

“Deep Learning Based Multi- Scenario Image Enhancement for Low-Light, Hazy & Under Water Environment”

Ranvir Kumar¹, Prof. Sarwesh Site²,

¹M.Tech Student , Department of Computer Science and Engineering, All Saint's College of Technology, Bhopal, India

Affiliated to Rajiv Gandhi Proudhyogiki Vishwavidyalaya (RGPV) ranvirdevops@gmail.com

² Associate Professor, Department of Computer Science and Engineering, All Saints College of Technology, Bhopal, India

Affiliated to Rajiv Gandhi Proudhyogiki Vishwavidyalaya (RGPV) er.sarwesh@gmail.com

Abstract –

Image enhancement is a branch of computer vision that focuses on improving the interpretation, analysis, and usability of visual data captured under adverse conditions. In this subject, Deep Learning techniques are widely used, and most state-of-the-art research focuses on domain-specific enhancements, such as low-light, underwater, or cloudy circumstances. However, these studies restrict their usability to specific domains, which limits their flexibility and scalability. Image enhancement plays a critical role in modern computer vision, enabling better interpretation, analysis and usability of visual data captured under adverse conditions. While deep learning has significantly advanced image restoration tasks, most state-of-the-art models are tailored to specific domains such as low-light, underwater or hazy environments, limiting their flexibility and scalability. An integrated framework for image enhancement based on an attention-guided U-Net architecture for multi-domain enhancement is presented in this thesis. The recommended approach incorporates attention mechanisms to selectively emphasize significant areas in the input image and makes use of U-Net's structural advantages for multiscale feature learning. A carefully curated tri domain paired dataset was constructed using three benchmark datasets: LoLI-Street for low-light images, EUVP for underwater scenes and RESIDE for hazy conditions. Supervised learning with Mean Squared Error (MSE) loss is used to train the model with the goal of restoring both visual quality and structural integrity. To assess the model, a number of tests are conducted throughout all three domains using standard metrics like Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), along with domain-specific measures such as UIQM (underwater), EPI (low-light), and FADE (hazy). The model achieved strong performance, with PSNR values up to 36.36 dB and SSIM scores exceeding 0.97 in the low-light domain while also demonstrating effective generalization in underwater and hazy conditions. Qualitative results further validated the framework to enhance images perceptually, restoring brightness, contrast, clarity and detail. In conclusion, the proposed attention-based U-Net offers a flexible and effective solution for cross domain image enhancement. By unifying multiple enhancement tasks into a single network,

this work establishes the basis for more scalable and generalized image restoration algorithms that can be used in practical applications like aerial imaging, autonomous driving, surveillance, and underwater research.

Keywords

Image Enhancement, Multi-Domain Image Enhancement, Image Restoration, Degraded Image Processing, Deep Learning, U-Net Architecture, Transformer-based Models, Attention Mechanism, PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index), Image Quality Assessment

1. INTRODUCTION

1.1 Background and Motivations

In contemporary research, the demand for high quality visual data has grown significantly across various domains such as autonomous vehicles, surveillance, underwater exploration and environmental monitoring. However, in many real world scenarios, images captured under challenging conditions such as poor lighting, hazy weather or underwater environments often suffer from minimal contrast, poor visibility and warping of color [1]. These distorted images affect not just visual perception but also the success of computer vision tasks such as detection, recognition, and segmentation. Histogram equalization and filtering-based approaches are examples of standard image enhancement techniques that have been widely utilized to improve image quality[1]. However, their performance often remains limited to specific domains and fails to generalize across varying degradation types [2]. With the advancement of deep learning, models like Convolutional Neural Networks (CNNs) and U-Nets have shown remarkable success in learning complex patterns from data, enabling more effective enhancement. Still, most existing solutions are domain specific, trained separately for either low-light, underwater or hazy images, which restricts their adaptability and real world utility. The growing need for an unified solution prompted the creation of a single framework that could manage several types of degradations using a shared architecture. The attention-augmented U-Net architecture used in this work allows the model to focus on

the most instructive spatial characteristics in an image. By designing a unified model that can learn across three distinct degradation domains, low-light, underwater and hazy, we aim to offer a generalized solution for image enhancement that is both efficient and effective

1.1.1 Image Enhancement

Image enhancement refers to the process of improving the visual appearance or the interpretability of an image. It aims to emphasize meaningful features and suppress undesired distortions without altering the underlying content [3]. This process is especially useful in applications where images are affected by poor lighting, environmental noise or weather conditions. Enhancement can be subjective (for human viewing) or objective (to support computer vision systems in tasks like detection or segmentation). Traditional image enhancement methods fall into two major categories: spatial domain and frequency domain techniques [3].

Spatial methods operate directly on pixel 2 1.1 Background and Motivations values and include approaches such as histogram equalization, contrast stretching and unsharp masking. Frequency domain methods on the other hand, modify image components after transforming them using techniques like the Fourier Transform, often targeting specific frequency bands for enhancement. Deep learning-based techniques have consequently gained popularity. By explicitly identifying patterns in data, these techniques can more effectively adjust to a range of challenging circumstances.

1.1.2 Challenges

In real world conditions, images are often captured under sub-optimal environments that introduce domain specific degradations. Each type of degradation presents unique challenges to both human observers and automated systems and understanding them is essential for designing effective enhancement models.

Low-light Low brightness, low contrast, and high noise levels are common problems with photos taken in low light. Common scenarios include night time surveillance, low-light photography and indoor imaging with insufficient illumination. Such images frequently contain large shadow regions, blurred details and noise that obscures important features. Traditional brightness adjustment or contrast enhancement techniques are generally inadequate, as they tend to amplify noise along with useful content [4].

Underwater Numerous physical constraints resulting from the way light interacts with the aquatic environment impact underwater imaging. As light travels through water, it gets absorbed and scattered, leading to a strong color cast usually blue or green along with reduced contrast and visibility [5]. Particles suspended in water further scatter the light, resulting in hazy appearances and blurring.

Hazy images typically arise in outdoor environments where fog, mist or pollution scatters the light between objects and

the camera. This scattering reduces scene contrast and color vibrancy making the image appear washed out and unclear. A key characteristic of haze is the presence of depth dependent degradation: objects farther from the camera appear more blurred or faded than those nearby [6].

1.2 PROBLEM STATEMENT

Despite significant progress in image enhancement, most existing methods are designed to address a specific type of degradation. Usually, these models are developed and refined for a specific domain, limiting their effectiveness when deployed in real world scenarios where degradations are not only diverse but sometimes occur simultaneously. As a result, multiple domain specific models must be maintained, increasing computational cost, storage requirements and development effort. Moreover, domain specific models lack the adaptability required for general use. A model trained to enhance low-light images may fail when applied to underwater scenes or hazy environments, as each domain presents unique challenges in terms of color distortion, visibility and noisy characteristics. This lack of generalization restricts their usability in systems that require robust performance across varying environmental conditions. Additionally, handcrafted enhancement techniques often rely on fixed rules or assumptions, which may not hold across different contexts. While Deep Learning techniques have demonstrated encouraging outcomes in domain-specific tasks, few studies make an effort to integrate several domains into a single architecture. Therefore, there is a pressing need for a unified framework capable of handling multiple degradation types with minimal compromise in performance. The problem lies not only in enhancing each type of image effectively but also in designing a scalable, efficient model that can generalize well across low-light, underwater and hazy images while preserving important features, restoring natural appearance and avoiding over enhancement or artifacts.

1.3 OBJECTIVES

This dissertation's main goal is to propose a single, deep learning-based framework for enhancing images captured under three challenging visual conditions, low-light, underwater and hazy environments using an Attention-based U-Net architecture. By creating a more generalizable strategy that can successfully manage many kinds of image deterioration, this study aims to overcome the limitations of domain-specific enhancement techniques. The key objectives of this research are clarified as follows: 1. To analyze the unique degradation characteristics associated with low-light, underwater, hazy images and understand the challenges they pose to traditional and deep learning-based enhancement techniques. 2. To design a unified image enhancement framework that effectively restores image quality across all three domains, low-light, underwater and hazy using a shared model. 3. To establish a U-Net-based architecture that performs pixel-wise improvement while preserving spatial and

contextual information using an encoder-decoder structure with skip connections and attention gates.

2 LITERATURE REVIEW

2.1 RELATED TOOLS AND TECHNOLOGIES

Image enhancement has changed drastically since deep learning became standard, offering improved performance over traditional methods in tasks such as contrast correction, dehazing, denoising and color restoration [7]. Among the architectures explored in literature, Convolutional Neural Networks (CNNs) have formed the backbone of many enhancement models due to their ability to learn spatial hierarchies and patterns directly from pixel data. A widely recognized architecture in this domain is the U-Net, originally developed for biomedical image segmentation and later extended to applications in image restoration and enhancement[8]. Its decoder-encoding architecture and skip connections enable both high-level contextual consciousness and low-level detail preservation. Recent developments have included attention mechanisms into U-Net to improve the model's ability to focus on visually significant areas[9]. This attention-based modification forms the foundation of the framework proposed in this thesis. For image enhancement, additional designs including Transformers and Generative Adversarial Networks (GANs) have been proposed in addition to U-Net variations. GANs are particularly known for producing visually realistic outputs but often come with challenges such as training instability and mode collapse. Despite their strength, transformer-based models often need a lot of data and processing power, which can make them less useful in some situations citedosovitskiy-2021. These limitations contributed to the decision to prioritize a more efficient and adaptable architecture in the present work.

The research community has also developed numerous domain-specific datasets to 6 2.2 Related Works support enhancement tasks. Common datasets include the LOL [10] dataset for low-light images, EUVP [11] and UIEB [12] for underwater images and RESIDE [13] for hazy scenes. These datasets, which offer a variety of real-world and artificial images, differ in complexity and style. In work, a representative and balanced tri-domain dataset was constructed by selecting and organizing relevant samples from these sources to enable the training of a unified model. Finally, evaluation of image enhancement models often relies on both generic and domain specific metrics. Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), which evaluate reconstruction quality at the pixel level, are two often used quantitative metrics. Additionally, task specific metrics like Underwater Image Quality Measure (UIQM) and Fog Aware Density Evaluator (FADE) are employed to better capture perceptual quality in their respective domains.

2.2 RELATED WORKS

The field of image enhancement has evolved significantly due to its diverse applications. In this section, we introduce some recent advancements in low-light, underwater and hazy image enhancement, each contributing to the development of effective solutions for image quality improvement.

2.2.1 LOW-LIGHT IMAGE

Enhancement Due to insufficient lighting during capture, low-light photos frequently have poor illumination, excessive noise levels, and a loss of texture information. These degradations not only affect visual quality but also hinder downstream tasks like object detection and recognition. Enhancing such images involves addressing issues of brightness, contrast and noise suppression while preserving fine structural information. Li et al. [14] Offered Kinder, a better version of U-Net that brightens and sharpens low-light photos with improved skip connections and customized loss mechanisms. It excels at restoring color and detail but occasionally amplifies noise in very dark regions as the network does not explicitly model sensor specific noise under extreme illumination. Park et al. [15] introduced DEDU, Dual Enhancing Dense U-Net, which fuses dense blocks, decoupled fully connected attention and cross-stage partial connections. It jointly performs brightness enhancement and denoising, achieving strong PSNR/SSIM results. However, its computational complexity and GPU memory demand may limit real time deployment. Lim et al. [16] combines channel-spatial attention modules with a resizing sub-network to improve low-light contrast.

2.2.2 UNDERWATER IMAGE ENHANCEMENT

The physical characteristics of water, such as light scattering, absorption, and turbidity, present special difficulties for underwater photography. These elements cause a decrease in visibility, color distortion and haziness making underwater image enhancement a complex task. Effective enhancement requires restoring true colors, improving contrast and recovering details lost due to environmental degradation. Recent approaches have explored physics based models and learning based methods to tackle these issues. Wang et al. present OUNet-JL, a U-Net variation that combines a strengthen-operatesubtract reconstruction module (SOSFM), a spatial multi-scale fused with channel attention (SMFM), and a multi-residual block (MRM) [20]. Combining four setbacks functions during training, the model excels in restoring detail, correcting color imbalance and reducing noise. However, its intricate architecture increases computational load, which may limit performance on less powerful devices. Zhu et al. [21] builds on U-Net with enhanced skip connections and dynamic channel attention to better handle color cast and regional contrast. It achieves strong perceptual quality but occasionally struggles with over-saturation in regions heavily affected by blue-green tones. Yang et al. [22] worked on LU2Net, a

streamlined U-Net that uses axial depthwise convolution and channel attention to balance performance and efficiency. It processes images up to eight times faster than many recent models while maintaining competitive quality though slightly trailing behind heavier models in very low visibility conditions. Han et al. [23] showed a semi-supervised approach using limited labeled samples paired with abundant unlabeled images in a UNet architecture, improving adaptability across various water types.

2.2.3 HAZY IMAGE ENHANCEMENT

The crucial role of hazy image enhancement in remote sensing, autonomous driving, and outdoor vision equipment has drawn a lot of concern. Haze, caused by particles in the atmosphere scattering light, leads to reduced contrast, faded colors and depth-dependent visibility loss. Addressing these issues involves estimating scene depth or transmission maps to recover the original image clarity. Traditional approaches rely on physical models like the atmospheric scattering model, while recent learning based methods leverage deep networks to infer clean images directly from hazy inputs. Zamir et al. [26] merges the global context capability of the Swin Transformer with a U-Net backbone. It achieves strong dehazing across varying haze densities. However, sliding-window transformers can introduce latency in high resolution inputs. Chen et al. established DEA-Net: Detail-Enhanced Attention Network, which outperforms various SOTA models by using content-guided attention blocks and detail-enhanced convolutions [27]. However, its reliance on attention maps may falter in extremely dense haze. Zhang and Tao [28] made a lightweight, multi-scale encoder-decoder CNN that provides efficient dehazing with competitive accuracy.

3 PROPOSED WORKS

3.1 PROPOSED METHODOLOGY

This thesis proposes a unified deep learning based framework for image enhancement that effectively addresses three major types of visual degradation: low-light, underwater and hazy conditions. Unlike conventional approaches that rely on separate models for each domain, the proposed framework is designed to generalize across multiple degradation types using a single architecture. This unified approach not only reduces the computational burden of maintaining separate models but also enhances scalability and real world applicability. An improved U-Net architecture combined with attention techniques forms the foundation of the suggested system. The U-Net structure is ideal to gain knowledge both fine-grained features and broad context because of its encoder-decoder pathway and skip links. However, to improve its domain adaptability and focus on critical visual regions, attention modules are incorporated into the network. These modules help the model selectively emphasize features relevant to enhancement tasks, such as texture preservation in hazy images, color correction in underwater scenes and contrast

recovery in low-light inputs. To train the model effectively, a custom tri-domain paired dataset was constructed by combining representative samples from publicly available datasets. The model may learn domain-specific and shared enhancement features under supervision thanks to this dataset, which contains input and ground truth image pairings for each type of degradation. The model is trained using a combination of loss functions that account for pixel-wise accuracy, structural similarity and perceptual quality, ensuring balanced restoration across different domains. The proposed framework aims to achieve high quality visual enhancement using a single, scalable and interpretable model. By leveraging the strengths of U-Net and attention mechanisms and by training on a diverse dataset, the framework is expected to deliver consistent results across varying environmental conditions without the need for retraining or model switching.

Dataset To train and evaluate the proposed unified image enhancement framework, a custom tridomain paired dataset was constructed, covering three challenging visual conditions: lowlight, underwater and hazy environments. For each domain, paired images consisting of degraded inputs and corresponding clean ground truth references were collected and organized in a consistent format suitable for supervised learning as shown in Figure 3.1.



Figure 3.1: An overview of the dataset.

For the low-light domain, the dataset was curated from the LoLI-Street dataset, which contains real world dim street images paired with their well-lit ground truth counterparts. This dataset captures common challenges such as poor illumination, shadow noise and contrast loss, making it suitable for training a model to handle natural low-light scenarios. For the underwater domain, image pairs were extracted from the paired subset of the EUVP (Enhancing Underwater Visual Perception) dataset. EUVP includes a diverse collection of underwater scenes with varying visibility and color cast issues. The paired subset provides low quality input images alongside corresponding high quality references, allowing the model to learn effective color correction, detail restoration and contrast enhancement under underwater conditions.

3.3 DATA PRE-PROCESSING

To guarantee consistent training behavior and improve a deep learning model's capacity for generalization, effective data pre-processing is an essential step. In this work, a systematic pre-processing pipeline was applied to the tri-domain paired dataset before feeding it into the unified enhancement model. **Image Resizing and Normalization** Every image from the three domains a low light, submerged, and hazy was scaled to a uniform spatial resolution of 256×256 pixels. Resizing ensures uniformity across batches during training and simplifies model architecture requirements. After resizing, pixel intensity values were normalized to the $[0, 1]$ range by dividing each pixel by 255. This normalization step helps in stabilizing gradient updates and accelerates convergence during training. **Data Augmentation** During training, a number of data augmentation strategies were used to improve the model's reliability and prevent overfitting. These featured random flipping in both horizontal and vertical orientations, color property changes, brightness and contrast alterations, and random rotations. To preserve the correct mapping required for supervised learning, the input photos and their corresponding ground truth images received the same modifications simultaneously. **Domain Labeling** Each image pair was associated with a domain label as shown in Figure 3.2 to inform the model of the degradation type (e.g., low-light, underwater or hazy).

	gt_path	domain_label
2689	/kaggle/input/poly-dataset/tri_dataset/llight_...	0
9243	/kaggle/input/poly-dataset/tri_dataset/hazy_gt...	2
3816	/kaggle/input/poly-dataset/tri_dataset/uwater_...	1
287	/kaggle/input/poly-dataset/tri_dataset/llight_...	0
8174	/kaggle/input/poly-dataset/tri_dataset/hazy_gt...	2

Figure 3.2: Domain Labeling

These labels were used as conditional inputs during training, enabling the network to adapt its enhancement behavior based on the specific degradation characteristics. This mechanism is essential for achieving domain aware processing in a unified model.

Dataset Splitting

The pre-processed dataset was split into training, validation and testing sets in an 70:15:15 ratio. The data was split into training (70%), validation (15%) and testing (15%) subsets. The final dataset structure was organized into consistent folders such as lowlight_input, underwater_input, hazy_input for inputs and corresponding ground truth folders for ground truths as shown in Figure 3.3. This split was performed separately for each domain to ensure balanced representation and reliable evaluation. No image from the same scenario appeared in more than one split because great care was taken to prevent data leakage across subsets.

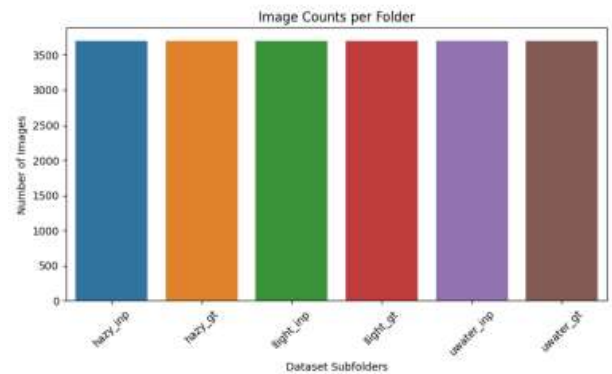
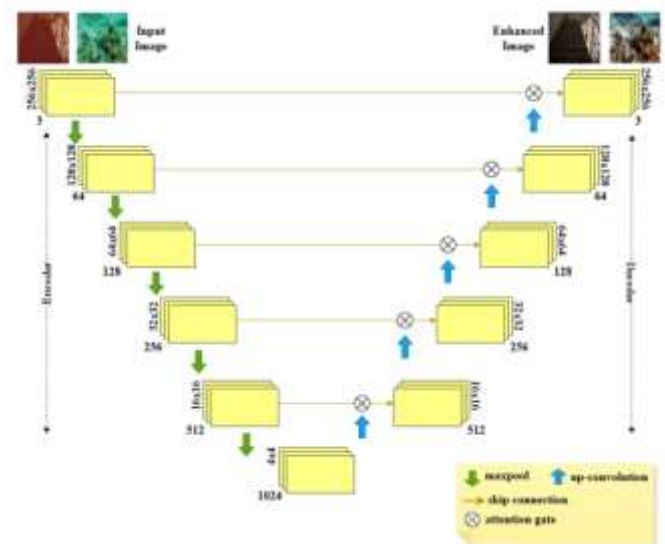


Figure 3.3: Dataset Distribution

This structured and consistent pre-processing pipeline played a vital role in enabling the attention based U-Net to learn meaningful enhancement mappings across all three domains while maintaining domain adaptability and minimizing overfitting. This unified structure ensures that all domains can be processed through a shared data pipeline during training and evaluation.

3.4 MODEL ARCHITECTURE

This section outlines the technique and architecture of the anticipated unified model designed for enhancing both low-light, underwater and hazy images.



In order to accomplish targeted feature modulation, the suggested model architecture shown in Figure 3.4 its extends the traditional U-Net framework by adding Attention Gates into the skip connections.

Encoder Path The encoder uses 2×2 max-pooling operations for down sampling after five stages of convolutional processing. Each level consists of two consecutive 3×3 convolutional layers, followed by ReLU activation and Batch Normalization. The input image of size $256 \times 256 \times 3$ is transformed progressively to smaller spatial dimensions while increasing feature depth from 64 channels at

the first level to 1024 channels at bottleneck at the very bottom. This approach uses hierarchical representation to capture both high-level and low-level features. For skip connections, the down-sampled feature maps are kept and sent to the bottom. Each convolutional block in the encoder produces the following outcome

$$C_{di} = \sigma(W * M + b),$$

where:

- W represents the convolution filter weights,
- M is the input feature map,
- b is the bias term,
- σ is the activation function (ReLU),
- $*$ signifies the convolution operation.

Feature Extraction and Selection

Each convolution layer in the network employs learnable filters (also referred to as kernels) of size 3×3 , designed to extract local features from the input images. In order to preserve and recreate small details throughout the picture enhancement process, these filters are crucial in distinguishing spatial elements including edges, textures, and gradients. The filters are initialized randomly and are optimized during training via backpropagation to capture features that are specific to low-light and underwater degradations. The use of small sized filters (3×3) allows the model to learn complex visual patterns while preserving spatial locality. The proposed architecture performs feature extraction primarily through its encoder which consists of stacked convolution layers at each level. These layers apply 3×3 learnable filters to the input and intermediate feature maps, capturing spatial patterns such as edges, textures and structural variations.

Attention Gates

Each skip link incorporates specific attention approaches whenever the encoder features are concatenated with the decoder outputs. Each Attention Gate receives two inputs: the encoder map of characteristics and the matching decoder stage's triggering signal. It computes attention coefficients through 1×1 convolutions followed by batch normalization, a Sigmoid function for output activation and a ReLU activation for adding non-linearity. These coefficients are then applied to the encoder features via element wise multiplication, producing modulated features that emphasize spatially relevant information. For a given gating signal g from the decoder and encoder feature map M , the attention coefficient ψ is computed as:

$$\psi = \sigma(BN(P_x \cdot g + P_y \cdot M + b))$$

where:

- P_x and P_y are learnable parameters from 1×1 convolutions applied to x and y ,
- B_N denotes Batch Normalization,
- σ is the Sigmoid activation function,
- b is a bias term.

The modulated encoder features are then obtained as:

$$M' = M \cdot \psi$$

This operation enables the network to emphasize relevant spatial regions while suppressing noise and irrelevant details, which is particularly beneficial for enhancing degraded low-light, underwater and hazy images.

Decoder Path

The decoder reconstructs the enhanced image by successively up-sampling the feature maps using transposed convolutions (up-convolutions). At each level the up-sampled output is concatenated with the corresponding encoder features filtered through the attention gate. The concatenated feature map is passed through two convolution layers with Batch Normalization and ReLU activation. The spatial dimensions are restored in stages: from 16×16 to 256×256 . The up-sampling operation is mathematically expressed as:

$$C_{ups} = W^{TC} * M,$$

- where W^{TC} represents the transposed convolution filter weights,
- where M is the input feature map.

Final Output Layer

The final layer applies a 1×1 convolution to reduce the output channels to 3, corresponding to the RGB image format. This generates the final enhanced image, where each pixel value is reconstructed through direct regression from the decoder final feature map. This operation allows for pixel-wise enhancement tailored to visual restoration rather than classification or segmentation objectives.

4 EXPERIMENTS AND RESULTS

4.1 HARDWARE CONFIGURATION, SOFTWARE AND LIBRARIES

The experimental evaluation of the proposed framework for enhancing low-light, underwater, and hazy images was carried out in a structured and reproducible setup. All experimentation was carried out on Kaggle, leveraging its cloud based infrastructure equipped with GPU support. Hardware configurations included NVIDIA Tesla P100, Intel Xeon Processors (2 vCPUs) and 16 GB system RAM. Additionally, up to 100 GB of cloud storage was used for data and model

checkpoints. The GPU accelerated the training process significantly, enabling efficient parallel processing of convolution operations within the U-Net model, which is particularly critical for handling large datasets. The software stack employed open source tools and libraries for implementation, training and evaluation. The experiments were conducted in a Linux based Kaggle environment using Python 3.7. The framework relied on PyTorch (v1.10+) for building and training the U-Net architecture supported by Torchvision for image transformations. Additional libraries included NumPy and Pandas for data manipulation and organization, Pillow (PIL) for image processing and Scikit-learn for dataset splitting. Metrics like PSNR and SSIM were computed using Scikit-image, while Matplotlib was used for visualizing input, ground truth and enhanced outputs. This setup ensured a robust and efficient environment for model development and performance evaluation.

4.2 Model Training

The proposed Attention based U-Net model was trained in a supervised setting using the curated tri-domain paired dataset consisting of low-light, underwater and hazy images. The primary goal during training was to minimize the pixel-wise difference between the enhanced outputs and their corresponding ground truth images.

Optimization

The Adam optimizer, a sophisticated version of stochastic gradient descent that modifies the learning rate during training, was used to train the model. To guarantee a suitable trade-off between quick convergence and efficient generalization, an initial learning rate of $1e-4$ was used. To prevent overfitting, an early stopping approach based on validation loss was used during the 100 epochs of training.

Loss Function

To train the Attention U-Net model for enhancing all three domain images, we diminish the pixel-by-pixel discrepancy between the anticipated image and its ground truth counterpart. The model is trained using the Mean Squared Error (MSE) loss function, which measures the average squared difference between the predicted pixel values and the corresponding ground truth intensities. This objective guides the network to generate outputs that closely match the target images in both structure and visual appearance. Mathematically, MSE is defined as:

$$L_{MSE} = \frac{1}{P} \sum_{j=1}^P (Y_{true,j} - Y_{pred,j})^2$$

where P : number of pixels in the image, Y_{pred} the predicted output and Y_{true} the ground truth. This loss function is

minimized during training to ensure accurate reconstruction of brightness, contrast and structural details.

Data Sampling and Augmentation

During training, mini-batches were drawn from all three domains to ensure balanced exposure to each type of degradation. Each batch contained a mix of low-light, underwater and hazy images to encourage shared feature learning across domains. Input images and their corresponding ground truths were resized to 256×256 pixels and normalized to the range $[-1,1]$ using mean and standard deviation normalization. The following augmentations were applied during training: Horizontal and vertical flips, Random rotations, Brightness and contrast adjustments. These augmentations helped improve model robustness and reduce overfitting by simulating variability in the input conditions.

Adaptive Enhancement using Shared Encoder and Decoder

The proposed framework achieves adaptive enhancement for low-light, underwater and hazy images using a shared Attention U-Net architecture. Unlike models that require separate networks for different image degradation types, our approach makes use of a common encoder-decoder backbone to process diverse inputs within a unified architecture. The encoder path learns to extract hierarchical features from low-light, underwater and hazy images, capturing essential patterns such as textures, edges and global context. By sharing the encoder, the model generalizes common image enhancement features like noise suppression, contrast improvement and structural detail preservation. To address the domain specific characteristics (e.g., brightness enhancement in low-light images and color correction in under-water images), Attention Gates are introduced in the skip connections. These gates dynamically emphasize relevant spatial features while suppressing irrelevant activations, allowing the model to adaptively focus on degraded regions depending on the input image type. The decoder path reconstructs enhanced images by integrating high level semantic information from the encoder with attention weighted features, ensuring precise restoration of visual quality.

Observation During Training We can See in Figure 4.1 throughout training, loss curves showed steady convergence without signs of instability or overfitting. Adequate generalization was demonstrated by the validation loss right after the training loss. While early epochs yielded outputs with washed out textures, progressive training led to noticeable improvements in contrast, clarity and color accuracy. Visual inspection confirmed that the model could enhance visibility in dark images, correct color cast in underwater scenes and reduce haze induced blur, all using a shared network trained with a simple yet effective objective function.

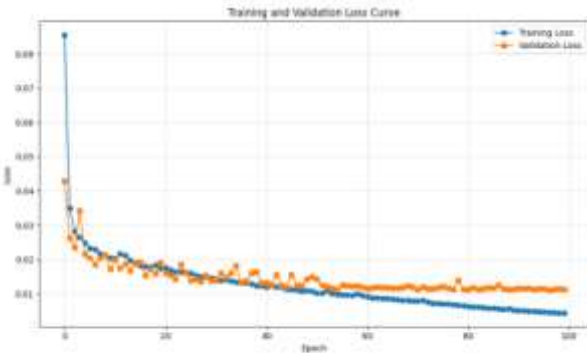


Figure 4.1: Training and Validation Loss Curve

4.3 EVALUATION

Both quantitative metrics and qualitative analytic techniques were used to fully assess the efficacy of the suggested attention-based U-Net framework. The evaluation was performed across all three target domains low-light, underwater and hazy using the corresponding test sets from the curated tri-domain paired dataset. The evaluation phase aimed to measure how well the model improves visual quality across various types of image degradation while preserving consistency and perceptual realism.

4.3.1 QUANTITATIVE METRICS

A combination of general purpose and domain specific metrics was used to evaluate the model. These include Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), Underwater Image Quality Measure (UIQM) [31], Exposure based Perceptual Index (EPI) [32] and Fog Aware Density Evaluator (FADE) [33]. For underwater images, PSNR and SSIM may not fully reflect color fidelity and visual appeal. Therefore, UIQM was used as a domain specific metric. UIQM combines three components: UICM (colorfulness), UISM (sharpness), UICoM (contrast). This composite score is designed to evaluate overall underwater image quality, with higher scores indicating more visually pleasing results.

4.4 RESULTS

This section presents the outcomes of the proposed attention-based U-Net framework across all three target domains low-light, underwater and hazy on the unified test set. The evaluation includes both quantitative and qualitative analyses, supported by domain specific image quality metrics and visual inspections. Additionally, key observations are discussed to reflect on the strengths and limitations of the model.

4.4.1 QUANTITATIVE EVALUATION

Table 4.1: Evaluation scores across different image domains

Domain	PSNR(dB) (↑)	SSIM (↑)
Low-Light	36.36	0.9731
Underwater	24.14	0.8461
Hazy	26.89	0.9336
Overall	29.13	0.9176

Table 4.2: Domain-specific perceptual metric scores

Domain	Low-Light	Underwater	Hazy
Metric	EPI(↓): 0.262	UIQM(↑): 24.308	FADE(↓): 0.049

Note: All evaluations were conducted on the same tri-domain dataset comprising low-light, underwater, and hazy images.

Standard measures including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) were used to objectively assess the effectiveness of the suggested attention-based U-Net model, along with domain-specific measures including EPI, UIQM, and FADE to provide a comprehensive statistical assessment. These metrics were chosen to reflect both the pixel-wise accuracy of reconstruction and the perceptual quality of the enhanced images across each domain. The achieved scores are shown below in Table 4.1 and Table 4.2. In the low-light domain, the model achieved a PSNR of 36.36 dB and an SSIM of 0.9731. In image enhancement tasks, a PSNR value exceeding 30 dB is generally regarded as a strong result, suggesting that the reconstructed image closely matches the ground truth in terms of pixel-level accuracy[34]. Similarly, SSIM values above 0.95 denote high structural similarity [35], suggesting that the model has effectively restored image details such as edges, textures and contrast.

For the underwater domain, the model obtained a PSNR of 24.14 dB and an SSIM of 0.8461. Although these values are slightly lower than those observed in the low-light domain, they are still within acceptable ranges. In underwater image enhancement, a PSNR above 22 dB and an SSIM above 0.80 are considered reasonable due to the inherently complex distortions caused by water scattering, color absorption and turbidity [34, 35]. The UIQM (Underwater Image Quality Measure) score for this domain was 24.308, which falls within the typical good-quality range (20-30) [31]. This suggests that the enhanced images showed notable improvements in color balance, sharpness and visibility, even though underwater images are generally more difficult to process due to their non-uniform degradation

For hazy images, the model achieved a PSNR of 26.89 dB along with an SSIM value of 0.9336. These scores indicate that the model was effective at recovering visibility and texture in foggy or misty scenes. PSNR values in the range of 25-30 dB are commonly accepted as high-quality restorations in the context of dehazing [34]. Additionally, the SSIM value near 0.93 suggests strong structural consistency between the enhanced outputs and the ground truth [35]. FADE (Fog Aware Density Evaluator) score was 0.0498, which is notably

low indicating successful suppression of haze content. In FADE, lower scores represent reduced fog density and improved scene clarity and a value below 0.1 is generally associated with perceptually clean, dehazed outputs [33]. When evaluated on the entire unified test set, which included all three domains, the model achieved an average PSNR of 29.13 dB and an SSIM of 0.9176. These results demonstrate that the model was not only capable of domain specific enhancement but also performed consistently when tested on a diverse mixture of degraded inputs. A PSNR close to 30 dB and SSIM above 0.90 for a unified framework indicate strong generalization and robustness, an important goal of this work. These metrics affirm the effectiveness of the proposed attention based U-Net model in handling a wide variety of image degradation types using a single architecture.

4.4.2 QUALITATIVE EVALUATION

Visual inspection is a critical component of image enhancement assessment, particularly in domains like underwater and hazy imaging, where distortions are often complex, nonuniform and subjective in nature. For this evaluation, side-by-side comparisons were generated showing the original degraded input, the enhanced output and the corresponding ground truth image. Representative samples from all three domains - low-light, underwater and hazy were selected and arranged into grid views for better interpretability.

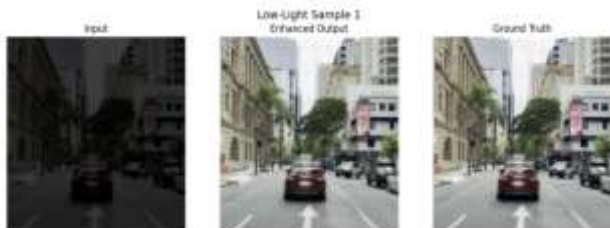


Figure 4.2: Visual grid of enhanced low-light image along with degraded input and ground truth

In the low-light domain, the enhanced images showed substantial improvements in brightness, contrast, and detail recovery. The model was able to brighten dark scenes without introducing unnatural artifacts or overexposure. As seen in the provided visual grid Figure 4.2, the outputs preserved structural details such as road markings, vehicle textures, and building edges. Color tones were well balanced, and noise amplification, a common challenge in low-light enhancement was effectively minimized



Figure 4.3: Visual grid of enhanced underwater image along with degraded input and ground truth

For the underwater domain, the visual results indicated notable improvement in clarity, color fidelity and sharpness. Underwater images typically suffer from a blue-green color cast, low contrast and detail suppression due to light absorption and scattering. The enhanced output in Figure 4.3 exhibited successful color correction, with warmer tones restored in coral, marine life and background textures. Fine features such as fish patterns, sea plants and water gradients were distinctly visible in the enhanced images.



Figure 4.4: Visual grid of enhanced hazy image along with degraded input and ground truth

In the case of hazy images, the model demonstrated strong performance in restoring visibility and depth. Haze often leads to loss of contrast, muted colors and diminished object boundaries, especially in distant regions. The enhanced outputs in this domain showed clear skies, more saturated building textures and improved foreground background separation. Distant structures became sharper and the depth perception of the scene was effectively recovered. Compared to the hazy inputs, the outputs revealed details previously obscured, bringing the images closer in quality to the clear ground truth references.

4.5 OBSERVATIONS

The training and evaluation process revealed several key observations about the behavior and effectiveness of the proposed attention based U-Net model across the three targeted image degradation domains. These insights not only validate the performance of the model but also help identify its current limitations and future potential for improvement.

To assess the accuracy and consistency of the enhanced outputs across different domains, scatter plots were generated by plotting predicted pixel intensities (from the model output) against their corresponding ground truth values. An ideal enhancement would result in all points falling along the diagonal line ($y=x$), indicating perfect prediction.

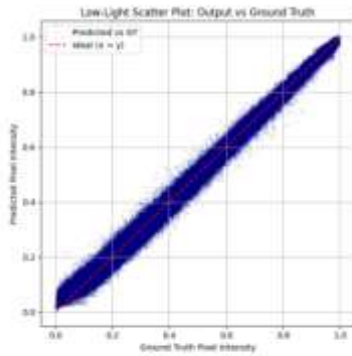


Figure 4.5: Scatter plot of pixel intensities for Low-light images

The narrow spread and sharp alignment demonstrate low prediction variance, which corresponds well with the high PSNR (36.36 dB) and SSIM (0.9731) observed in this domain. It also suggests that the model is highly reliable when handling structured low-light environments with consistent contrast patterns. In the underwater domain, the presented scatter plot 4.6 reveals a wider and more diffuse distribution compared to the other domains. While the overall trend still follows the diagonal, the deviations are more pronounced, particularly at mid to high intensity ranges. This dispersion is likely due to the complex nature of underwater distortions, including non-linear color shifts, scattering and refraction, which are challenging to recover using pixel-wise losses alone. The scatter behavior aligns with the lower SSIM (0.8461) and UIQM (24.3), indicating that while the model performs well, there remains room for improvement, particularly in preserving fine details and dynamic contrast in underwater environments.

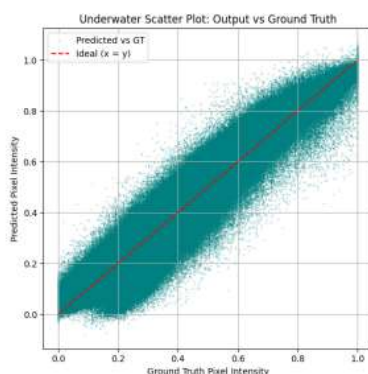


Figure 4.6: Scatter plot of pixel intensities for Underwater images

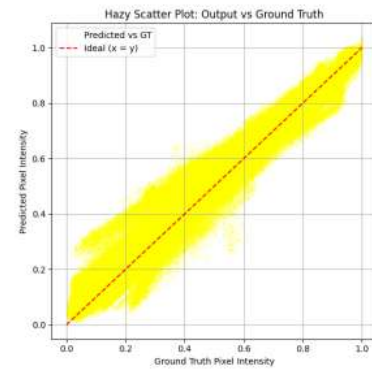


Figure 4.7: Scatter plot of pixel intensities for Hazy images

The scatter figure in Figure 4.6 for the hazy domain likewise demonstrates a significant correlation between ground truth and anticipated intensities, but with a little greater dispersion than in low-light. The spread is broader near the lower intensity values, which may reflect difficulty in perfectly reconstructing heavily faded or low contrast areas affected by haze.

4.5.1 ERROR VISUALIZATION

To gain deeper insight into the spatial accuracy of the enhanced outputs, error heatmaps were generated for representative samples from each domain. Darker regions (red to black) indicate smaller discrepancies, while brighter areas (yellow to white) represent higher levels of error.



Figure 4.8: Error heatmap of enhanced low-light image along with ground truth

In the low-light domain, the error heatmap in Figure 4.8 shows that the majority of the frame has very low error intensity, particularly in structured areas like buildings, roads and vehicles. The error values are concentrated near small texture details such as tree leaves or reflective surfaces like car windows and license plates.

However, with peak error values remaining below 0.08 and largely localized, the performance of the model in low-light conditions is both quantitatively and perceptually robust. In the underwater domain, the error heatmap in Figure 4.9 reveals more prominent and dispersed differences. The highest error regions are observed around fine textures in the coral and jellyfish structures, particularly where transparency, depth or lighting varies sharply. This aligns with expectations, as underwater imagery presents more complex visual degradations that are difficult to model with MSE alone.

With peak error values reaching 0.35 in localized spots, the heatmap confirms that this domain remains the most challenging for the model, especially in scenes with intricate patterns or strong chromatic distortions. Nevertheless, the overall enhancement still significantly improves visibility and color balance and the structural coherence of the scene remains intact.



Figure 4.9: Error heatmap of enhanced low-light image along with ground truth



Figure 4.10: Error heatmap of enhanced low-light image along with ground truth

The hazy domain displays a more distributed pattern of error in Figure 4.10. The heat map highlights increased error along object boundaries and regions with strong contrast transitions, such as vehicle edges and human silhouettes. These regions correspond to the areas where haze most strongly distorts the image and where precise restoration is most challenging. Although some over or under enhancement is visible in these zones, the overall error values remain moderate, with the highest intensity around 0.14. The model still manages to maintain good structural consistency and recovers depth effectively, but fine boundary textures could be further improved with perceptual or edge aware loss functions. Overall, these heat maps support the quantitative results and offer a spatial understanding of behavior of the model. The attention based U-Net shows high precision in uniform and well lit regions while facing more error in areas with fine grained variation, depth ambiguity or non-uniform lighting - typical challenges in real world image degradation. These results imply that the model may benefit from additional improvement in edge sensitive areas, perhaps through perceptual guided learning or multi-scale loss integration in subsequent research.

5. CONCLUSION AND FUTURE WORK

5.1 CONCLUSIONS

Due to its numerous uses in fields including augmented reality, photographic, telemetry, remote sensing and undersea research image enhancement has garnered a lot of attention in recent years. Although deep learning has made significant progress in this field, the majority of cutting edge methods are

still domain-specific; concentrating on improving photographs impacted by either low light, haze, or underwater distortions rarely all three inside a single, cohesive system. In order to close this gap, a generic image improvement framework that can handle several difficult visual degradations with single model architecture was proposed and evaluated in this thesis. The proposed framework is built upon an attention based U-Net, which retains the strengths of the original U-Net in extracting multi scale spatial, features and enhances it further through the integration of attention modules. These attention mechanisms empower the network to selectively focus on regions of interest, leading to improved restoration of fine structures, contrast and perceptual details. Importantly, the same model is used across all domains without modification or retraining, thus supporting the central hypothesis that a unified enhancement approach is feasible, efficient and effective. To support the development and evaluation of this approach, a tri-domain paired dataset was constructed by combining representative image pairs from three publicly available and widely recognized datasets: LoLI-Street for low-light images, EUVP for underwater images, and RESIDE for hazy images. Each image pair consisted of a degraded input and its corresponding high quality reference, enabling the model to be trained in a supervised manner.

The Adam optimizer and a straightforward but efficient Mean Squared Error (MSE) loss function were utilized to train the model. Despite the absence of complex loss components such as perceptual loss, the model demonstrated high accuracy and visual fidelity. Evaluation was performed using both general purpose metrics - PSNR and SSIM, as well as domain specific metrics like UIQM for underwater enhancement, EPI for low-light exposure quality and FADE for haze density estimation. The results were highly promising. On the low-light test set, the model achieved a PSNR of 36.36 dB and an SSIM of 0.9731, which are considered excellent by both academic and industry standards. In the underwater domain, the model delivered a PSNR of 24.14 dB, SSIM of 0.8461, and a UIQM score of 24.308, reflecting significant improvements in visual clarity and color balance. The overall performance across the unified test set remained consistent, with an average PSNR of 29.13 dB and SSIM of 0.9176. Qualitative analysis further confirmed these findings. Visual comparisons showed that the model successfully brightened dark regions, removed haze and corrected underwater color casts, producing outputs that were perceptually close to the ground truth. Error heat maps and pixel intensity scatter plots provided deeper insights into the behavior of the model, revealing high spatial precision and structural coherence, particularly in the low-light and hazy domains. In conclusion, this thesis has demonstrated that a single, well designed deep learning model, augmented with attention mechanisms, can perform effective enhancement across diverse degradation domains. The unified framework eliminates the need for domain specific networks, reducing complexity and making deployment more scalable in real world applications. This work contributes a significant step

toward generalized, efficient and high quality image enhancement under natural and adverse environmental conditions.

5.2 FUTURE WORKS

While the results achieved in this thesis demonstrate the feasibility and effectiveness of a unified image enhancement framework, there remains opportunity to expand and improve the system in future research. These enhancements may target model performance, training strategies, architectural improvements and broader applications:

- Architectural Enhancements - Explore hybrid architectures combining transformer based encoders and convolutional decoders for improved contextual modeling.
- Dataset Expansion - Extend the tri-domain dataset to include other challenging conditions such as rain, snow or sand, promoting broader generalizability.
- Unsupervised and Semi-supervised Learning - Leverage unpaired data with unsupervised or semi-supervised methods to mitigate the lack of paired ground truth in domains like underwater and hazy images.
- Loss Function Improvements - Incorporate a multi-loss strategy combining perceptual and Total Variation (TV) losses to improve perceptual quality in underwater and hazy domains.

REFERENCES

- [1] S. Anwar and N. Barnes, "Densely residual laplacian super-resolution," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 3, pp. 1192–1204, 2020.
- [2] K. G. Lore, A. Akintayo, and S. Sarkar, "Llnet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognition*, vol. 61, pp. 650–662, 2017.
- [3] R. C. Gonzalez, *Digital image processing*. Pearson education india, 2009.
- [4] Y. Zhang, X. Di, B. Zhang, and C. Wang, "Self-supervised image enhancement network: Training with low light images only," *arXiv preprint arXiv:2002.11300*, 2020.
- [5] D. Akkaynak and T. Treibitz, "Sea-thru: A method for removing water from underwater images," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 1682–1691.
- [6] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior," *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 12, pp. 2341–2353, 2010.
- [7] Z. Wang, J. Chen, and S. C. H. Hoi, "Deep Learning for Image SuperResolution: A survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 10, pp. 3365–3387, 3 2020. [Online]. Available: <https://doi.org/10.1109/tpami.2020.2982166>
- [8] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical image computing and computerassisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5–9, 2015, proceedings, part III* 18. Springer, 2015, pp. 234–241.
- [9] O. Oktay, J. Schlemper, L. L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N. Y. Hammerla, B. Kainz et al., "Attention u-net: Learning where to look for the pancreas," *arXiv preprint arXiv:1804.03999*, 2018.
- [10] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep retinex decomposition for low-light enhancement," *arXiv preprint arXiv:1808.04560*, 2018.
- [11] M. J. Islam, Y. Xia, and J. Sattar, "Fast underwater image enhancement for improved visual perception," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 3227–3234, 2020.
- [12] C. Li, C. Guo, W. Ren, R. Cong, J. Hou, S. Kwong, and D. Tao, "An underwater image enhancement benchmark dataset and beyond," *IEEE transactions on image processing*, vol. 29, pp. 4376–4389, 2019.
- [13] B. Li, W. Ren, D. Fu, D. Tao, D. Feng, W. Zeng, and Z. Wang, "Benchmarking single-image dehazing and beyond," *IEEE Transactions on Image Processing*, vol. 28, no. 1, pp. 492–505, 2019.
- [14] H. Li, Y. Yu, and Y. Wang, "Kinder: Dark light image enhancement method based on improved u-net," in *2023 International Conference on Computer Applications Technology (CCAT)*. IEEE, 2023, pp. 318–323.
- [15] H. Park, H. Lim, and D. Jang, "Dedu: Dual-enhancing dense-unet for lowlight image enhancement and denoise," *IEEE Access*, vol. 12, pp. 24 071–24 078, 2024.
- [16] C. C. Lim, Y. P. Loh, and L.-K. Wong, "Lau-net: A low light image enhancer with attention and resizing mechanisms," *Signal Processing: Image Communication*, vol. 115, p. 116971, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S092359652300053X>
- [17] J. Y. Park, C. W. Park, and I. K. Eom, "Ulbnet: Low-light image enhancement using u-shaped lightning back-projection," *Knowledge-Based Systems*, vol. 281, p. 111099, 2023. [Online]. Available:

<https://www.sciencedirect.com/science/article/pii/S0950705123008493>

[18] X. Wei, X. Zhang, S. Wang, Y. Huang, and Y. Li, "Tsn-ca: A two-stage network with channel attention for low-light image enhancement," 2021. [Online]. Available: <https://arxiv.org/abs/2110.02477>

[19] Y. Jin, Y. Sun, J. Liang, X. Yan, Z. Hou, S. Zheng, Y. Chen, and X. Chen, "A hybrid framework for curve estimation based low light image enhancement," *Scientific Reports*, vol. 15, no. 1, 3 2025. [Online]. Available: <https://doi.org/10.1038/s41598-025-92161-y>

[20] X. Wang, Z. Luo, W. Huang, Y. Zhang, and R. Hu, "Optimized unet framework with a joint loss function for underwater image enhancement," *Scientific Reports*, vol. 15, 03 2025.

[21] S. Zhu, Z. Geng, Y. Xie, Z. Zhang, H. Yan, X. Zhou, H. Jin, and X. Fan, "New underwater image enhancement algorithm based on improved u-net," *Water*, vol. 17, no. 6, 2025. [Online]. Available: <https://www.mdpi.com/2073-4441/17/6/808>

[22] H. Yang, J. Xu, Z. Lin, and J. He, "Lu2net: A lightweight network for real-time underwater image enhancement," 2024. [Online]. Available: <https://arxiv.org/abs/2406.14973>

[23] J. Han, X. Liu, D. Tan, and J. Zhu, "An improved semi-supervised learning-based u-net for underwater image enhancement," in *2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT)*, 2024, pp. 1389–1393.

[24] D. Zhang, W. Cao, J. Zhou, Y.-T. Peng, W. Zhang, and Z. Lin, "Two-branch underwater image enhancement and original resolution information optimization strategy in ocean observation," *Journal of Marine Science and Engineering*, vol. 11, no. 7, 2023. [Online]. Available: <https://www.mdpi.com/2077-1312/11/7/1285>

[25] J. Jiang, T. Ye, J. Bai, S. Chen, W. Chai, S. Jun, Y. Liu, and E. Chen, "Five a+ network: You only need 9k parameters for underwater image enhancement," 2023. [Online]. Available: <https://arxiv.org/abs/2305.08824>

[26] P. Maldonado-Quispe and H. Pedrini, "Swindehazing: Haze removal using u-net and swin transformer." Berlin, Heidelberg: Springer-Verlag, 2024, p. 104–117. [Online]. Available: https://doi.org/10.1007/978-3-031-76607-7_8

[27] Z. Chen, Z. He, and Z.-M. Lu, "Dea-net: Single image dehazing based on detail-enhanced convolution and content-guided attention," 2023. [Online]. Available: <https://arxiv.org/abs/2301.04805>

[28] J. Zhang and D. Tao, "Famed-net: A fast and accurate multi-scale end-to-end dehazing network," *IEEE Transactions on Image Processing*, vol. 29, p. 72–84, 2020. [Online]. Available: <http://dx.doi.org/10.1109/TIP.2019.2922837>

[29] H. Zhou, Z. Chen, Q. Li, and T. Tao, "Dehaze-unet: A lightweight network based on unet for single-image dehazing," *Electronics*, vol. 13, no. 11, 2024. [Online]. Available: <https://www.mdpi.com/2079-9292/13/11/2082>

[30] L. Lu, Q. Xiong, B. Xu, and D. Chu, "MixDehazeNet: Mix structure block for image dehazing network," *2022 International Joint Conference on Neural Networks (IJCNN)*, pp. 1–10, 6 2024. [Online]. Available: <https://doi.org/10.1109/ijcnn60899.2024.10651326>

[31] K. Panetta, C. Gao, and S. Agaian, "Human-Visual-System-Inspired underwater image quality measures," *IEEE Journal of Oceanic Engineering*, vol. 41, no. 3, pp. 541–551, 10 2015. [Online]. Available: <https://doi.org/10.1109/joe.2015.2469915>

[32] F. Sattar, L. Floreby, G. Salomonsson, and B. Lovstrom, "Image enhancement based on a nonlinear multiscale method," *IEEE transactions on image processing*, vol. 6, no. 6, pp. 888–895, 1997.

[33] N. Hautiere, J.-P. Tarel, D. Aubert, and E. Dumont, "Blind contrast enhancement assessment by gradient ratioing at visible edges," *Image Analysis and Stereology*, vol. 27, no. 2, pp. 87–95, 2008.

[34] U. Sara, M. Akter, and M. S. Uddin, "Image quality assessment through fsim, ssim, mse and psnrâa comparative study," *Journal of Computer and Communications*, vol. 7, no. 3, pp. 8–18, 2019. [35] 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.