Noise Suppressed Real-Time Enhancing Environment using Deep Learning

[1]Mrs. Roopashree C S, [2]Anu Jagdeesh, [3]Jeslita Viyola Almeida

Department of Information Science & Engineering
B.N.M Institute of Technology
Bangalore – 560070, Karnataka, India.

***Abstract*- MIRNet is an innovative architecture designed to improve image restoration by addressing the trade-offs between spatial precision and contextual robustness in conventional approaches. It maintains high-resolution representations throughout the network, ensuring precise spatial detail, while also integrating multi-resolution data for broader contextual understanding. This is achieved through a multi-scale residual block featuring parallel convolution streams at different resolutions, allowing for diverse feature extraction. The architecture enables information sharing across these streams, which is key to capturing contextual information from varying scales. To further enhance performance, spatial and channel attention mechanisms are used, focusing the network on critical areas and features. An attention-based aggregation system then combines multi-scale features into a cohesive output, which balances fine detail with broader context. Experiments with MIRNet on various benchmark datasets have shown its effectiveness, achieving state-of-the-art results in tasks like image denoising, super-resolution, and image enhancement.**

***Index terms*: MIRNet, Image Restoration, Spatial Precision, Feature Extraction, Image Denoising, Super-Resolution, Benchmark Datasets.**

I. INTRODUCTION

Photographs are precious records of memories, but they often degrade over time due to improper storage, leading to damage such as scratches, film grain, sepia effects, and color loss. Traditional restoration methods relied on manual retouching or inpainting techniques to repair localized damage, which could be laborious and lacked the capability to fix broader spatial degradation issues. With the advancement of deep learning, specifically convolutional neural networks (CNNs), there has been a shift towards automated restoration techniques that leverage large datasets to train models for specific restoration tasks. The Multi-scale Information Restoration Network (MIRNet) provides a solution to these issues by combining high-resolution representation with multi-scale contextual processing. It achieves this through a unique architecture that maintains precise spatial details while processing features at different scales. The key modifications made to MIRNet reduce its computational cost while enhancing its performance, making it more suitable for large-scale photo restoration. By using parallel convolution streams at varying resolutions, MIRNet is able to retain high-resolution features while also integrating broader context, allowing for more accurate restoration. This multi-scale approach enables the network to remove unwanted degraded content while preserving fine spatial details, resulting in restored images that retain their original clarity and quality.

II. LITERATURE REVIEW

T. Sunil Kumar Reddy and A. Dileep Kumar [1] introduces an image restoration method using Variational Autoencoders (VAEs) and a triplet domain translation network to tackle noise removal and quality enhancement in historical images. Their approach excels at restoring both real historical photos and synthetic data, outperforming existing methods in visual quality restoration. However, the authors highlight a potential limitation: the method's reliance on synthetic image pairings might reduce its effectiveness when dealing with unique or uncommon forms of image deterioration, suggesting a need for further research into its adaptability and broader applicability.

Ting Nie, Xiaofeng Wang, Hongxing Liu, Mingxuan Li, Shenkai Nong, Hangfei Yuan, Yuchen Zhao, and Liang Huang [2] the research paper by Ting Nie and colleagues explores various noise reduction methods in image processing, emphasizing the effectiveness of the median filter for tackling image noise. Alongside the median filter, the authors also discuss customized Wiener and fuzzy filters, presenting a practical guide for noise filtering. However, the paper briefly mentions the increased computational complexity of median filters without delving into the details, suggesting a need for further investigation to understand the impact on real-world applications.

Abdalla Mohamed Hambal, Dr. Zhijun Pei, and Faustini Libent Ishabailu [3] The paper by Abdalla Mohamed

Hambal, Dr. Zhijun Pei, and Faustini Libent Ishabailu categorizes different types of image noise and emphasizes the importance of noise removal in image processing. It focuses on various noise reduction methods, especially the median filter, while also introducing adaptive Wiener and fuzzy filters to aid practitioners in filter selection. However, the paper has limitations, particularly regarding its mention of median filter complexity without detailed exploration, and its 2017 publication date, which may affect its relevance to current noise reduction practices given ongoing advancements in the field.

Bhuvan UnhelkarYuo Guo, Yuxu Lu, Ryan Wen Liu, Meifang Yang, and Kwok Tai Chui [4] focuses on improving low-light maritime images for better maritime traffic management through a two-step approach. Initially, the Retinex algorithm is used to optimize illumination, brightening the images and enhancing overall visual quality. The second step involves deep learning-based denoising, reducing noise and improving clarity, resulting in significantly enhanced low-light images suitable for maritime applications.

Aswathy K. Cherian, Eswaran Poovammal, Ninan Sajeeth Philip, Kadiyala Ramana, Saurabh Singh, and In-Ho Ra [5] this introduces a deep learning-based algorithm for enhancing underwater images, employing a combination of a trigonometric-Gaussian filter for denoising and Contrast Limited Adaptive Histogram Equalization (CLAHE) for contrast enhancement. While the approach demonstrates improved image quality in underwater scenarios, the paper could benefit from a deeper discussion on computational complexities and we could add the insights into real-world applications and training data characteristics.

Chunwei Tian, Lunke Fei, Wenxian Zheng, Yong Xu, Wangmeng Zuo, and Chia-Wen Lin [6] this explores deep learning techniques for image denoising, with a focus on enhancing low-light maritime images to improve visibility and noise control in maritime traffic management. It provides a comprehensive overview of Convolutional Neural Networks (CNNs) for denoising, comparing performance across publicly available datasets, but is limited in technical details and discussions on practical challenges, potentially impacting its relevance given advancements in noise reduction technologies.

III. PROPOSED NETWORK

A. CNN Model

Convolutional Neural Network (CNN) for detecting tomato leaf diseases, the process starts with collecting and preparing data. This involves assembling a diverse dataset containing images of tomato leaves, encompassing both healthy specimens and those affected by common diseases such as early blight, late blight, and leaf mold. Ensuring a balanced representation of classes in the dataset is crucial. Before inputting the images into the network, standard preprocessing steps are applied, including resizing them to a consistent size, converting them to a suitable color space (e.g., RGB), and normalizing pixel values.

Moving forward, the architecture of the CNN is carefully designed to suit the task of image classification. Typically, this entails constructing a convolutional base composed of convolutional layers followed by max-pooling layers to extract pertinent features from the input images. Various architectural configurations are experimented with, including adjustments in the number of convolutional layers, filter sizes, and pooling strategies. Additionally, fully connected layers are appended to the architecture to facilitate classification based on the extracted features. Activation functions such as ReLU are utilized to introduce non-linearity into the network.

B. MIRNet Model

The MIRNet (Multi-scale Information Restoration Network) model is a sophisticated deep learning architecture designed for image restoration tasks like denoising, super-resolution, and image enhancement. It uniquely combines high-resolution representation with multi-scale contextual information to achieve state-of-the-art results. The core concept involves maintaining spatial precision while simultaneously incorporating broader contextual understanding. This is achieved through a multi-scale residual block structure, where parallel convolutional streams operate at different resolutions, allowing the model to learn from both high-resolution details and low-resolution context. MIRNet can create a richer feature representation for improved image restoration outcomes.

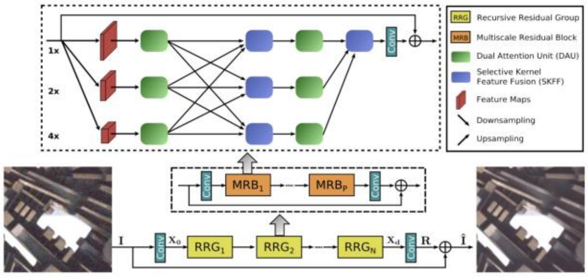

Fig. 1 MIRNet Model

MIRNet employs attention mechanisms to further refine feature extraction. The Dual Attention Unit (DAU) integrates both spatial and channel attention, allowing the network to focus on the most relevant parts of the image while suppressing less informative features. This strategic use of attention enhances the model's ability to deliver high-quality restorations without losing crucial details. To ensure effective training, MIRNet uses a combination of residual resizing modules, a robust loss function, and advanced optimizers, contributing to its consistent performance across various datasets. These design choices enable MIRNet to achieve significant improvements in image quality, making it a powerful tool in the domain of image restoration.This model contains
1. Recursive Residual Block (RRG)
2. Multiscale Residual Block (MRB)
3. Dual Attention Unit (DAU)
4. Selective Kernel Feature Fusion (SKFF)
5. Feature Maps

1. Recursive Residual Block

Recursive Residual Block is a building block used in deep learning models to improve image restoration by enabling deeper networks without causing the degradation problem. It applies residual learning, where each block learns the residual or difference between the input and the desired output. This recursive structure allows the block to be used multiple times within a network, providing a more efficient way to extract features without adding excessive computational complexity. The RRB's recursive nature ensures that the network maintains high-resolution features while still incorporating contextual information, leading to improved performance in tasks like image restoration and enhancement.

2. Multiscale Residual Block

Multiscale Residual Block (MRB) is designed to extract features at multiple scales while retaining spatial precision. It consists of parallel convolutional branches that operate at different resolutions, enabling the network to capture both fine-grained details and broader contextual information. This multi-scale approach allows the network to combine the strengths of high-resolution representation with the benefits of contextual understanding, resulting in enhanced performance for image restoration tasks. The MRB's ability to facilitate information exchange across different scales is key to creating a comprehensive feature set that improves the quality of the final output.

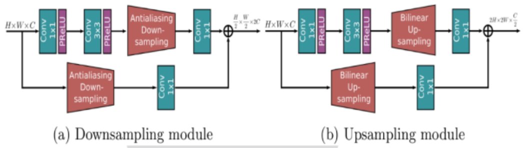

Fig. 2 Multiscale Residual Block

3. Dual Attention Unit

Dual Attention Unit introduces a mechanism to focus the network's attention on the most relevant features, enhancing the quality of the extracted information. It combines spatial attention, which focuses on important regions within an image, with channel attention, which emphasizes significant feature channels. This dual attention approach helps suppress less informative parts of the image while emphasizing critical details, leading to more effective and precise image restoration. The DAU's combination of spatial and channel attention provides a powerful tool for refining feature extraction and improving the overall performance of deep learning-based image processing tasks.

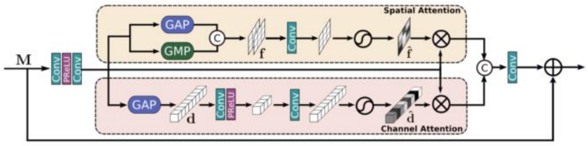

Fig. 3 Dual Attention Unit

4. Selective Kernel Feature Fusion

Selective Kernel Feature Fusion module is designed to dynamically adjust receptive fields, allowing a network to combine information from multiple scales while preserving high-resolution details. By selectively fusing features from different scales, the SKFF provides a flexible and efficient way to balance detailed spatial information with contextual knowledge. This adaptability makes it particularly useful for tasks where varying levels of detail are required. The SKFF's selective fusion strategy contributes to enhanced image restoration by allowing the network to focus on the most relevant features across different resolutions.

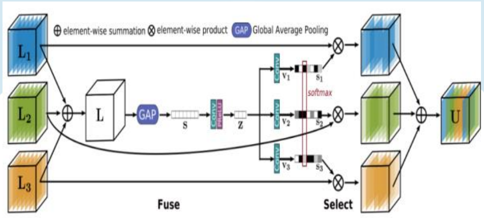

Fig. 4 Selective Kernel Feature Fusion

5. Feature Maps

Feature Maps are a crucial component in convolutional neural networks, representing the output of convolutional layers. They contain the features extracted from the input data and are used by the network to make decisions or further process the information. Feature maps can vary in size and number, depending on the architecture and task, providing insights into the network's focus and attention during processing. They are essential for understanding how a network learns and extracts features, serving as the foundation for advanced operations like attention mechanisms, feature fusion, and multi-scale processing.

IV. RESULTS

When evaluating the performance of our trained MIRNet model, we employ specific metrics and loss functions tailored to the task of image restoration. The model is trained using the Charbonnier Loss function, which is effective for image reconstruction tasks as it provides robustness against outliers and noise. Additionally, we utilize the Adam optimizer with a learning rate set to 1e-4 to efficiently optimize the model parameters during training. To assess the quality of our model's outputs, we measure the Peak Signal-to-Noise Ratio (PSNR) as a key performance metric. PSNR quantifies the quality of reconstructed images by comparing them to the original ground truth images. It calculates the ratio between the maximum possible signal power and the power of corrupting noise that affects the image quality. Higher PSNR values indicate better image quality, with larger values representing less distortion and higher fidelity in the reconstructed images.

Fig.1 Train and Validation Loss over Epochs

By leveraging these specific metrics and loss functions, we can accurately evaluate the effectiveness and robustness of our MIRNet model in restoring image quality, ensuring that it produces visually appealing and faithful results for various image restoration tasks. The use of PSNR as a performance metric allows us to quantitatively assess the model's ability to preserve image details and minimize reconstruction errors compared to ground truth references.

Fig.2 Train and Validation PSNR over Epochs

In addition to using the Peak Signal-to-Noise Ratio (PSNR) as a performance metric for evaluating our trained MIRNet model, we also consider other important metrics such as the Structural Similarity Index Measure (SSIM) and Mean Absolute Error (MAE) rate.

```
In [22]: ssim_score = compute_ssim(output_image, gt_image)
         mae_score = compute_mae(output_image, gt_image)

         print(f"Structural Similarity Index Measure Score: {ssim_score*100.0:.3f}")
         print(f"Mean Absolute Error Rate: {mae_score:.3f}")

         Structural Similarity Index Measure Score: 79.402
         Mean Absolute Error Rate: 0.171
```

Fig 3. SSIM and MAE Value

The Structural Similarity Index Measure (SSIM) score of 79.402 indicates the perceptual quality of reconstructed images compared to the ground truth by evaluating luminance, contrast, and structure. A higher SSIM score, closer to 1, means a greater resemblance in terms of texture and structure, signifying better image fidelity and detail preservation. Additionally, the Mean Absolute Error (MAE) of 0.171 represents the average absolute difference between the pixel values of the reconstructed and ground truth images. A lower MAE value reflects smaller average errors, indicating greater accuracy and reduced deviation in image reconstruction.

Fig 4. Model Performance

V. FINDINGS AND DISCUSSION

- **Real-time Capability:** While MIRNet is complex, its architecture is designed for efficient processing, making it suitable for real-time image enhancement. The use of residual resizing modules and optimized network structures contributes to fast and reliable noise suppression, which is crucial in real-time applications.
- **Versatile Application**: MIRNet's design allows it to be applied to a variety of image restoration tasks, including image denoising, super-resolution, and image enhancement. This versatility makes it a valuable tool in environments where real-time noise suppression and image enhancement are required.
- **High-Quality Outputs:** The combination of multi-scale processing, attention mechanisms, and residual learning enables MIRNet to produce high-quality outputs with minimized noise and artifacts. This quality is critical in applications where image clarity and precision are essential, such as medical imaging, surveillance, and photography.
- **Training and Adaptability:** MIRNet's architecture can be trained with different loss functions and optimized to achieve the desired results, offering adaptability to various noise suppression scenarios.

A. DISCUSSION

- **Computational Complexity:** MIRNet's multi-scale design and attention mechanisms, while enhancing performance, can result in high computational demands. This complexity can pose challenges for real-time applications, requiring substantial processing power and memory, which might not be feasible in resource-limited environments.

- **Latency:** Real-time image enhancement requires rapid processing to maintain a smooth user experience. The computational complexity of MIRNet may introduce latency, causing delays in image rendering or real-time processing. Balancing model accuracy with speed is crucial to ensure real-time capabilities.
- **Model Optimization:** To achieve real-time performance, MIRNet may need significant optimization, such as pruning, quantization, or using specialized hardware like GPUs or TPUs. This optimization process can be complex and requires specialized expertise, adding to the challenge of deploying the model in real-time scenarios.
- **Generalization:** MIRNet's performance depends on the quality and diversity of the training data. In real-time applications, the model must generalize well across different environments and conditions. If the model is not robust to variations in lighting, noise types, or other factors, it may not perform as expected in real-time settings.
- **Maintenance and Updates:** As real-time applications evolve, models like MIRNet must be updated to address new challenges and maintain performance. The need for ongoing maintenance, retraining, and fine-tuning can be challenging, especially in dynamic real-time environments where conditions change frequently.
- **Integration with Existing Systems:** Incorporating MIRNet into existing real-time systems may require significant integration work. Ensuring compatibility with existing software and hardware, along with managing dependencies, can pose a challenge during deployment.
- **Security and Privacy:** Real-time image processing may involve sensitive data, raising concerns about security and privacy. Implementing MIRNet in such contexts requires careful consideration of data protection and secure processing to prevent unauthorized access or data leakage.

VI. CONCLUSION

MIRNet, or the Multi-scale Residual Network, is a deep learning model designed for image restoration tasks, focusing on multi-scale feature extraction and global context integration. The unique network of triplet domains translations mentioned in your description suggests a structure where information is processed across different domains, potentially involving variations in image types like grayscale and RGB, as well as different resolutions. In MIRNet, this concept aligns with its multi-scale approach, where features are extracted at different scales to capture both fine details and broader contextual information, contributing to its robust performance in restoration tasks. The partial nonlocal block that recovers latent features could correspond to MIRNet's use of multi-scale residual connections and nonlocal operations to enhance global context understanding, aiding in scratch inpainting and other restoration techniques. The limitation in handling complex shading highlights a common challenge in image restoration, where MIRNet's effectiveness might be constrained by the quality and variety of its training data, emphasizing the importance of comprehensive datasets to improve generalization and handle intricate shading issues in vintage photographs.

REFERENCES:

1. T. Sunil Kumar Reddy and A. Dileep Kumar, Noise Suppressed Image Enhancing Environment, Volume 8, Issue 6 of November-December 2022
2. Nie, T., Wang, X., Liu, H., Li, M., Nong, S., Yuan, H., Zhao, Y., & Huang, L. (2022). Enhancement and Noise Suppression of Single Low-Light Grayscale Images. Remote Sensing, 14, 3398.
3. Hambal, A. M., Pei, Z., & Ishabailu, F. L. (2017). Image Noise Reduction and Filtering Techniques. Volume 6 Issue 3, March 2017.
4. Y. Guo, Y. Lu, R. W. Liu, M. Yang, and K. T. Chui, "Low-Light Image Enhancement With Regularized Illumination Optimization and Deep Noise Suppression," IEEE Access, vol. 8, pp. 155742-155753, 2020.
5. Cherian, A.K., Poovammal, E., Philip, N.S., Ramana, K., Singh, S., & Ra, I.-H. (2021). Deep Learning Based Filtering Algorithm for Noise Removal in Underwater Images. Water, 13, 2742.
6. Tian, C., Fei, L., Zheng, W., Xu, Y., Zuo, W., & Ling, C.-W. (2020, August 3). Deep Learning on Image Denoising: An Overview PrePrint submitted to arXiv 1912.13171v4 [eess.IV].