A MAJOR PROJECT -II REPORT
ON
CONTENT GUIDED ATTENTION FOR UNDERWATER IMAGE ENHANCEMENT

Submitted in Partial Fulfilment of the Requirements
For The Award of the Degree of

MASTER OF TECHNOLOGY
IN
COMPUTER SCIENCE & ENGINEERING

Submitted By:
Minal Tandekar
2K22/CSE/11

Under the Supervision of
Dr. Anil Singh Parihar
Professor
Department of Computer Science and Engineering
Delhi Technological University

Department of Computer Science and Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

MAY, 2024

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to **Dr. Anil Singh Parihar**, Professor at the Department of Computer Science and Engineering, Delhi Technological University, for his invaluable guidance and unwavering encouragement throughout this research. Her vast knowledge, motivation, expertise, and insightful feedback have been instrumental in every aspect of preparing this research plan.

I am also grateful to **Prof. Dr. Vinod Kumar**, Head of the Department, for his valuable insights, suggestions, and meticulous evaluation of my research work. His expertise and scholarly guidance have significantly enhanced the quality of this thesis.

My heartfelt thanks go out to the esteemed faculty members of the Department of Computer Science and Engineering at Delhi Technological University. I extend my gratitude to my colleagues and friends for their unwavering support and encouragement during this challenging journey. Their intellectual exchanges, constructive critiques, and camaraderie have enriched my research experience and made it truly fulfilling.

While it is impossible to name everyone individually, I want to acknowledge the collective efforts and contributions of all those who have been part of this journey. Their constant love, encouragement, and support have been indispensable in completing this MTech thesis.

Minal Tandekar

(2K22/CSE/11)

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE DECLARATION

I MINAL TANDEKAR hereby certify that the work which is being presented in the thesis entitled **Content Guided attention for Underwater Image Enhancement** in partial fulfillment of the requirements for the award of the Degree of Master of Technology submitted in the Department of Computer Science and Engineering, Delhi Technological University in an authentic record of my work carried out during the period from August 2022 to May 2024 under the supervision of Dr. Anil Singh Parihar.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Minal
Tandekar

This is to certify that the student has incorporated all the corrections suggested by the examiner in the thesis and that the statement made by the candidate is correct to the best of our knowledge.

Signature of Supervisor(s) **Signature of External Examiner**

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Minal Tandekar (Roll no 2K22/CSE/11)** has carried out their project work presented in this thesis entitled "**Content Guided attention for Underwater Image Enhancement"** for the award of **Master of Technology** from the Department of Computer Science and Engineering, Delhi Technological University, Delhi under my supervision. The thesis embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Dr. Anil Singh Parihar
Professor,
Department of Computer Science and
Engineering,
DTU-Delhi,
India

Date:

Content Guided attention for Underwater Image Enhancement

Minal Tandekar

ABSTRACT

This thesis addresses the challenge of enhancing underwater image clarity using deep learning techniques, a crucial advancement for applications in marine biology, underwater archaeology, and environmental monitoring. Traditional enhancement methods typically fail to address the harsh light distortions encountered underwater, such as discoloration and blur due to light absorption and scattering This study uses Enhanced- . Encoder FUnIE-GAN (EEF-GAN), which is an updated version of Fast Underwater Image Enhancement comes GAN (FUnIEGAN), which is designed to overcome these challenges by adding new encoder structures The modified encoder uses traditional convolution side convolution difference serves to enhance feature extraction, thereby significantly improving image recognition Empirical results from extensive testing on UIEB dataset show that EEF -GN: peak signal The model outperforms existing models is available in several metrics including -to-noise ratio (PSNR) and structural similarity index (SSIM) giving a PSNR of 22.94 dB and a SSIM of 0.8926, a clear and accurate underwater image for comparison a to baseline models like WaterNet and UGAN Underscoring its effectiveness to create These findings not only demonstrate the feasibility of using generative anti-nets for real-time image enhancement in complex underwater environments but also demonstrate the potential of such technologies this has in other easily identifiable imaging applications Preferences involve enhancing images and optimizing the model This function by extending the capabilities of deep learning models Contributes to the environment extensive environmental mapping projects, providing new tools for research and conservation efforts.

Keywords: Underwater Image Enhancement, Deep Learning, GAN, Image Quality Metrics, CNN, Feature Extraction, Image Clarity and Quality, Convolution, Real-Time Image Processing.

TABLE OF CONTENT

LIST OF TABLE(S)

LIST OF FIGURE(S)

LIST OF ABBREVIATION(S)

CNN	Convolutional Neural Network
UGAN	Underwater Generative Adversarial Network
GAN	Generative Adversarial Network
FUnIE-GAN	Fast Underwater Image Enhancement Generative Adversarial Network
RAUE-Net	Residual Attention U-Net for Underwater Image Enhancement
MSDR-Net	Multi-Scale Deep Residual Network
USP-Net	Underwater Single Image Processing Network
UCC-Net	Underwater Color Correction Network
MISD	Medical Image Segmentation Dataset
UIEB	Underwater Image Enhancement Benchmark
EUVP	Enhanced Underwater Vision via Paired Samples
PSNR	Peak Signal-to-Noise Ratio
ILSVRC	ImageNet Large Scale Visual Recognition Challenge
dB	decibels
SSIM	Structural Similarity Index
MSE	Mean Squared Error
RMSE	Root Mean Squared Error
RUIE	Real-World Underwater Image Enhancement
EEF-GAN	Enhanced-Encoder Fast Underwater Image Enhancement Generative Adversarial Network

CHAPTER 1

INTRODUCTION

1.1 OVERVIEW

Underwater image enhancement refers to the process of improving the visual appearance of images taken underwater. This is an important area of study because of the variety of optical challenges encountered underwater
The environment. When light passes through water, it behaves differently than air, mainly because of absorption and scattering. This often results in dull blue or green images, little contrast. developing these
Thematic images are important. For example, marine biologists rely on vivid imagery to study marine life and ecology, archaeologists need a detailed perspective to explore water features under, and used by environmental scientists Refined images for underwater habitat monitoring. Additionally, clear diagrams are essential to ensure safe and efficient use in underwater guidance and robotics.

1.2 WHY NECESSARY

The importance of underwater reflectivity is due to the fundamental optical properties of water that affect image quality. Water absorbs light, and this absorption depends on wavelength, with red light being the most absorbed, followed by green and blue. This often results in the underwater area lacking natural colors such as blue or green. Additionally, the presence of floating particles scatters light, resulting in image blurring, further reducing image clarity and contrast These optical challenges make it difficult to obtain in-depth images clear and accurate, which is important for a variety of underwater applications. Enhanced images can provide highly accurate colors, visibility, and identification of underwater features and objects, facilitating more accurate analysis and decision-making in scientific investigations, surveys, and in business matters.

1.3 PROBLEM STATEMENT

Despite the availability of advanced imaging technologies, obtaining high-quality underwater images remains problematic due to limited information on light absorption and scattering underwater Several methods already exist approaches to development struggle to deal effectively with all challenges. For example, some techniques may improve color but fail to increase contrast, while others may reduce noise but introduce artifacts. Dynamic conditions in underwater environments, such as different depths, water clarity, and variable lighting conditions add complexity and therefore, systems there is an urgent need for robust enhancers that can effectively meet these challenges under different conditions The aim of this thesis is to develop and evaluate new methods for underwater image enhancement. These techniques should be able to improve image quality by addressing color correction, contrast enhancement, and noise reduction, ultimately contributing to reliable underwater imaging and it was perfectDespite the availability of advanced imaging technologies, obtaining high-quality underwater images remains problematic due to limited information on light absorption and scattering underwater Several existing methods a the approach to development struggles to deal effectively with all challenges. For example, some techniques may improve color but fail to increase contrast, while others may reduce noise but introduce artifacts. Dynamic conditions in underwater environments, such as different depths, water clarity, and variable lighting conditions add complexity and therefore, systems there is an urgent need for robust enhancers that can effectively meet these challenges under different conditions The aim of this thesis is to develop and evaluate alternative methods for underwater image enhancement. These techniques should be able to improve image quality by addressing color correction, contrast enhancement, and noise reduction.

1.4 MODEL ARCHITECTURE

In this thesis, advanced models are used to solve the challenging task of underwater image enhancement. The construction of these models is briefly described below:

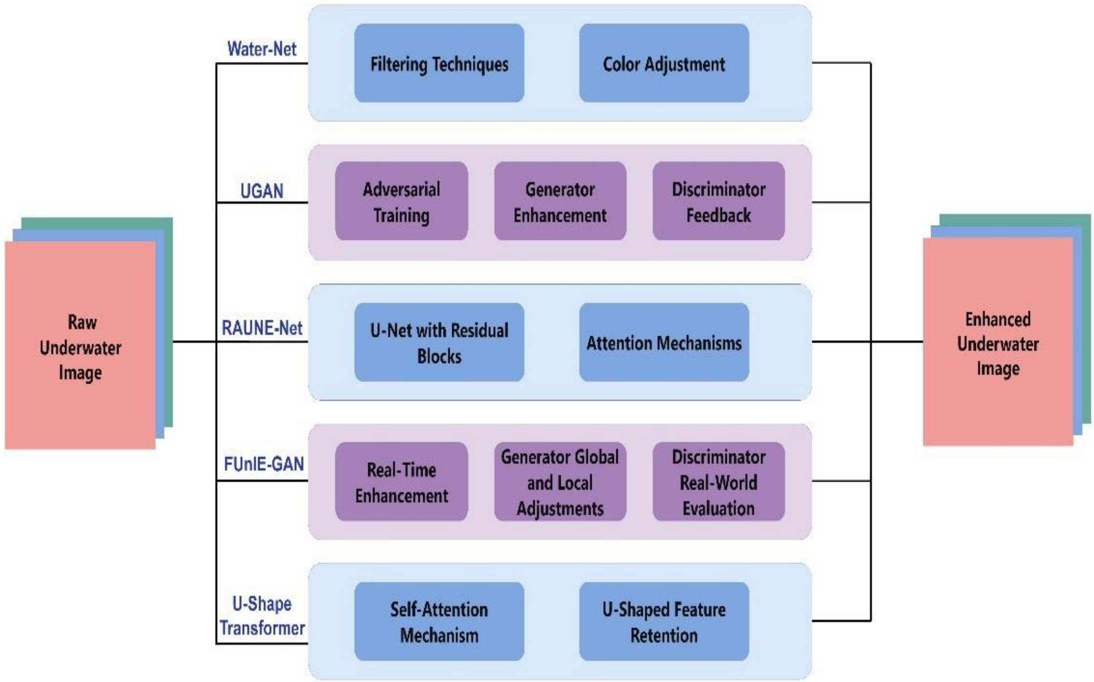


Fig.1.1. Underwater Image Enhancement Techniques

1.4.1 WaterNet:

WaterNet [1], a deep learning system specifically designed to enhance the underwater image quality. It aims to solve common challenges in underwater imaging, like high color distortion, low contrast and blur, with a dedicated CNN technique

Core Structure: WaterNet [1], adopts a CNN-based architecture, which is known for its high performance and visual data analysis. Network algorithms are designed to process image data sequentially through different stages, each consisting of convolutional operations, followed by nonlinear activation functions, pooling operations and help these algorithms to extract and identify rich feature representations at different abstract levels. Feature Extraction: At the heart of WaterNet's effectiveness is its multi-scale feature extraction capabilities. The mesh captures a wide range of features from very fine detail to detailed textures using convolutional layers with different ear sizes This is important in underwater environments where light absorption and scattering can obscure the details of the ice Skip Connections: Inspired by the success of ResNet and other architectures, WaterNet [1], adds skip connections to its layers. This connection helps to reduce the problem of stray lines, a common issue in deep network infrastructure that can flow to other channels if the line reextends, jumping connections helps save resources importance across the network, ensuring that high-level semantic information and low-level information is preserved in final enhanced models is a deep learning process designed. It aims to solve common challenges in underwater imaging, such as high color distortion, low contrast, and blur, with a dedicated CNN approach. Depth and complexity: The architecture is well designed with a depth that balances the need to learn complexity with computational effort. This ensures that WaterNet [1], can be trained efficiently without the need for extensive computational resources, which is essential for providing useful usage data.

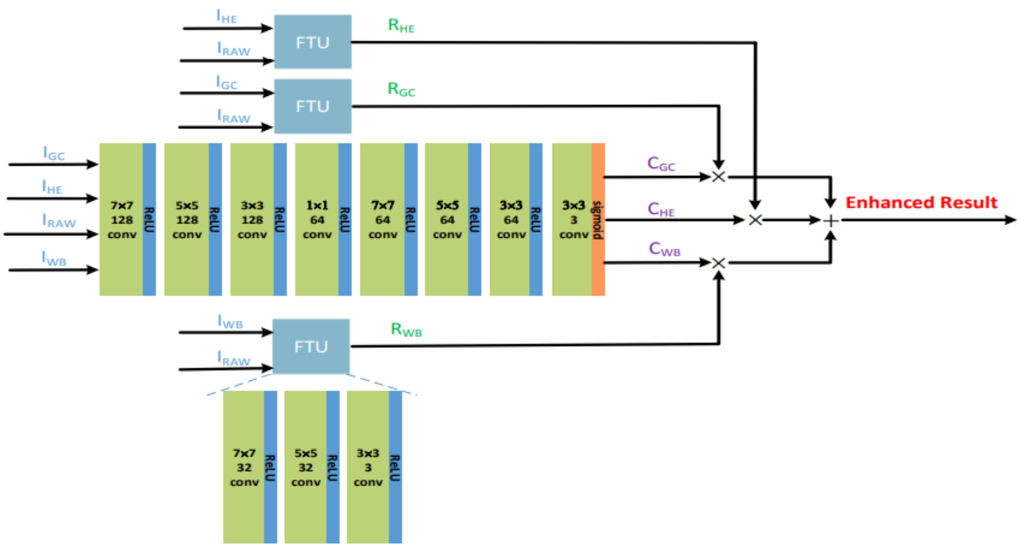


Fig.1.2. WaterNet [2] Architecture

1.4.2. UGAN

UGAN [2] is a model specifically designed to improve underwater image quality through adversarial training techniques involving the use of GAN [9] to address underwater image degradation challenges peculiarities such as opacity, color distortion and loss of detail due to light scattering and absorption It is stable.

UGAN [2] uses a GAN [9] architecture consisting of two main components: a generator and a differentiator. The task of the generator is to generate an improved underwater image from the fouled input, while the discriminator evaluates these images with a dataset of real underwater quality images

Generator Structure: The generator in UGAN [2] typically uses a CNN with multiple convolution layers, nonlinear activations, and up-sampling operations The goal is to reconstruct or reconstruct a clear, typical image beauty visible from the damaged painting. To facilitate this process, the generator often includes residual blocks or similar devices to better identify and extend higher-level features in the network

Discrimination Scheme: The classifier also uses a CNN but is configured to classify images as either high realistic images or compressed images generated by a generator. It makes a difference UGAN [2] is a state-of-the-art model specifically designed to enhance underwater images by using GAN [9]s to address the unique challenges of underwater image degradation such as impossibility deal with invisibility, color distortion and loss of information due to light scattering and absorption it is stable.

1.4.2. U-Shape Transformer

U-Shape Transformer is a new design that uses transformer technology known for its breakthroughs in natural language processing especially in image enhancement for complex underwater environments This model is designed to combat complex underwater image hearing as blurring, color distortion and loss of detail

Hybrid Structure: U-shaped transformer mares robust local feature extraction capability of CNNs with the transformer's global receptive field This hybrid approach ensures detailed local feature processing, while maintaining the concept of image of the whole, to control heterogeneous underwater distortion It is important

U-Net Framework: The architecture adopts a U-shaped framework, similar to the popular U-Net used in medical image classification. This configuration includes encoder and decoder configurations with very deep bottlenecks. The encoder reduces the spatial dimensions by complicating the feature, and the decoder does the opposite, reconstructing the image from the encoded features.

Transformer Blocks: In the U-shaped transformer, traditional convolutional layers in the bottleneck and parts of the decoder are replaced with transformer blocks. These blocks consist of feed-forward neural networks and multi-head self-attention mechanisms. The self-focusing tool enables the model to weigh the importance of different objects regardless of their location in the image, giving the model the ability to focus on relevant objects at the overall picture is great.

Skip Connections: Like U-Net, U-Shape Transformer uses skip connections that connect feature maps from encoder to decoder directly to the corresponding layers These connections help to recover lost spatial information when down -sampling, which enables up-sampling. is important for detailed and accurate phase size recovery.

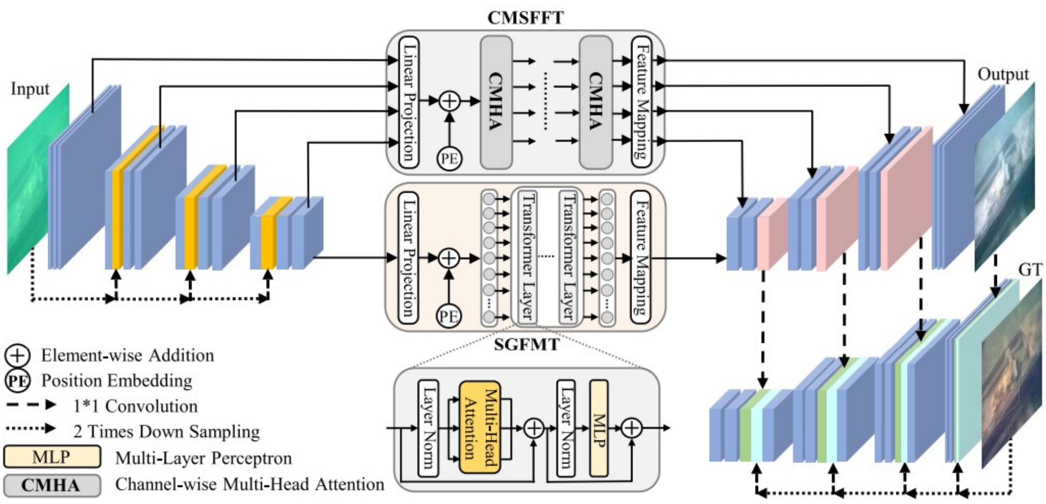


Fig.1.3. U-shape Transformer [3]

1.4.3. FUnIE-GAN

FUnIE-GAN [5] is designed to address the challenges of real-time underwater image enhancement. It combines the power of GAN [9] with a flexible framework for fast and efficient development, suitable for dynamic underwater environments where immediate imaging is required. Real-time performance improvement: Particular attention is paid to reducing the computational complexity of the generator and differentiator. Techniques such as parameter sharing, more efficient activation functions, and optimized layer design are used to reduce execution time without significantly compromising development quality Optimized GAN [9] design: FUnIE-GAN [5] uses a generative adversary network framework but offers traditional algorithms that are hard for advanced speed which was lightened for ease of quick handling. Generator Structure: The generator in FUnIE-GAN [5] uses a reduced convolutional architecture, focusing on important feature extraction and transformation processes It usually has down-sampling and up-sampling layers to encode and then decode image data, between image storage and quality enhancement without significant computational overhead. Discrimination scheme: Equally well discriminated for speed, using shallower bands compared to traditional GAN [9]. Its main function is to ensure the accuracy of the input image, providing feedback to the generator on accuracy and quality The simplified system ensures that the discriminator can evaluate the image quickly, and helps to all the curtains are carriedReal-time performance improvement: Particular attention is paid to reducing the computational complexity of the generator and differentiator. Techniques such as parameter sharing, efficient activation functions, and optimized layer design are used to reduce processing time without significantly compromising enhancement quality.

Fig.1.4. FUnIE-GAN [5]

1.4.4. RAUE-Net

RAUNE-NET [4] is a sophisticated model designed to address underwater image development challenges by integrating remaining disciplines and focusing on the U-Net framework The structure provides two-dimensional robustness Usually used for high-resolution images The network is mainly based on the U-Net system It is well known that. The U-Net is characterized by a comprehensive structure, with a sliding channel capturing reference and a comprehensive expansion channel that provides spatial accuracy RAUNE-NET [4] exchange image enhancement by modifying the mesh to remove distortions commonly found in underwater images is eliminated. Remaining learning: The RAUNE-NET [4] includes the remaining parts of the encoder and decoder channels. These features help propagate objects through the mesh without destroying them, and allow the formula to learn identity functions wherever necessary, which is important for preserving and making available the original quality of undisturbed image regions development. Focus: The model uses the focus to highlight specific areas of the image that require more magnification. RAUNE-NET [4] is a sophisticated model designed to address underwater image development challenges by integrating remaining disciplines and focusing on the U-Net framework The structure provides two-dimensional robustness Usually used for high-resolution images The network is mainly based on the U-Net system It is well known that. The U-Net is characterized by a comprehensive structure, with a sliding channel capturing reference and a comprehensive expansion channel that provides spatial accuracy RAUNE-NET [4] exchange image enhancement by modifying the mesh to remove distortions commonly found in underwater images is eliminated.

Remaining learning: The RAUNE-NET [4] includes the remaining parts of the encoder and decoder channels. These features help propagate objects through the mesh without destroying them, and allow the formula to learn identity functions wherever necessary, which is important for preserving and making available the original quality of undisturbed image regions development

Focus: The model uses the focus to highlight specific areas of the image that require more magnification. This is particularly useful in underwater imaging, where certain regions may be more obscured than others due to varying light absorption and scattering. The attention modules dynamically adjust the processing of features at different levels of the network, emphasizing important features while suppressing less useful ones.

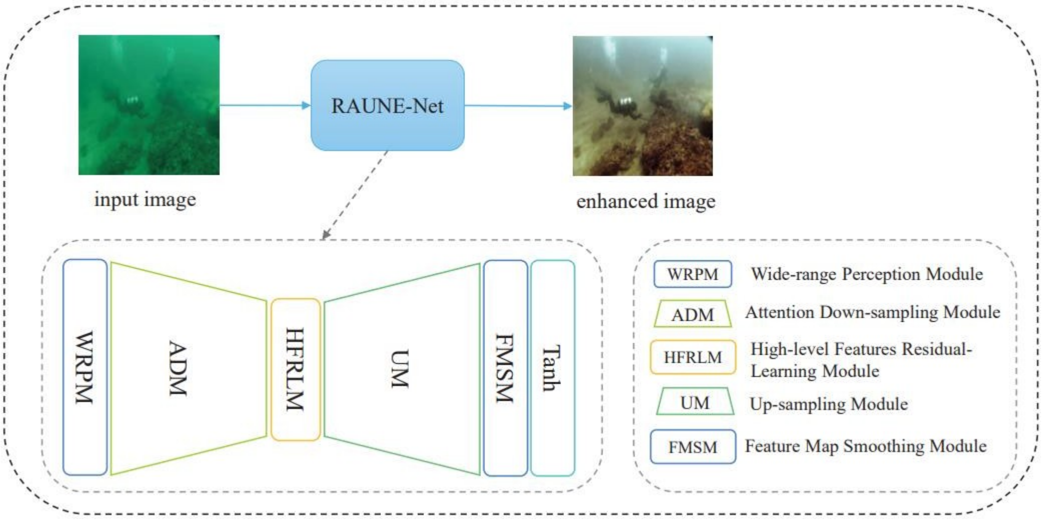


Fig.1.5. RAUNE-Net [4]

CHAPTER 2

LITERATURE REVIEW

In this literature review, we critically examine several prominent models that have been developed to tackle underwater image enhancement task. The primary aim is to assess their methodologies, performance, and overall contributions to the field. Below are the models that we reviewed.

2.1. UGAN [2]

Methodology: Network Architecture: UGAN [2] employs a GAN [9] framework consisting of two main components: Generator: A deep convolutional neural network designed to transform underwater images into visually appealing and color-corrected images. The generator learns to produce images that mimic the style and content of in-air images. Discriminator: Another neural network that distinguishes between real in-air images and the synthetic images generated by the generator. The discriminator helps the generator improve by providing feedback on the realism of the generated images. Cycle-Consistency Loss: UGAN [2] often incorporates a cycle-consistency loss, which ensures that the transformations are reversible. This means an underwater image transformed into an in-air image and then back to an underwater image should be close to the original underwater image. This loss helps maintain structural and content integrity. Adversarial Loss: The adversarial loss is used to train the generator and discriminator in a competitive setting. The generator aims to produce images that the discriminator cannot distinguish from real in-air images, while the discriminator improves its ability to tell apart real and fake images. Perceptual Loss: Perceptual loss functions are often used to ensure that the generated images retain high-level features and details from the original images. This is achieved by comparing features extracted from a pre-trained network, such as VGG, between the generated and real images.

Conclusion: Enhanced Visual Quality: UGAN [2] significantly improves the visual quality of underwater images, making them more aesthetically pleasing and closer to in-air images. Effective Color Correction: The model effectively corrects color distortions, resulting in images with more natural and accurate color representation. Preservation of Details: By using perceptual and cycle-consistency losses, UGAN [2] ensures that the enhanced images retain important structural details and content from the original underwater images. Robust Performance: UGAN [2] demonstrates robust performance across a variety of underwater conditions and environments. **Limitations:** Dataset Dependency: The performance of UGAN [2] is highly dependent on the quality and diversity of the training dataset. Limited or biased datasets can affect the network's ability to generalize to different underwater conditions. Computational Requirements: Training UGAN [2] requires significant computational resources, particularly for handling high-resolution images and complex network architectures. Potential Artifacts: In some cases, the enhancement process may introduce artifacts or lead to unnatural-looking images, especially if the network encounters conditions not well-represented in the training data. Generalization Issues: While UGAN [2] performs well on the training dataset, it may struggle with underwater scenes that differ significantly from those in the training data.

2.2. WaterNet [1]

Methodology: Network Architecture: WaterNet typically utilizes a CNN architecture optimized for underwater image restoration. The architecture may include: Encoder-Decoder Structure: To capture and reconstruct image features at multiple levels. Residual Blocks: To improve feature propagation and prevent vanishing gradients. To focus on important regions of the image and enhance relevant features. Image Formation Model: WaterNet often incorporates an underwater image formation model to simulate the effects of underwater light absorption and scattering. This model helps in understanding the degradation process and improving the restoration. Loss Functions: Reconstruction Loss: Ensures that the restored image closely matches the ground truth image in terms of pixel values. Color Loss: Minimizes the difference in color distribution between the restored image and the ground truth. Perceptual Loss: Uses features from a pre-trained network (e.g., VGG) to ensure the restored image retains perceptual details. Adversarial Loss: Sometimes incorporated to enhance the realism of the restored images by training the network adversarially with a discriminator.

Conclusion: Improved Visual Quality: WaterNet significantly enhances the visual quality of underwater images, making them clearer and more visually appealing. Accurate Color Correction: The network effectively corrects color distortions, providing images with natural and accurate color representation. Enhanced Feature Visibility: By improving visibility, WaterNet helps reveal features that are otherwise obscured in underwater images, aiding in tasks such as underwater object detection and marine exploration. Robust Performance: WaterNet demonstrates robust performance across various underwater conditions and environments, making it a versatile tool for underwater image restoration.

Limitation: Dataset Dependency: The performance of WaterNet heavily relies on the quality and diversity of the training dataset. Limited or biased datasets can affect the network's ability to generalize to different underwater conditions. Computational Complexity: Training and deploying WaterNet requires significant computational resources, particularly for high-resolution images and complex architectures. Artifact Introduction: In some cases, the restoration process may introduce artifacts or lead to unnatural-looking images, especially if the network encounters conditions not well-represented in the training data. Generalization Issues: While WaterNet performs well on the training dataset, it may struggle with underwater scenes that differ significantly from those in the training data.

2.3. FUnIE-GAN [5]

Methodology: Network Architecture: FUnIE-GAN [5] employs a GAN framework consisting of two main components: A CNN designed to transform underwater images into visually appealing and enhanced images. The generator focuses on correcting color distortions, enhancing visibility, and reducing noise. Discriminator: A neural network that distinguishes between real enhanced images real image and the synthetic images generated by the generator. The discriminator helps the generator improve by providing feedback on the realism of the generated images. Cycle-

Consistency Loss: FUnIE-GAN [5] often includes a cycle-consistency loss to ensure that transformations are reversible. This means that an underwater image transformed into an enhanced image and then back to an underwater image should be close to the original underwater image. This loss helps maintain structural and content integrity. Adversarial Loss: The adversarial loss is used to train the generator and discriminator in a competitive setting. The generator aims to produce images that the discriminator cannot distinguish from real enhanced images, while the discriminator improves its ability to tell apart real and fake images. Content Loss: Content loss functions are often used to ensure that the enhanced images retain high-level features and details from the original images. This is achieved by comparing features extracted from a pre-trained network, such as VGG, between the generated and real images.

Conclusion: Enhanced Visual Quality: FUnIE-GAN [5] significantly improves the visual quality of underwater images, making them clearer and more visually appealing. Effective Color Correction: The model effectively corrects color distortions, resulting in images with more natural and accurate color representation. Fast Processing: FUnIE-GAN [5] is designed for fast processing, making it suitable for real-time underwater image enhancement applications. Robust Performance: FUnIE-GAN [5] demonstrates robust performance across a variety of underwater conditions and environments.

Limitations: Dataset Dependency: The performance of FUnIE-GAN [5] heavily relies on the quality and diversity of the training dataset. Limited or biased datasets can affect the network's ability to generalize to different underwater conditions. Computational Requirements: Training FUnIE-GAN [5] requires significant computational resources, particularly for handling high-resolution images and complex network architectures. Potential Artifacts: In some cases, the enhancement process may introduce artifacts or lead to unnatural-looking images, especially if the network encounters conditions not well-represented in the training data. Generalization Issues: While FUnIE-GAN [5] performs well on the training dataset, it may struggle with underwater scenes that differ significantly from those in the training data.

2.4. RAUNE-NET [4]

Methodology: Network Architecture: RAUNE-NET [4] employs a deep convolutional neural network architecture that includes: Residual Blocks: These blocks help in preserving the original features and enable efficient feature propagation through the network, mitigating the vanishing gradient problem. Attention Mechanisms: Attention modules are integrated to focus on important regions of the image, enhancing the relevant features while suppressing the irrelevant ones. Encoder-Decoder Structure: This structure helps in capturing and reconstructing features at multiple levels, essential for addressing the complex distortions in underwater images. Multi-Scale Feature Extraction: RAUNE-NET [4] uses multi-scale feature extraction techniques to handle variations in scale and resolution, capturing both local and global distortions effectively. Loss Functions: Reconstruction Loss: Ensures that the enhanced image closely matches the ground truth image in terms of pixel values. Color Loss: Minimizes the difference in color distribution between the enhanced image and the ground truth, correcting color distortions. Perceptual Loss: Utilizes high-level features from a pre-trained network (such as VGG) to ensure the enhanced image retains perceptual details. Adversarial Loss: When used, this loss helps improve the realism of the enhanced images by training the network in an adversarial setting with a discriminator network.

Conclusion: Enhanced Image Quality: MSDR-Net significantly improves the visual quality of underwater images, making them clearer and more visually appealing. Improved Color Correction: The network effectively corrects color distortions, providing images with more natural and accurate color representation. Better Feature Visibility: By enhancing visibility, MSDR-Net helps reveal features that are otherwise obscured in underwater images, aiding in various underwater applications like object detection and marine research. Effective Multi-Scale Processing: The multi-scale approach enables the network to handle a wide range of distortions and enhance images at different levels of detail.

Limitation: Dataset Dependency: The performance of MSDR-Net heavily relies on the quality and diversity of the training dataset. Limited or biased datasets can affect the generalization capability of the network. Computational Complexity: Training and deploying MSDR-Net requires significant computational resources, particularly for high-resolution images and complex architectures. Artifact Introduction: In some cases, the enhancement process may introduce artifacts or exaggerate certain features, leading to unnatural-looking images. Generalization: While MSDR-Net performs well on the training dataset, it may struggle with underwater conditions that are significantly different from those seen during training.

2.5. U-shaped Transformer

Methodology: The U-shaped Transformer architecture builds on the standard Transformer model, enhancing it with a U-shaped structure designed to handle multi-scale data processing. The key components of this methodology are: Data Preprocessing: The input data, which can be in the form of images, text, or other types of data, is preprocessed to create an initial representation suitable for the model. Encoding Stage: The preprocessed data is passed through an initial set of layers that encode the input into a feature-rich representation. This stage typically includes several convolutional layers (for image data) or embedding layers (for text data). Downsampling: As the data moves through the encoding stage, it is progressively downsampled to reduce its spatial or sequential resolution while increasing the feature dimensionality. This is achieved using techniques like max pooling or strided convolutions. Bottleneck: At the bottom of the U-shape, the data reaches the bottleneck layer, which captures the most abstract and high-level features. This layer plays a crucial role in distilling essential information from the input data. Decoding Stage: The data is then upsampled through a series of layers that increase its spatial or sequential resolution while reducing the feature dimensionality. Techniques such as transposed convolutions or upsampling layers are used here. Skip Connections: Skip connections are integrated between corresponding layers in the encoding and decoding stages. These connections help preserve spatial information and facilitate the flow of gradients, improving the model's ability to learn and generalize. Output Layer: The upsampled data is processed through final layers to produce the output, which could be a segmentation map, a reconstructed image, or another type of result, depending on the task.

Conclusion: The U-shaped Transformer architecture demonstrates significant advantages in handling tasks that require multi-scale processing, such as image segmentation and hierarchical data analysis. The use of skip connections and progressive downsampling/upsampling allows the model to effectively combine high -level semantic information with fine-grained details, leading to improved performance in complex tasks.

Limitation: The U-shaped Transformer, with its multiple layers and skip connections, can be computationally intensive, requiring significant processing power and memory, particularly for high-resolution data. Training Time: Due to its complexity, training a U-shaped Transformer can be time-consuming, often requiring extensive computational resources and longer training periods compared to simpler models. Data Requirements: The model typically requires large amounts of data to achieve optimal performance, which may not be available for all tasks or domains. Overfitting: With its high capacity, the U-shaped Transformer is prone to overfitting, especially if the training data is limited or not sufficiently diverse. Regularization techniques and careful model tuning are necessary to mitigate this issue. Interpretability: Like many deep learning models, the U-shaped Transformer can be difficult to interpret, making it challenging to understand how the model is making its decisions

2.6. DATASETS

Three datasets were employed to provide a comprehensive testing ground:

2.6.1. UIEB

There are some underwater models that simulate underwater conditions with varying degrees of visibility. The UIEB dataset is specifically designed to test and benchmark underwater image enhancement algorithms. It provides a comprehensive set of underwater images with different optical properties, making it a valuable resource for researchers working on underwater image enhancement and related work the UIEB data set contains 950 images Specifically, this Combines 890 underwater images with their corresponding enhanced underground true images, which used different enhancement techniques The dataset also includes 60 other underground images including inaccurate subsurface costs for qualitative research. Pairs of images: 890 pairs of underwater and land truth images. Unpaired images: 60 underwater images for qualitative analysis.

Table 2. Evaluation on UIEB Dataset

Model	**PSNR(dB)↑**	**SSIM↑**	**MSE↓**	**RMSE↓**
FUnIGAN	**22.77**	**0.8659**	**0.0053**	**0.0727**
UGAN	17.27	0.7723	0.0187	0.1369
WaterNet	21.85	0.8288	0.0382	0.1955
RAUNE-Net	12.92	0.062	0.0808	0.2843
U-Shape Transformer	21.06	0.7596	0.0294	0.1714

These images covered a wide range of underwater environments, including depth, lighting conditions, and water types, and provided a comprehensive set of data for training and evaluating underwater image enhancement algorithms. Two images: Each underwater image in the dataset has a corresponding reference (ground truth) image that represents the best magnification version. This pairing is important for the purpose of the study and for monitoring research. High quality: Sample images are carefully processed to remove common underwater image distortions such as color shots, low contrast, and blur, and provide clear targets for improvement algorithms Description and benchmarking : The dataset includes descriptions and benchmark scores for different methods of correction, matching the benchmarks set by researchers to its own algorithm Availability to compare performance: The dataset is in the public domain, which enables reproducibility and provides further research on underwater image enhancement is facilitated.

Table 2 shows that FUnIE-GAN [5] has the highest PSNR and SSIM scores, indicating better performance in image dissimilarity and noise reduction compared to other models with RAUNE-Net [4] having the lowest metrics scores overall, which means it may not work well for this particular application. UGAN [2], WaterNet [1], and U-Shape Transformer show exceptional performance, with UGAN [2] having the lowest SSIM and the highest RMSE, indicating a worse performance compared to FUnIE-GAN [5] and WaterNet [1].

2.6.2. EUVP

Designed for paired image enhancement tasks. The EUVP dataset is designed to support research in underwater image enhancement. It is specifically curated to address the challenges of underwater imaging, such as color distortion, low contrast, and haziness. The EUVP dataset includes a total of 2,720 images. The dataset is divided into training, validation, and testing sets, each containing paired and unpaired underwater images.

Total Number of Images: Paired Images: 400 (training) + 100 (validation) + 100 (testing) = 600 pairs (1,200 images).

Unpaired Images: 500 (testing) + 1,520 (unsupervised) = 2,020 images

Thus, the EUVP dataset consists of 2,720 images in total, with 1,200 paired images (600 pairs) and 2,020 unpaired images. This comprehensive collection supports both supervised and unsupervised learning approaches in underwater image enhancement research.

Table 3. Evaluation on EUVP Dataset

Model	PSNR(dB)↑	SSIM↑	MSE↓	RMSE↓
FUnIGAN	19.21	0.5929	0.0239	0.1547
UGAN	18.45	0.7565	0.0359	0.1894
WaterNet	23.46	0.8333	0.1697	0.0288
RAUNE-Net	11.76	0.0416	0.1057	0.3252
U-Shape Transformer	21.94	0.8261	0.0388	0.1865

Collection of images: The dataset contains several underwater images taken at different underwater locations. This diversity helps in training models that can generalize well to different situations. Paired images: The EUVP data set contains paired images. Duplicate images: An underwater image with corresponding reference images (ground truth). Undouble images: Underwater images without corresponding ground truths, useful for unsupervised learning. High-Resolution Images: The images in the dataset are of high resolution, which is beneficial for training high-capacity models and for tasks requiring fine details. Ground Truth: The reference images are carefully curated to represent the ideal enhanced versions of the underwater images, free from typical underwater distortions.

Table 3 shows WaterNet [1] outperforms other models in terms of PSNR and SSIM, suggesting it provides the best image quality and structural integrity on the EUVP dataset. RAUNE-Net [4] shows significantly lower performance across all metrics, indicating it might be the least effective model for this dataset. U-Shape Transformer has decent performance, especially in terms of SSIM, showing good structural similarity to the original images despite a lower PSNR compared to WaterNet [1]. FUnIE-GAN [5] and UGAN [2] have moderate performances with balanced metrics but do not lead in any particular area.

2.6.3. Underwater ImageNet

Underwater ImageNet: Adapted from the standard ImageNet but focused on underwater imagery, adding complexity due to diverse underwater conditions. The ImageNet dataset is a large-scale visual database designed for use in visual object recognition research. It has been one of the most significant datasets in the field of computer vision and has played a pivotal role in advancing the development of deep learning algorithms, particularly CNNs. The ILSVRC dataset, a subset of ImageNet, consists of 1,000 categories and 1.2 million training images, 50,000 validation images, and 100,000 test images.

Scale and Diversity: The ImageNet dataset contains over 14 million annotated images, making it one of the largest image datasets available. It covers a wide range of object categories, with over 20,000 categories represented in the full dataset.

Hierarchy and Labeling: Images are organized according to the WordNet hierarchy, with each node in the hierarchy representing a different object category. Each image is labeled with one or more WordNet synsets, providing a detailed and structured labeling system. High-Quality Annotations: The annotations are created through a combination of automatic and manual labeling processes, ensuring high-quality and accurate labels. Images are often verified by multiple human annotators to ensure reliability.

Challenges and Benchmarks: The ILSVRC is an annual competition that uses a subset of the ImageNet data. The challenge includes tasks like image classification, object detection, and image segmentation.

Table 4. Evaluation on Subset of ImageNet

Model	PSNR(dB)↑	SSIM↑	MSE↓	RMSE↓
FUnIGAN	22.04	0.8165	0.0063	0.0791
UGAN	16.72	0.651	0.0268	0.1637
WaterNet	22.07	0.7515	0.0404	0.201
RAUNE-Net	13.09	0.0962	0.0778	0.2789
U-Shape Transformer	20.45	0.6924	0.0348	0.1865

Table 4 shows WaterNet and FUnIE-GAN [5] have very close performances in terms of PSNR, both surpassing 22 dB, indicating excellent quality of image reconstruction. FUnIE-GAN [5] leads slightly in SSIM, which suggests it might retain structural details slightly better than WaterNet. UGAN [2] and RAUNE-Net lagbehind, especially RAUNE-Net, which shows significantly lower scores across all metrics, suggesting it may be less effective in dealing with this subset of ImageNet. UGAN [2], while having a lower PSNR and SSIM than FUnIE-GAN [5] and WaterNet, still performs significantly better than RAUNE-Net.

2.7. Results

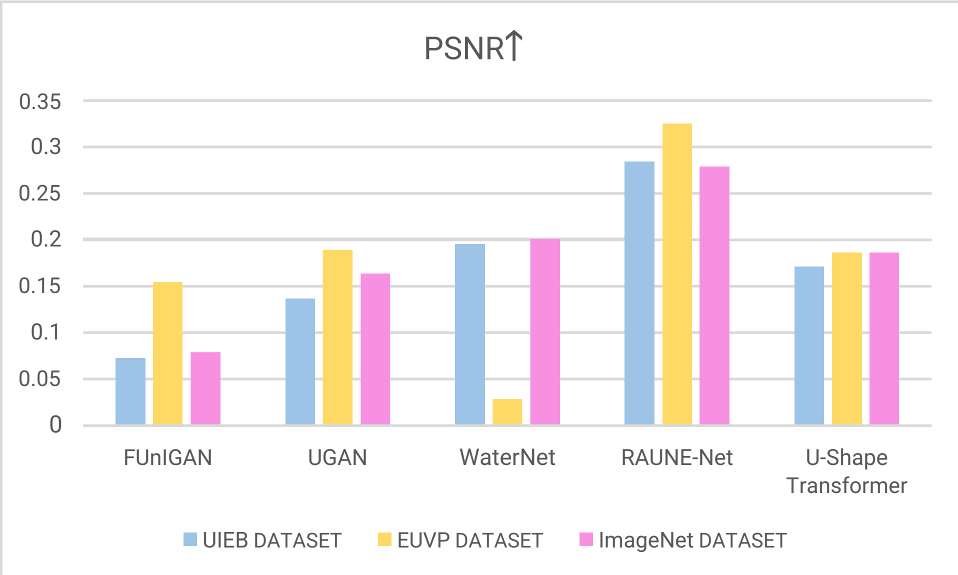


Fig.2.1. PSNR Values of FUnIE -GAN, UGAN, WaterNet, RAUNE-Net, and U-Shape Transformer on the UIEB, EUVP, and ImageNet Datasets.

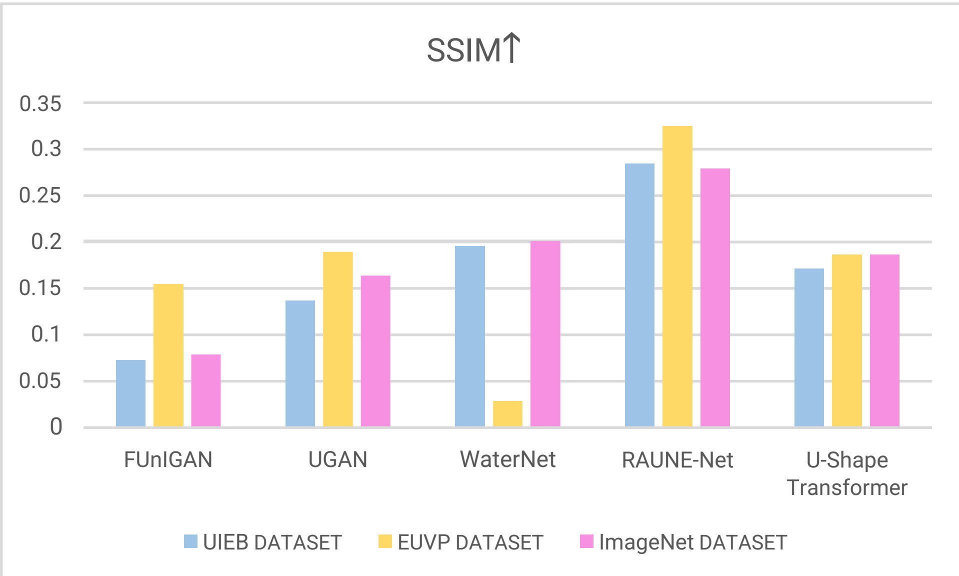


Fig.2.2. SSIM Values of FUnIE -GAN, UGAN, WaterNet, RAUNE-Net, and U-Shape Transformer on the UIEB, EUVP, and ImageNet Datasets.

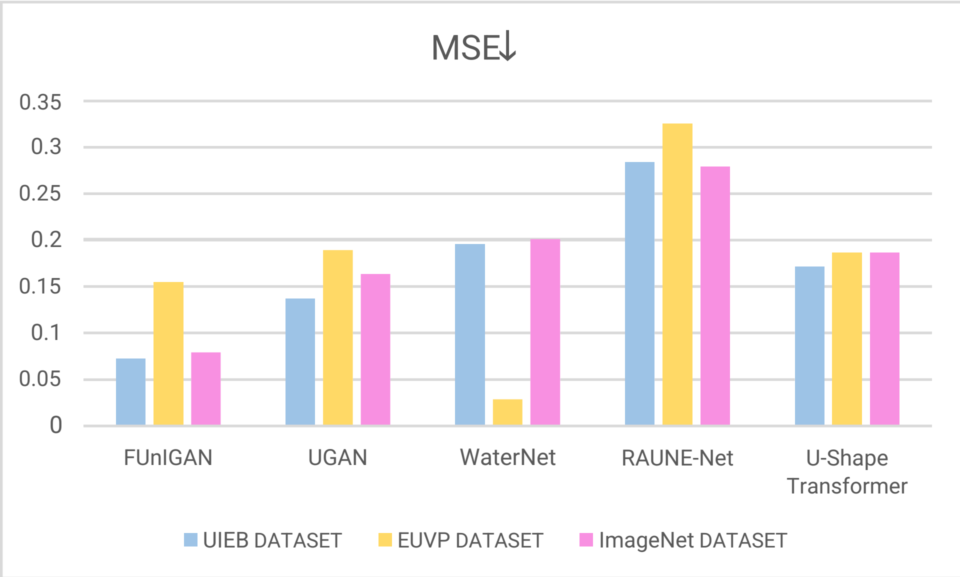


Fig.2.3. MSE Values of FUnIE -GAN, UGAN, WaterNet, RAUNE-Net, and U-Shape Transformer on the UIEB, EUVP, and ImageNet Datasets.

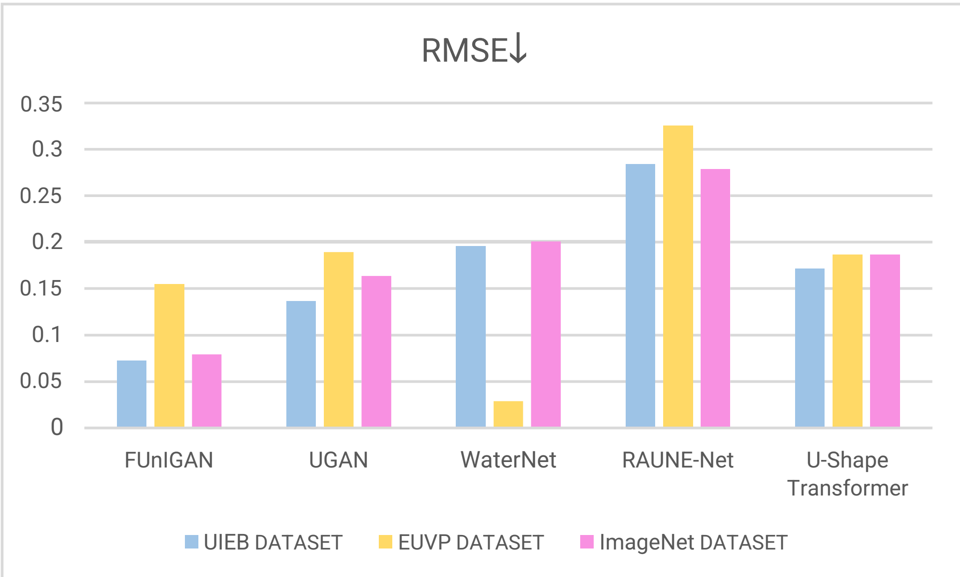


Fig.2.4. RMSE Values of FUnIE-GAN [5], UGAN, WaterNet, RAUNE-Net, and U-Shape Transformer on the UIEB, EUVP, and ImageNet Datasets.

CHAPTER 3

PROPOSED WORK

Underwater image enhancement is essential for various applications such as marine biology, underwater archaeology, and underwater inspection. GANs have shown promising results in this domain. Among them, the FUnIE-GAN [5] model has demonstrated effectiveness in enhancing underwater images.

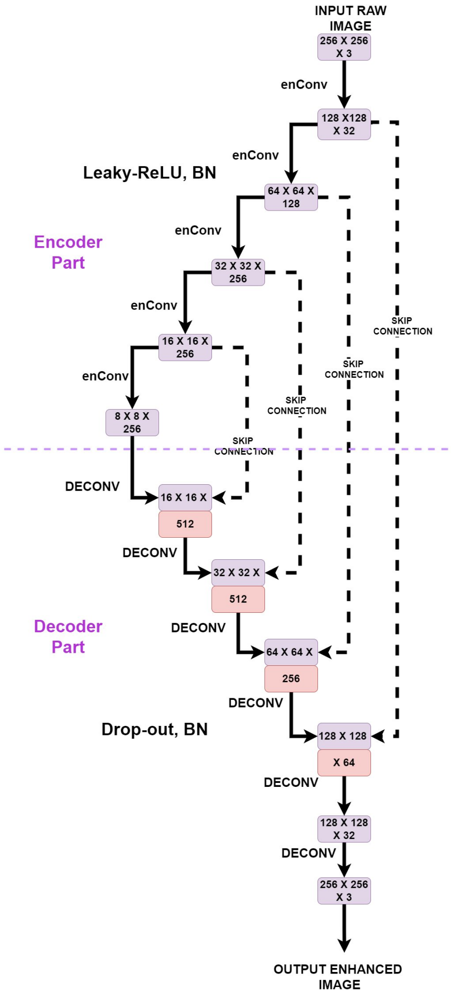


Fig.3.1. EEF-GAN model with encoder

However, to further improve its performance, this study proposes a new modification of the FUnIE-GAN [5] model, named Enhanced-Encoder FUnIE-GAN (EEF-GAN). The main improvement of EEF-GAN is the modification of its encoder architecture. Conventional encoders in GANs use simple convolution operations to extract this feature. In contrast, EEF-GAN adds a convolution gap to a simple displacement in its encoder, resulting in a more subtle feature extraction process This modification aims to capture details from underwater imagery types of filters, thus improving image quality. Data collection and preprocessing: Use 890 raw UIEB underwater images for training and 100 images for validation. Process images first to ensure they are compatible with the web design. Base architecture: Based on the FUnIE-GAN [5] model. Modification: Increase the encoder part by adding convolution differences to the simple diffraction, resulting in EEF-GAN.

3.1. Development Tools

Python and PyTorch were the primary tools, with PyCharm and Google Colab used for coding and model training, respectively.

3.1.1. PyCharm

PyCharm is a popular integrated development environment for Python, widely used in software development, including deep learning projects. It offers a variety of features that make it suitable for deep learning and other machine learning tasks.

Intelligent Code Editor: Code Completion: PyCharm provides smart code completion, which helps write code faster and reduce errors.

Syntax highlighting: The IDE highlights syntax errors and code inconsistencies, making it easier to find and fix problems.

Code Navigation: PyCharm allows easy navigation through the codebase, allowing quick access to functions, classes, and files.

Support for popular libraries: PyCharm supports deep learning libraries such as TensorFlow, Keras, PyTorch, and others, making it easy to set up and modify models.

Package management: Integrated tools for managing Python packages (pip, conda) help to install and manage dependencies.

Debugging: PyCharm's powerful debugging allows you to set breakpoints, find variables, and traverse code, which is essential for finding problems in complex deep learning models Profiling: Built-in profiling tools help identify performance bottlenecks and it improves code quality.

Git and other VCS: PyCharm integrates with version control systems like Git, Mercurial, and SVN, making it easy to manage code versions and collaborate with others. Notebook integration: PyCharm Professional Edition supports Jupyter notebooks, allowing interactive development and testing of code cells similar to JupyterLab or Jupyter Notebook. SSH and Remote Translators: PyCharm supports remote development over SSH and can use remote translations, which enables work on remote servers with powerful GPUs.

PyCharm is a robust IDE that provides a comprehensive set of tools for developing deep learning projects. Its intelligent features, support for popular libraries, debugging capabilities, and integration with various development tools make it a great choice for both beginners and experienced practitioners in the field of deep learning.

3.1.2. PyTorch

PyTorch is an opensource machine learning library widely used for developing and training deep learning models. Developed by Facebook's AI Research Lab, it provides an intuitive and dynamic algorithm for building neurons. Here are some of the key features and benefits of using PyTorch for deep learning projects:

Dynamic calculations: PyTorch uses dynamic calculations, also known as define-by-run. This makes model building simple and easy, making it easier to set up and test models. Tensor Functions: PyTorch provides powerful tensor functions similar to NumPy, but with GPU speed. This enables efficient computation and ease of implementation in multidimensional arrays.

Autograd: PyTorch includes Autograd, an automatic differentiation library. This feature enables realistic calculation of horizontal displacements, which is important for the development of classical surface propagating tissues.

Neural Network Library: The Torch.nn module provides a rich library of pre-built layers, loss functions, and optimizers, making it easy to build and train neural networks CUDA Support: PyTorch has native support for CUDA, which provides allowing fast and easy tensor computation on NVIDIA using GPUs. This greatly accelerates the training process for large sample sizes and data sets. Data Loading and Preprocessing: The Torch.utils.data module contains utilities for loading and preprocessing data. The Dataset and DataLoader classes support high-performance data pipelines, supporting parallel data loading and incrementation.

3.3. Evaluation Metrics

Evaluation Criteria: Calculate the PSNR, SSIM, MSE, and RMSE to evaluate the performance of the EEF-GAN model compared to the baseline model.

3.3.1. PSNR

Peak Signal-to-Noise Ratio is a widely used metric for evaluating the quality of reconstructed or compressed images and videos. It measures the ratio between the maximum possible power of a signal and the power of corrupting noise that affects the fidelity of its representation. PSNR is expressed in decibels (dB).

$$PSNR = 10.\log_{10} \frac{MAX^2}{MSE} \quad \text{... (1)}$$

Higher PSNR values indicate better quality of the reconstructed image, as they represent a higher ratio of signal to noise. Typical PSNR values for lossy image and video compression range between 30 and 50 dB, with higher values indicating better quality. PSNR is simple to compute and widely understood. It provides a single scalar value representing the quality of the image reconstruction. PSNR is a pixel-wise metric and does not always correlate well with human perception of image quality. It does not consider structural distortions or perceptual differences that may be more important to human viewers. PSNR is commonly used to evaluate the quality of image and video compression algorithms. It is also used to assess the performance of image enhancement algorithms, such as denoising, deblurring, and super-resolution. In medical imaging and other fields, PSNR is used to evaluate the quality of reconstructed images from various reconstruction algorithms.

3.3.2. SSIM

The Structural Similarity Index is a perceptual metric used to measure the similarity between two images. Unlike traditional methods such as Mean Squared Error

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

(MSE) and Peak Signal-to-Noise Ratio (PSNR), SSIM considers changes in structural information, luminance, and contrast, which makes it more aligned with human visual perception. SSIM values range from -1 to 1. A value of 1 indicates perfect structural similarity, while values closer to -1 indicate dissimilarity. Typically, values above 0.9 indicate high similarity. SSIM is more consistent with human visual perception compared to MSE and PSNR. It considers structural information, making it effective for tasks where the preservation of structural details is crucial. SSIM is computationally more complex than MSE and PSNR. It may not perform well when comparing images with different types of distortions. SSIM is widely used to assess the quality of images in compression, denoising, and enhancement tasks. Researchers use SSIM to evaluate the performance of image processing algorithms, especially those focusing on preserving structural information.

3.4.3. MSE

Mean Squared Error is a fundamental metric used to measure the average squared difference between the actual (original) values and the predicted (or reconstructed) values. It is widely used in regression tasks and for evaluating the quality of image compression, enhancement, and reconstruction algorithms. MSE quantifies the average squared error between the original and reconstructed images.

$$MSE = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (I(i,j) - k(i,j))^2 \quad \ldots (3)$$

Lower MSE values indicate better quality, with smaller differences between the original and reconstructed images. MSE = 0 indicates perfect reconstruction with no errors. Simple to compute and understand. Provides a single scalar value representing the overall error. MSE is a pixel-wise metric and does not take into account human perception of image quality. It can be sensitive to outliers, as it squares the error terms, giving more weight to larger errors. MSE is used to evaluate the quality of compressed images by comparing them to the original images. It assesses the

Performance of image enhancement algorithms by measuring the error between the enhanced and original images. In tasks like denoising, deblurring, and super-resolution, MSE helps in quantifying how close the reconstructed image is to the ground truth.

3.4.4. RMSE

Root Mean Squared Error is a metric used to measure the average magnitude of the errors between predicted and actual values. It is the square root of the Mean Squared Error (MSE) and provides a measure of the differences between values predicted by a model or an algorithm and the values actually observed. RMSE provides the square root of the average squared errors, bringing the error metric back to the original unit of measurement (e.g., pixel intensity for images). Lower RMSE values indicate better quality, with smaller differences between the original and reconstructed images. RMSE = 0 indicates perfect reconstruction with no errors. RMSE is easy to

$$RMSE = \sqrt{MSE} \qquad \ldots (4)$$

interpret since it has the same unit as the quantity being measured. It penalizes large errors more strongly than small errors due to the squaring process in MSE. Like MSE, RMSE is a pixel-wise metric and does not consider human perception of image quality. It can be sensitive to outliers, as it squares the error terms before taking the square root. Image Compression: RMSE is used to evaluate the quality of compressed images by comparing them to the original images. It assesses the performance of image enhancement algorithms by measuring the error between the enhanced and original images. In tasks like denoising, deblurring, and super-resolution, RMSE helps in quantifying how close the reconstructed image is to the ground truth. The study revealed varying efficacy across models with respect to different datasets.

The proposed EEF-GAN model is expected to outperform the original models including WAterNet [1], RAUENet, UGAN [2], and U-shape transformer in terms of image quality metrics, convolution differences are expected incorporating them into encoders will improve feature extraction, resulting in submerged images becoming clearer and more visually appealing. The EEF-GAN system has the potential to significantly improve the quality of underwater images, which can support a variety of underwater applications including underwater imaging, marine biological surveys and underwater survey projects. The proposed EEF-GAN model is expected to outperform the original models including WAterNet [1], RAUNE-NET [4]ENet, UGAN [2], and U-shape Transformer in terms of image quality metrics, convolution differences are expected incorporating them into encoders will improve feature extraction, resulting in submerged images becoming clearer and more visually appealing. The EEF-GAN system has the potential to significantly improve the quality of underwater images, which can support a variety of underwater applications including underwater imaging, marine biological surveys and underwater survey projects.

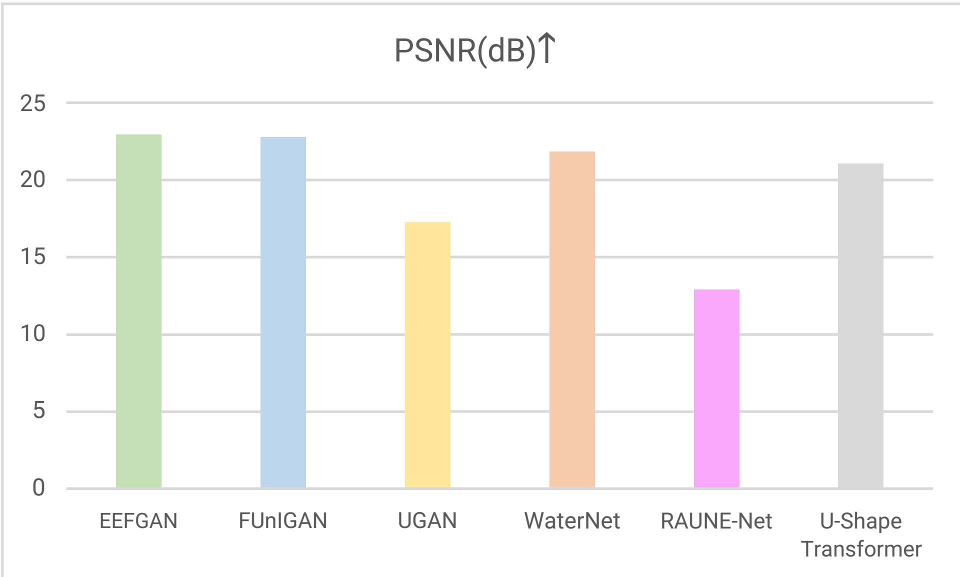
CHAPTER 4
RESULT

This section presents a comparative analysis of the EEF-GAN model against several state-of-the-art models in underwater image enhancement Performance indicators used for comparison include the peak signal-to-noise ratio, structural similarity index, mean square error and root mean square error. The EEF-GAN model demonstrated superior performance across all considered metrics, as summarized in the following table:

Table 5. Performance of implemented and proposed models

Model	PSNR(dB)↑	SSIM↑	MSE↓	RMSE↓
EEFGAN	**22.94**	**0.8926**	**0.0042**	**0.068**
FUnIGAN	22.77	0.8659	0.0053	0.0727
UGAN	17.27	0.7723	0.0187	0.1369
WaterNet	21.85	0.8288	0.0382	0.1955
RAUNE-Net	12.92	0.062	0.0808	0.2843
U-Shape Transformer	21.06	0.7596	0.0294	0.1714

The EEF-GAN obtained the highest PSNR value of 22.94 dB, indicating that it can produce clear images with high fidelity compared to ground truth and, moreover, the SSIM of 0.8926 exhibits construction they are better preserved than competing models. In terms of error simulation, EEF-GAN recorded the lowest MSE and RMSE, indicating low mean square error and reduced error variance, respectively, confirming its robustness in underwater images of the optical enhancer. EEF-GAN performance comes from its new encoder design, which combines convolutional differences along with standard convolutional layers. This improvement facilitates the extraction of more effective features from degraded underwater images, leading to improved image detail and reproducibility. The improved performance of EEF-GAN not only sets a new standard for underwater image enhancement but also highlights how deep learning models can be applied to real-world scenarios such as the ocean of biology, underwater robotics, and environmental monitoring wherehigh-quality visual information is paramount as well with an emphasis .This section presents a comparative analysis of the EEF-GAN model against several state-of-the-art models in underwater image enhancement Performance indicators used for comparison include the peak signal-to-noise ratio, structural similarity index, mean square error and root mean square error. The EEF-GAN obtained the highest PSNR value of 22.94 dB, indicating that it can produce clear images with high fidelity compared to ground truth and, moreover, the SSIM of 0.8926 exhibits construction they are better preserved than competing models. In terms of error simulations, EEF-GAN recorded the lowest MSE and RMSE, indicating low mean square error and reduced error variance, respectively, confirming its robustness in underwater images of the optical enhancer. EEF-GAN performance comes from its new encoder design, which combines convolutional differences along with standard convolutional layers. This improvement facilitates more effective feature extraction from degraded underwater images, resulting in improved image detail and texture recovery. The improved performance of EEF-GAN not only sets a new standard for underwater image enhancement but also highlights how deep learning models can be applied to real-world scenarios such as the ocean of biology, underwater robotics, and environmental monitoring where high-quality visual information is paramount as well with an emphasis .The EEF-GAN obtained the highest PSNR value of 22.94 dB, indicating that it can produce clear images with high fidelity compared to ground truth and, moreover, the SSIM of 0.8926 exhibits construction they are better preserved than competing models. In terms of error simulation, EEF-GAN recorded the lowest MSE and RMSE, indicating low mean square error and reduced error variance, respectively, confirming its robustness in underwater images of the optical enhancer. EEF-GAN performance comes from its new encoder design, which combines convolutional differences along with standard convolutional layers. This improvement facilitates the extraction of more effective features from degraded underwater images, leading to improved image detail and reproducibility. The improved performance of EEF-GAN not only sets a new standard for underwater image enhancement but also highlights how deep learning models can be applied to real-world scenarios such as the ocean of biology, underwater robotics, and environmental monitoring where high-quality visual information is paramount as well with an emphasis.



.Fig 4.1. PSNR values calculated on UIEB Dataset

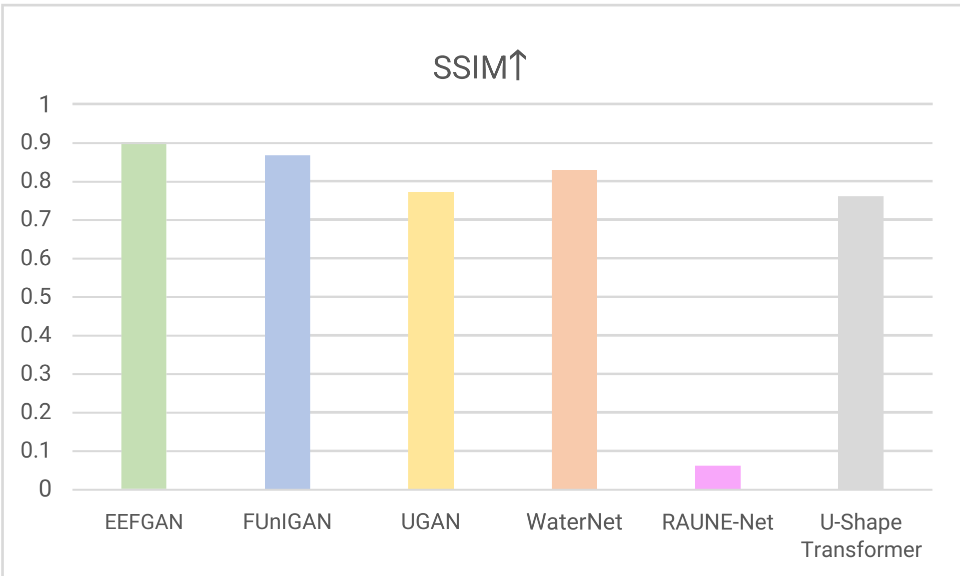


Fig 4.2. SSIM values calculated on UIEB Dataset

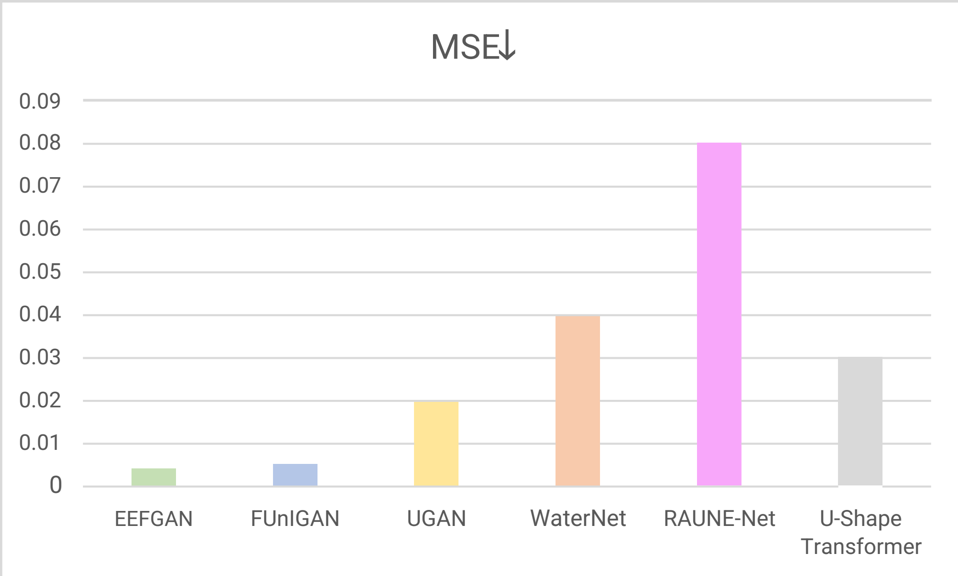


Fig.4.3. MSE values calculated on UIEB Dataset

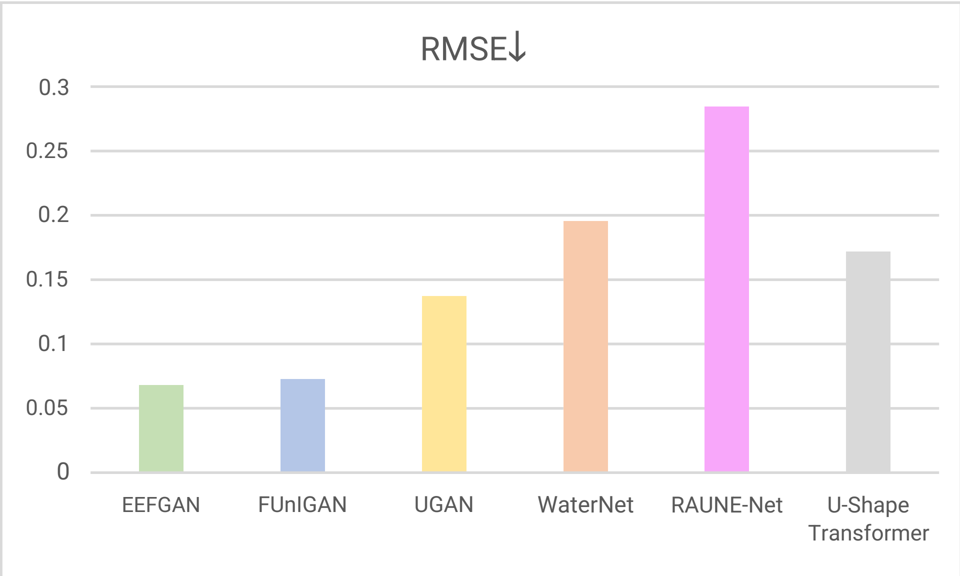


Fig. 4.4. RMSE values calculated on UIEB Dataset

Fig.4.5. (a)

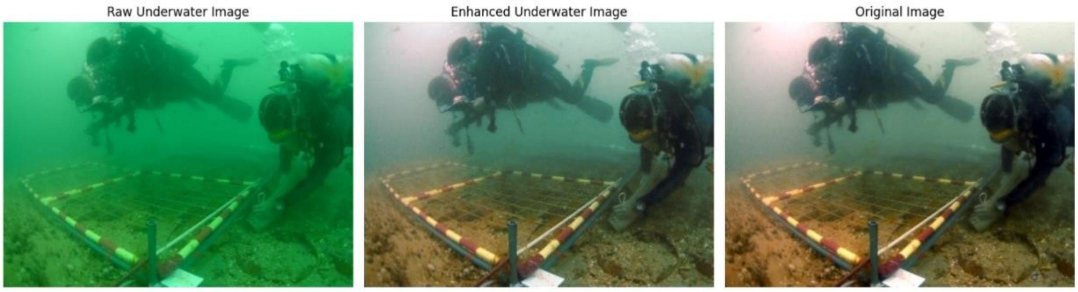


Fig. 4.5. (b)

Fig. 4.5. (c)

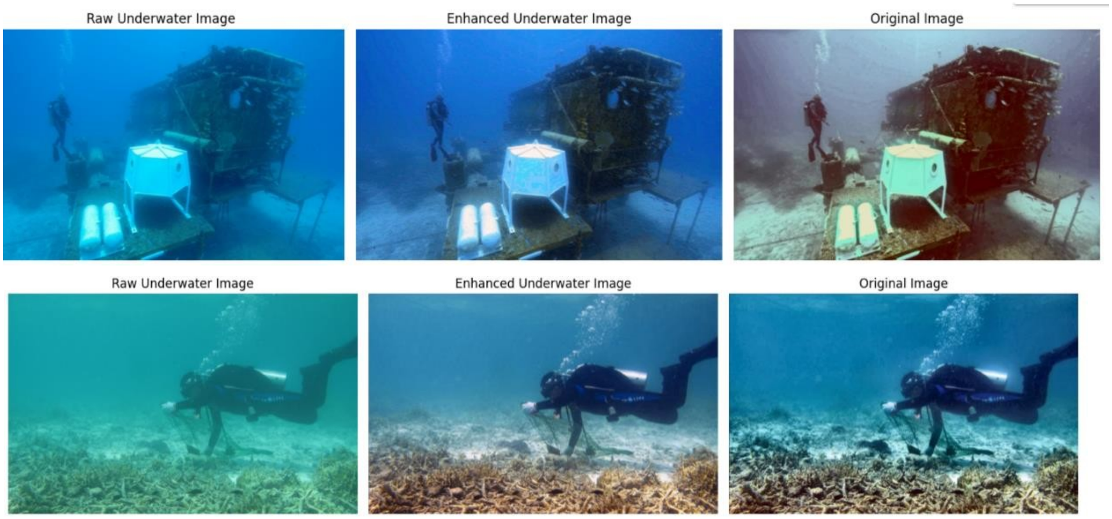


Fig. 4.5. (d) (e)

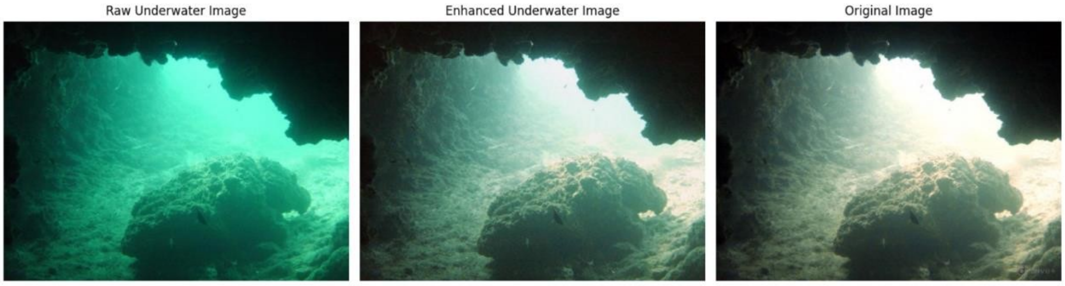


Fig. 4.5. (f)

Fig. 4.5.(g)

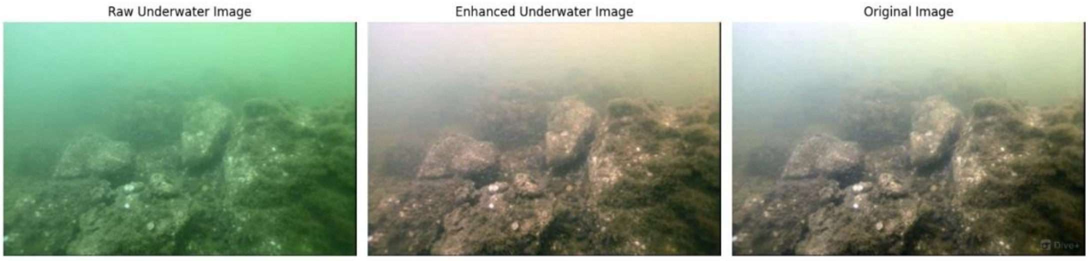


Fig. 4.5. (h)

Fig. 4.5. (i)

Fig. 4.5. (j), Resultant Images of Proposed Method (a to j)

CHAPTER 5

RESEARCH GAP & FUTURE SCOPE

The EEF-GAN model was inspired by the distinctive shortcomings of existing underwater image enhancement techniques Commonly used models typically deal with severe damage such as various lighting conditions and color a distortion that poorly handles underwater environmental characteristics, and results in detail and inaccurate color rendering. It greatly improves feature extraction and detail recovery, thus ensuring that the rendered images are attractive and realistic.

Looking ahead, the expansion potential of EEF-GAN is immense. Optimization of real-time modeling could transform the application of underwater robotics and live ocean surveys. Furthermore, adapting the EEF-GAN to other unpredictable imaging conditions, such as nighttime weather or fog, could expand its applicability. Future research may also seek to integrate this model into an autonomous vehicle for in-flight video processing, and test its effectiveness on different datasets to increase its robustness and versatility If we push these limitations, future research could enhance the impact of EEF-GAN, opening new frontiers in internal imaging and environmental monitoring.

CHAPTER 7
CONCLUSION

This study introduces the EEF-GAN, a new method for underwater image enhancement that significantly improves existing models in terms of clarity and image quality by affecting convolutional contrast together in the encoder part, EEF-GAN used methods such as FUnIE-GAN , UGAN, WaterNet, Comparatively, good performance was obtained in all metrics.

RAUNE-Net, and U-Shape Transformer, as reflected by the highest PSNR and SSIM values and the lowest MSE and RMSE.

The effectiveness of EEF-GAN highlights the power of particularly deep neurons in addressing specific imaging challenges, especially in areas where conventional imaging fails.

However, the study is not without limitations. While the performance of EEF-GAN is promising, it has been tested under controlled experimental conditions and especially on the UIEB dataset. Future research should look to validate the model and possibly improve its robustness in different downstream environments and under different degradation conditions.

Looking ahead, there is considerable scope for expansion of the EEF-GAN programme. Possible improvements could include the integration of more dynamic coding techniques or the use of transfer learning to adapt the model for similar tasks. In other visually challenging situations, such as fog or night scenes.

In addition, further research may seek to reduce the computational requirements of the model to facilitate its implementation in real-time applications. In conclusion, the development of EEF-GAN represents a significant step forward in the application of deep learning to enhance underwater imaging, providing a strong foundation for future advancements in this important area of research.

REFERENCES

[1] C. Li et al., "An Underwater Image Enhancement Benchmark Dataset and Beyond," in IEEE Transactions on Image Processing, vol. 29, pp. 4376-4389, 2020, doi: 10.1109/TIP.2019.2955241.

[2] Fabbri, C., Islam, M. J., & Sattar, J. (2018). Enhancing Underwater Imagery using Generative Adversarial Networks. ArXiv. /abs/1801.04011

[3] L. Peng, C. Zhu and L. Bian, "U-Shape Transformer for Underwater Image Enhancement," in IEEE Transactions on Image Processing, vol. 32, pp. 3066-3079, 2023, doi: 10.1109/TIP.2023.3276332.

[4] Peng, W., Zhou, C., Hu, R., Cao, J., & Liu, Y. (2023). RAUNE-Net: A Residual and Attention-Driven Underwater Image Enhancement Method. ArXiv. /abs/2311.00246

[5] Islam, M. J., Xia, Y., & Sattar, J. (2019). Fast Underwater Image Enhancement for Improved Visual Perception. ArXiv. /abs/1903.09766

[6] C. Fabbri, M. J. Islam and J. Sattar, ”Enhancing Underwater Imagery Using Generative Adversarial Networks,” 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, QLD, Australia, 2018, pp. 7159-7165, doi: 10.1109/ICRA.2018.8460552.

[7] J. Li, K. A. Skinner, R. M. Eustice and M. Johnson-Roberson, ”WaterGAN: Unsupervised Generative Network to Enable Real-Time Color Correction of Monocular Underwater Images,” in IEEE Robotics and Automation Letters, vol. 3, no. 1, pp. 387-394, Jan. 2018, doi: 10.1109/LRA.2017.2730363.

[8] Uplavikar, P., Wu, Z., Wang, Z.: All-in-one underwater image enhancement using domain-adversarial learning. arXiv preprint arXiv:1905.13342 (2019)

[9] Goodfellow, I. J., Mirza, M., Xu, B., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative Adversarial Networks. ArXiv. /abs/1406.2661

[10] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., and Fei-Fei, L. 2009. ImageNet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, pp. 248-255.

[11] A. K. Pandey and A. Singh Parihar, ”A Comparative Analysis of Deep Learning Based Human Action Recognition Algorithms,” 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-7, doi: 10.1109/ICCCNT56998.2023.10308200.

[12] Singh, K., Pandey, A., Agarwal, A. et al. FRN: Fusion and recalibration network for low-light image enhancement. Multimed Tools Appl 83, 12235–12252 (2024). https://doi.org/10.1007/s11042-023-15908-7

[13] A. S. Parihar and K. Singh, ”A study on Retinex based method for image enhancement,” 2018 2nd International Conference on Inventive Systems and

Control (ICISC), Coimbatore, India, 2018, pp. 619-624, doi: 10.1109/ICISC.2018.8398874.

[14] A. S. Parihar, Y. K. Gupta, Y. Singodia, V. Singh and K. Singh, ”A Comparative Study of Image Dehazing Algorithms,” 2020 5th International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India, 2020, pp. 766-771, doi: 10.1109/ICCES48766.2020.9138037.

[15] Singh, K., & Parihar, A. S. (2021). Variational optimization based single image dehazing. Journal of Visual Communication and Image Representation, 79, 103241. https://doi.org/10.1016/j.jvcir.2021.103241

[16] K. Singh and A. S. Parihar, ”A comparative analysis of illumination estimation based Image Enhancement techniques,” 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-5, doi: 10.1109/icETITE47903.2020.195.

[17] Singh, K., Parihar, A.S. Illumination estimation for nature preserving low-light image enhancement. Vis Comput 40, 121–136 (2024). https://doi.org/10.1007/s00371-023-02770-9

[18] Wang, N., Zhou, Y., Han, F., Zhu, H., and Yao, J. 2019. UWGAN: Underwater GAN for real-world underwater color restoration and dehazing. arXiv preprint arXiv:1912.10269.

[19] A. Pipara, U. Oza and S. Mandal, "Underwater Image Color Correction Using Ensemble Colorization Network," *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, Montreal, BC, Canada, 2021, pp. 2011-2020, doi: 10.1109/ICCVW54120.2021.00228.

[20] Mirza, M., & Osindero, S. (2014). Conditional Generative Adversarial Nets. *ArXiv*. /abs/1411.1784

[21] S. Han, J. Wang, Z. Pan and Z. Shen, "MSDR-Net: Multi-Scale Detail-Recovery Network for Single Image Deraining," *2022 China Automation Congress (CAC)*, Xiamen, China, 2022, pp. 4823-4828, doi: 10.1109/CAC57257.2022.10055299.

LIST OF PUBLICATIONS

1. Minal Tandekar, Anil Singh Parihar, "Deep Learning Approaches to Underwater Image Enhancement: Performance Metrics and Evaluations". Accepted at the International Conference on Intelligent Computing and Communication Techniques (ICICCT 2024).

   Paper ID: 665, Indexed by Scopus

Taylor & Francis
Taylor & Francis Group

Scopus

ICICCT

About The *ICICCT*

ICICCT-2024 is a non-profit conference and the objective is to provide a platform for academicians, researchers, scholars and students from various institutions, universities and industries in India and abroad, to exchange their research and innovative ideas in the field of Cloud, Communication and Internet of Things. We invite all students, research scholars, academicians, engineers, scientist and industrialist working in the field of Intelligent Computing, Communications and Techniques from all over the world. We warmly welcome all the authors to submit their original research in upcoming conference ICICCT-2024 to share their knowledge and experience among each other in collaboration with JNU and EVedant Foundation.

Papers submitted to ICICCT-2024 will undergo a strict double-blind review process. All papers that are accepted and presented in ICICCT-2024 will be published in the ICICCT-2024 Conference Proceeding. The proceeding will be published by Taylor & Francis and Indexing in Scopus.

SUBMIT ARTICLE | DOC TEMPLATE | LATEX TEMPLATE

Notification of acceptance of paper id 665 Inbox ×

Microsoft CMT <email@msr-cmt.org> May 10, 2024, 1:32 PM
to me

Dear Dr./ Prof. Minal Tandekar,

Congratulations...

Your paper / article paper id 665: Deep Learning Approaches to Underwater Image Enhancement: Performance Metrics and Evaluations has been accepted for publication in International Conference on Intelligent Computing and Communication Techniques at JNU New Delhi, India.

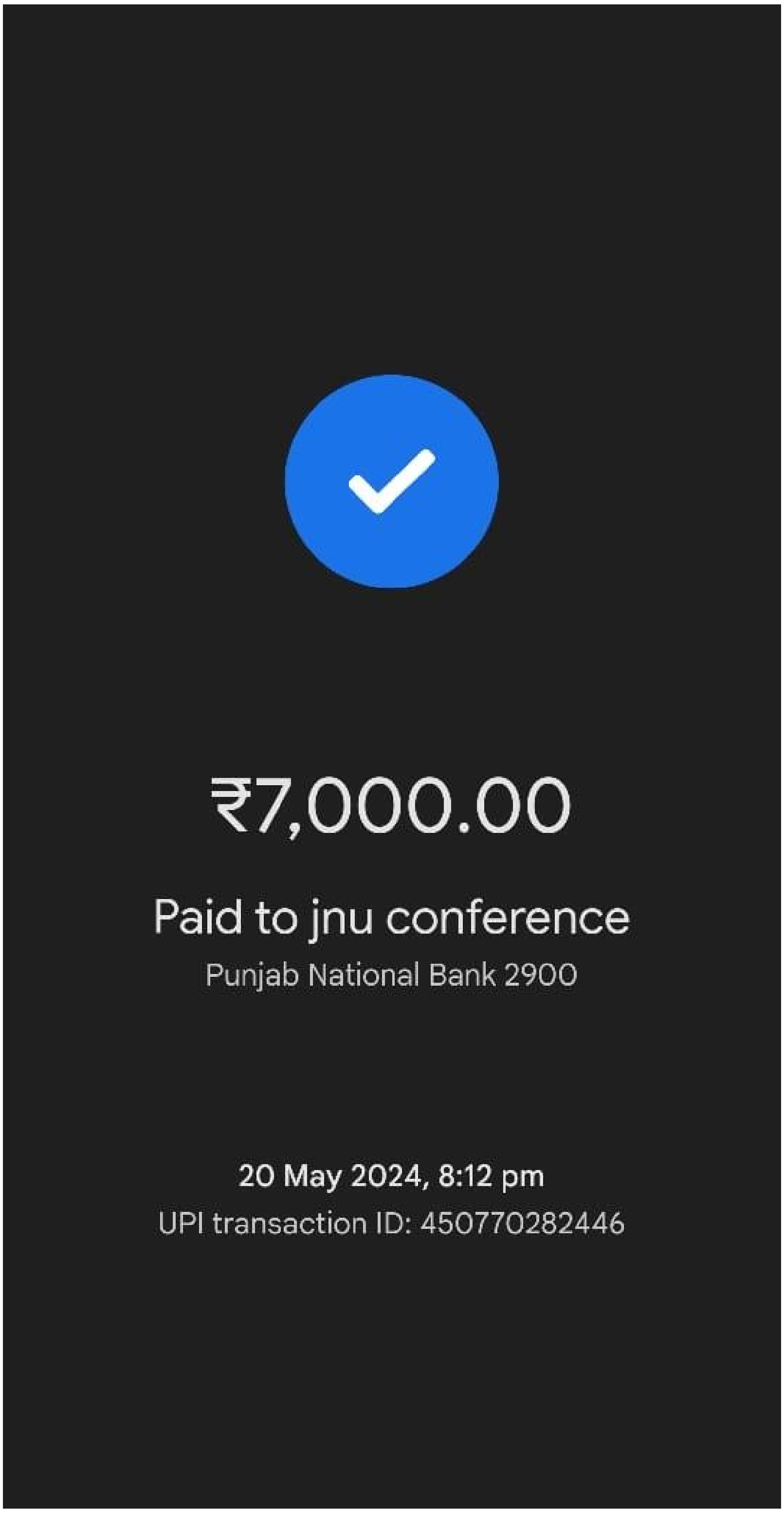
Kindly save your paper by given paper id only (eg. 346.docx, 346.pdf, 346_copyright.pdf)

Registration Link:

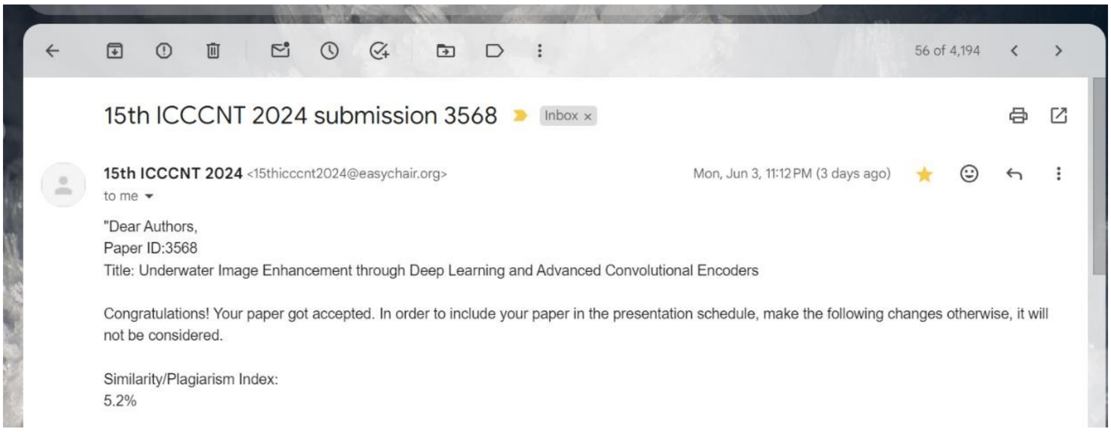
https://forms.gle/mSsHa8GMLtMkWuag8

Please ensure the following before registration and uploading camera ready paper.
1. Paper must be in Taylor and Frances Format.
Template and copyright with author instruction are given in below link: https://icicct.in/author_inst.html

2. Minimum 12 references should be cited in the paper and all references must be cited in the body. Please follow the template.
3. The typographical and grammatical errors must be carefully looked at your end.
4. Complete the copyright form (available at template folder).
5. The regular fee (Available in registration section) will be charged up to 6 pages and after that additional Rs.1000 for Indian authors / 10 USD for


₹7,000.00
Paid to jnu conference
Punjab National Bank 2900
20 May 2024, 8:12 pm
UPI transaction ID: 450770282446

2. Minal Tandekar, Anil Singh Parihar,” Underwater Image Enhancement through Deep Learning and Advanced Convolutional Encoders.” Accepted at the 15th International IEEE Conference on Computing Communication and Networking Technologies.


56 of 4,194

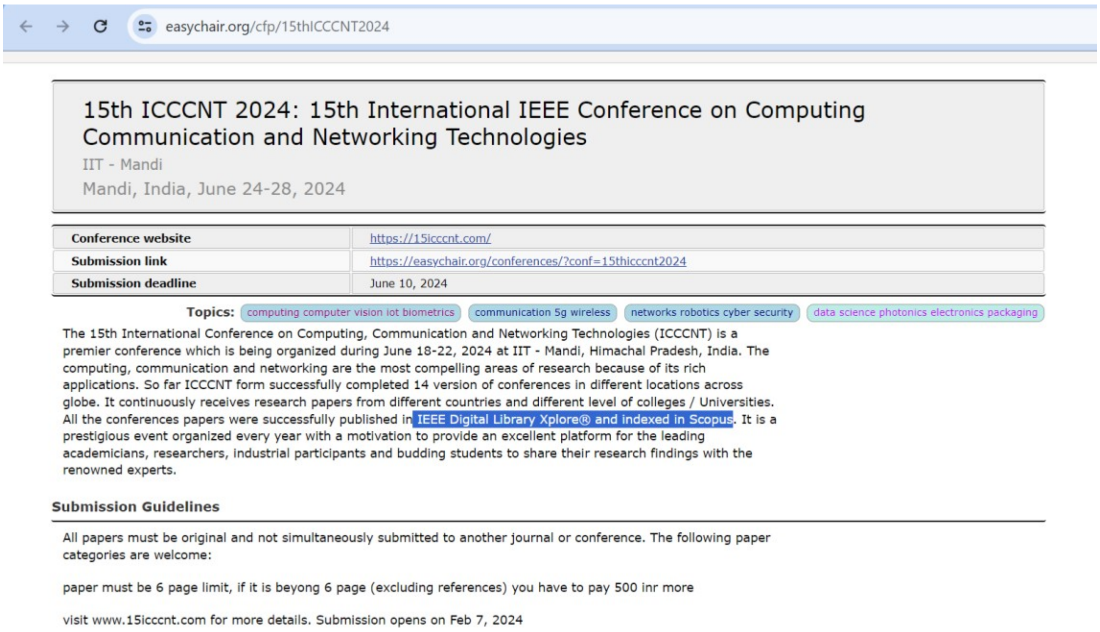
15th ICCCNT 2024 submission 3568 Inbox ×

15th ICCCNT 2024 <15thicccnt2024@easychair.org> Mon, Jun 3, 11:12 PM (3 days ago)
to me

"Dear Authors,
Paper ID:3568
Title: Underwater Image Enhancement through Deep Learning and Advanced Convolutional Encoders

Congratulations! Your paper got accepted. In order to include your paper in the presentation schedule, make the following changes otherwise, it will not be considered.

Similarity/Plagiarism Index:
5.2%


easychair.org/cfp/15thICCCNT2024

15th ICCCNT 2024: 15th International IEEE Conference on Computing Communication and Networking Technologies
IIT - Mandi
Mandi, India, June 24-28, 2024

Conference website	https://15icccnt.com/
Submission link	https://easychair.org/conferences/?conf=15thicccnt2024
Submission deadline	June 10, 2024

Topics: computing computer vision iot biometrics | communication 5g wireless | networks robotics cyber security | data science photonics electronics packaging

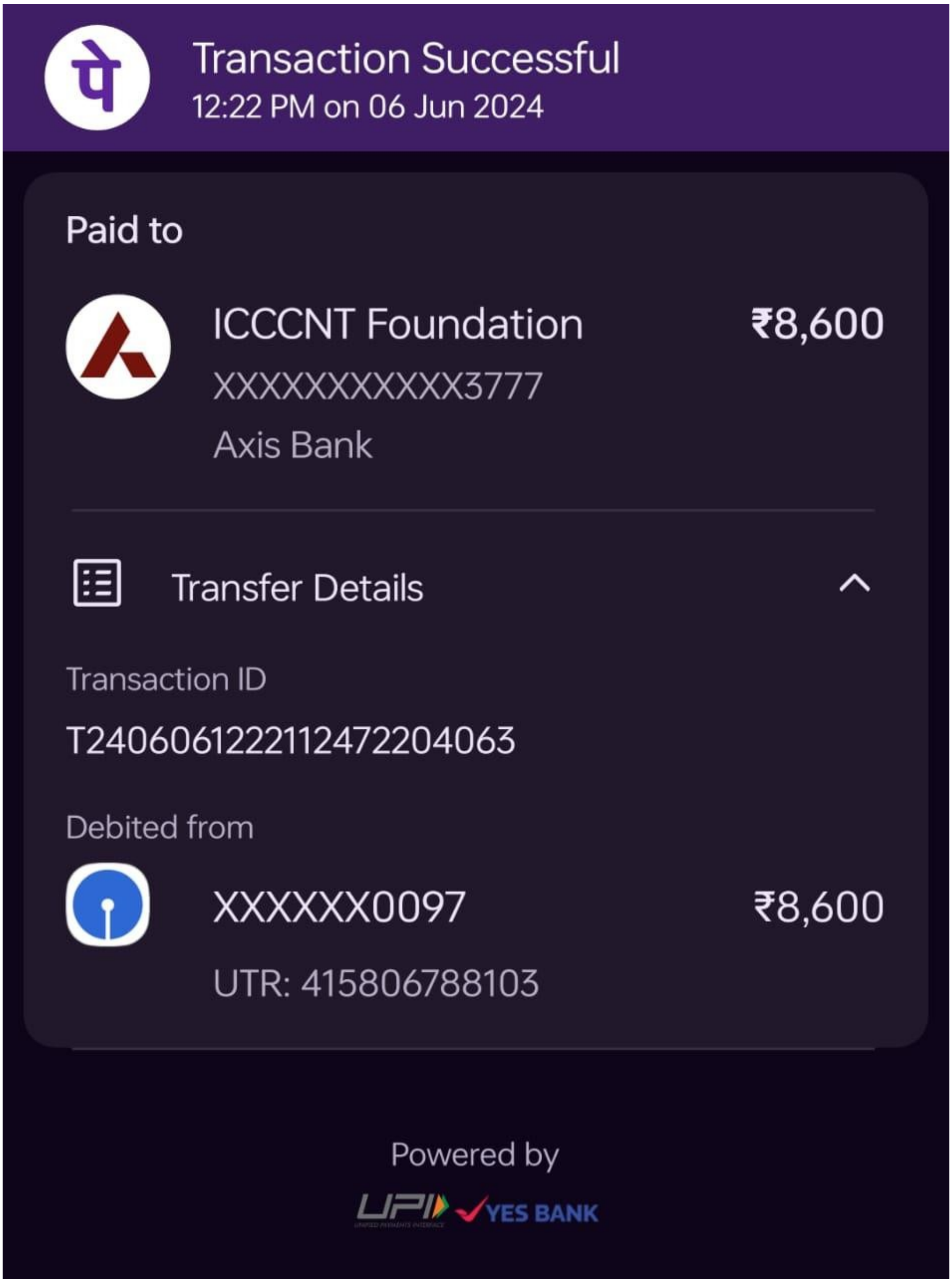
The 15th International Conference on Computing, Communication and Networking Technologies (ICCCNT) is a premier conference which is being organized during June 18-22, 2024 at IIT - Mandi, Himachal Pradesh, India. The computing, communication and networking are the most compelling areas of research because of its rich applications. So far ICCCNT form successfully completed 14 version of conferences in different locations across globe. It continuously receives research papers from different countries and different level of colleges / Universities. All the conferences papers were successfully published in IEEE Digital Library Xplore® and indexed in Scopus. It is a prestigious event organized every year with a motivation to provide an excellent platform for the leading academicians, researchers, industrial participants and budding students to share their research findings with the renowned experts.

Submission Guidelines

All papers must be original and not simultaneously submitted to another journal or conference. The following paper categories are welcome:

paper must be 6 page limit, if it is beyong 6 page (excluding references) you have to pay 500 inr more

visit www.15icccnt.com for more details. Submission opens on Feb 7, 2024



Transaction Successful
12:22 PM on 06 Jun 2024

Paid to

ICCCNT Foundation ₹8,600
XXXXXXXXXXX3777
Axis Bank

Transfer Details

Transaction ID
T2406061222112472204063

Debited from

XXXXXX0097 ₹8,600
UTR: 415806788103

Powered by
YES BANK

PAPER NAME

chapters.pdf

AUTHOR

Minal

WORD COUNT

9087 Words

CHARACTER COUNT

53653 Characters

PAGE COUNT

43 Pages

FILE SIZE

1.8MB

SUBMISSION DATE

May 31, 2024 10:52 AM GMT+5:30

REPORT DATE

May 31, 2024 10:52 AM GMT+5:30

15% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 9% Internet database
- 6% Publications database
- Crossref database
- Crossref Posted Content database
- 13% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Quoted material
- Cited material
- Small Matches (Less then 8 words)