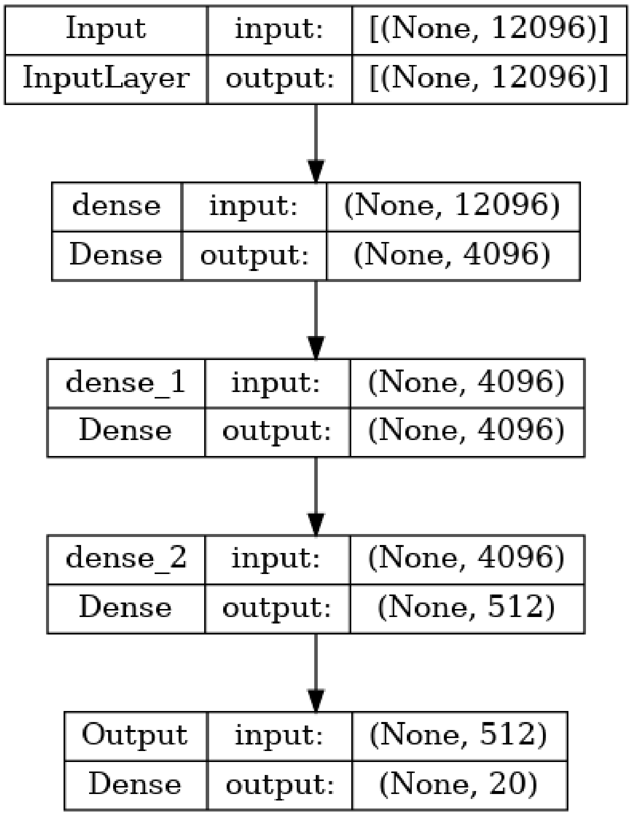

Figure 9: DQN Model Architecture

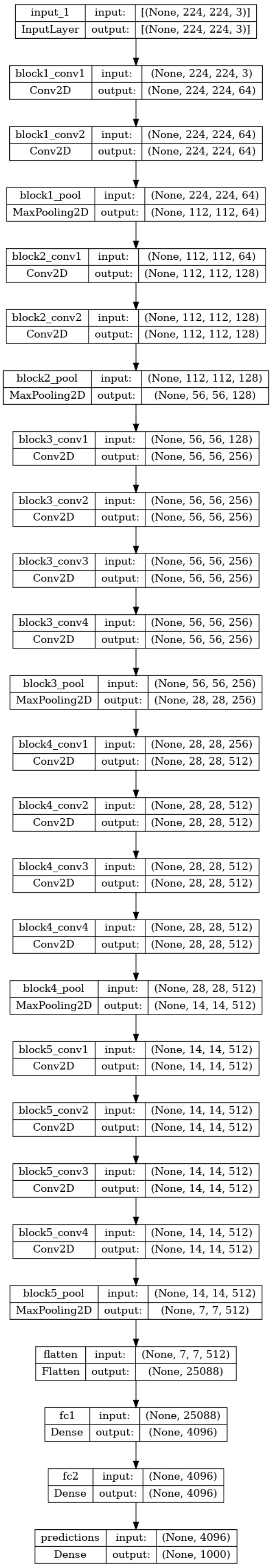

Figure 10: VGG Model Architecture

Parameter	Value
learning rate	1e-5
minimum learning rate	1e-8
learning decay step	0.96
discount rate	0.99
batch size	32
initial ϵ	1
minimum ϵ	0.1
epsilon decay step	0.99
episodes	20,000
max steps per episode	50
α	1
β	0.05

Table 1: Training parameters for the Deep Q-Network (DQN) model

2.6 Model Components and Process

The methodology involves several key steps and components:

Feature Extraction. Features are extracted from images to constitute the states. Color features are extracted using CIELAB histograms, and perceptual features are extracted using a pre-trained VGG-19 network. The feature extraction process involves combining Color Features and Perceptual Features to form a comprehensive representation of underwater images. Color features describe the global properties of an image's surface and are extracted using a CIELAB histogram. This involves utilizing the L* (lightness), a* (green to red), and b* (blue to yellow) axes of the CIELAB color space, dividing each axis into 20 intervals, and counting the number of pixels within each interval, resulting in a $20 \times 20 \times 20$ histogram. This process yields 8000 bins, forming the Color Features vector. On the other hand, perceptual features are context-dependent and inspired by human vision, extracted using a pre-trained VGG-NET 19 model. This pre-trained model, trained on the ImageNet dataset, extracts features from one of the fully connected layers, producing a 4096-dimensional vector. These features capture high-level abstractions and patterns within the image.

The extracted Color Features and Perceptual Features are then concatenated to form a single, comprehensive feature vector. The combined vector consists of 12,096 features, providing a rich representation of both the color distribution and perceptual characteristics of the underwater images.

Action Selection. Actions include image enhancement techniques such as brightness and contrast adjustments, color corrections, histogram equalization, and deblurring techniques.

Reward Calculation. The reward is calculated based on pixelwise loss (Mean Absolute Error, MAE) and perceptual loss (Mean Squared Error, MSE), combined using the formula $R = \alpha R_{pixelwise} + \beta R_{perceptual}$.

Training Process. The agent is trained by iteratively applying actions to enhance images and learning from the resulting rewards using an ϵ-greedy policy.

Category	Action	Description
Basic Adjustments	Brightness +5%	Increase the brightness of the image by 5%.
	Brightness -5%	Decrease the brightness of the image by 5%.
	Contrast +5%	Increase the contrast of the image by 5%.
	Contrast -5%	Decrease the contrast of the image by 5%.
	Color Saturation +5%	Increase the color saturation of the image by 5%.
	Color Saturation -5%	Decrease the color saturation of the image by 5%.
Color Adjustments	Red +5%	Increase the red channel intensity by 5%.
	Red -5%	Decrease the red channel intensity by 5%.
	Green +5%	Increase the green channel intensity by 5%.
	Green -5%	Decrease the green channel intensity by 5%.
	Blue +5%	Increase the blue channel intensity by 5%.
	Blue -5%	Decrease the blue channel intensity by 5%.
Corrections	Gamma +	Apply gamma correction to increase the brightness of darker regions.
	Gamma -	Apply gamma correction to decrease the brightness of brighter regions.
	HE	Enhance image contrast by redistributing intensity values.
	CLAHE	Enhance image contrast while limiting noise amplification.
	White Balance	Adjust the colors to make the image appear more natural.
Deblurring	Sharpen	Apply a sharpen filter to enhance the edges and details.
	Emboss	Apply an emboss filter to create a 3D relief effect.
	Dark Channel Prior (DCP)	Reduce haze and improve image clarity using the dark channel prior method.

Table 2: List of actions used for image enhancement, categorized by type, with descriptions.

Validation: Performance is validated on a separate dataset section using metrics such as MSE, PSNR, and SSIM.

2.7 Actions

The proposed methodology includes various image enhancement techniques categorized into four main types. Each type involves specific actions that the agent can take to improve the image quality. Table 2 provides a summary of these actions along with their descriptions.

3. RESULTS AND DISCUSSION

3.1 Results

The results of this study are presented in the form of tables and figures that summarize the key findings. Figure 11 shows the reward trend during the training process. An increasing trend indicates that the agent is learning to effectively improve raw images over time. Table 3 provides a detailed summary of the performance metrics (MSE, PSNR, and SSIM) for the proposed model.

Metric	Value
Average MSE	1858.0198193370575
Average PSNR	16.992879298748356
Average SSIM	0.7424274301354952

Table 3: Summary of performance metrics for the proposed DQN model.

The findings of this study have several important implications. Firstly, the improved metrics of the proposed DQN [7] model demonstrate its potential for real-world applications in underwater image enhancement. The use of reinforcement learning to iteratively enhance images provides a new approach to tackling the challenges posed by underwater imaging conditions. The improved MSE, PSNR, and SSIM values indicate that the proposed model effectively enhances image quality, making it a viable solution for applications such as marine biology, underwater archaeology, and pipeline inspection.

Episode	Steps	Episode Reward
1	50	-2.03209
1000	50000	-2.13461
2000	100000	0.278477
3000	150000	-2.60664
4000	200000	-0.675033
10000	500000	-0.785109
11000	550000	0.523002
12000	585665	-0.098416
13000	586667	-1.68988
14000	587667	0.158719
16000	589671	-0.0532499
17000	590678	-0.0574592
18000	591682	-0.717265
19000	592682	-0.461128
20000	593684	0.104513

Table 4: Summary of episode steps and corresponding rewards during the training process.

The significance of these improvements is highlighted by the specific challenges inherent in underwater imaging. For instance, light scattering and absorption in water lead to significant color distortion and loss of detail, issues that traditional image enhancement techniques often struggle to ad-

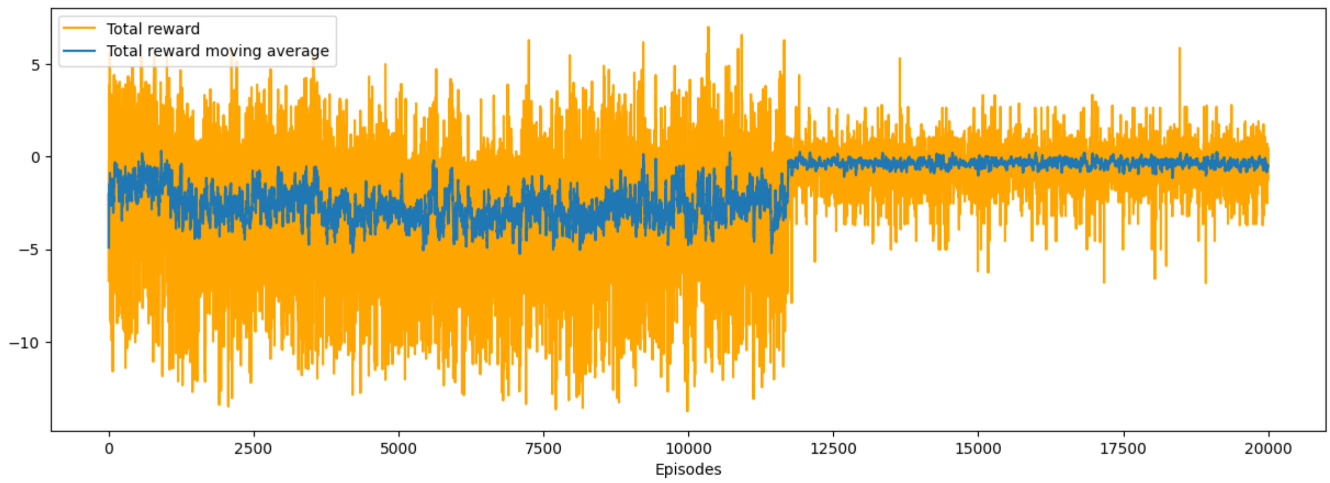

Figure 11: The reward trend during training with the above-listed hyperparameters. An increasing trend indicates that the agent is learning to effectively improve raw images.

dress comprehensively. By using reinforcement learning, the model adapts to these challenges dynamically, learning to apply the most effective enhancements for varying conditions.

Moreover, the integration of both pixelwise and perceptual rewards ensures that the enhancements not only improve numerical metrics but also lead to visually pleasing results. The perceptual reward, derived from features extracted by a pre-trained VGG-19 network, aligns the enhancements with human visual perception, addressing common issues such as over-saturation or unnatural color balances that purely pixelwise methods might introduce.

4. CONCLUSION

4.1 Limitations

Despite the promising results, this study has certain limitations. One of the primary limitations is the reliance on specific datasets (LSUI and UIEB [4]), which may not generalize well to other underwater environments. While these datasets provide a diverse range of conditions, they may not capture the full spectrum of challenges encountered in different underwater settings.

Additionally, the computational constraints limited the size of the model and the number of experiments conducted. Training deep reinforcement learning models is resource-intensive, and more extensive experimentation could explore a wider range of model architectures and hyperparameters. The model's performance may also be influenced by the specific choice of hyperparameters and the structure of the reward function, which could be further optimized.

Another limitation is the evaluation framework. While MSE, PSNR, and SSIM are widely used metrics, they may not fully capture the perceptual quality improvements perceived by human observers. Future work could incorporate subjective evaluations by human participants to complement the quantitative metrics.

4.2 Future Work

Future research should address the limitations identified in this study. One potential direction is exploring the use of more diverse datasets to improve the generalizability of the model. This could involve collecting underwater images from different geographical locations, depths, and environmental conditions to ensure the model's robustness across a wider range of scenarios.

Additionally, further optimization of the model architecture and training process could yield even better performance. Investigating the impact of different hyperparameters on model performance, such as the learning rate, discount factor, and reward coefficients, could provide insights into the optimal configurations for underwater image enhancement.

Finally, expanding the model's testing framework to include various image types beyond underwater images could evaluate its generalization capabilities and potential for broader application. For instance, applying the model to enhance images in other challenging environments, such as hazy or low-light conditions, could demonstrate its versatility.

5. REFERENCES

[1] Chen, C., Ying, V., and Laird, D. Deep q-learning with recurrent neural networks. *stanford cs229 course report 4* (2016), 3.

[2] Han, M., Lyu, Z., Qiu, T., and Xu, M. A review on intelligence dehazing and color restoration for underwater images. *IEEE Transactions on Systems, Man, and Cybernetics: Systems 50*, 5 (2020), 1820–1832.

[3] Le, N., Rathour, V. S., Yamazaki, K., Luu, K., and Savvides, M. Deep reinforcement learning in computer vision: a comprehensive survey. *Artificial Intelligence Review 55*, 4 (Apr 2022), 2733–2819.

[4] Li, C., Guo, C., Ren, W., Cong, R., Hou, J., Kwong, S., and Tao, D. An underwater image

enhancement benchmark dataset and beyond. *IEEE transactions on image processing 29* (2019), 4376–4389.

[5] Peng, L., Zhu, C., and Bian, L. U-shape transformer for underwater image enhancement. *IEEE Transactions on Image Processing 32* (2023), 3066–3079.

[6] Puterman, M. L. Markov decision processes. *Handbooks in operations research and management science 2* (1990), 331–434.

[7] Roderick, M., MacGlashan, J., and Tellex, S. Implementing the deep q-network. *arXiv preprint arXiv:1711.07478* (2017).

[8] Xu, Y., Zhang, Y., Wang, H., and Liu, X. Underwater image classification using deep convolutional neural networks and data augmentation. In *2017 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)* (2017), pp. 1–5.