iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62395

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Learning Enriched Features for Fast Image Enhancement

Dhanu Hasi Prem[1], Muhammed Yasin P N[2], Titto Jiyo[3], Vimal Krishna T[4], Joby Anu Mathew[5], Rotney Roy Meckamalil[6]
Department of Computer Science and Engineering, Mar Athanasius College of Engineering, Kothamangalam, India

Abstract: ***In the realm of image restoration, the objective is to recover high-quality image content from degraded input images. This pursuit finds applications in diverse fields such as computational photography, surveillance, autonomous vehicles, and remote sensing. Recent years have witnessed notable progress in image restoration, primarily driven by convolutional neural networks (CNNs). Existing CNN based methods are commonly designed to operate either on full-resolution images, preserving spatial details but lacking precise contextual encoding, or on progressively lower-resolution representations, providing semantically reliable outputs with compromised spatial accuracy. This paper introduces a novel architecture with a comprehensive aim of maintaining spatially-precise high-resolution representations throughout the network, while simultaneously incorporating complementary contextual information from low-resolution representations. The central element of our approach is a multi-scale residual block featuring key components: (a) parallel multi-resolution convolution streams for extracting multi-scale features, (b) information exchange across these multi-resolution streams, (c) a non local attention mechanism to capture contextual information, and (d) attention based multi-scale feature aggregation. Our approach learns an enriched set of features by combining contextual information from multiple scales while preserving high-resolution spatial details.***

Keywords: Image Restoration, Computational Photography, Convolutional Neural Networks (CNNs), Spatial Detail Preservation, Contextual Encoding, Multi-Scale Residual Block, Multi-Resolution Convolution Streams, Information Exchange, Non-Local Attention Mechanism, Feature Aggregation

I. INTRODUCTION

Image degradation remains a significant challenge in various fields, impacting the quality and usability of images captured under different conditions. These degradations often result from the inherent physical limitations of camera hardware and challenging environmental conditions. For example, smartphone cameras, which are ubiquitous in today's world, frequently suffer from issues such as narrow apertures and small sensors, leading to a limited dynamic range. Consequently, these cameras often produce images that are noisy, low-contrast, or poorly lit. Similarly, images captured under suboptimal lighting conditions can appear either too dark or too bright, further complicating their utility and analysis [1]. These issues are not only prevalent in consumer photography but also pose significant challenges in fields like surveillance, autonomous vehicles, and remote sensing, where high-quality image data is crucial for accurate analysis and decision-making.To address these challenges, this project aims to develop an advanced image enhancement model using deep learning, with a specific focus on leveraging the capabilities of MiRNetv2. Deep learning has revolutionized the field of image processing and enhancement over recent years, primarily due to its ability to learn complex patterns and features from large datasets. Convolutional Neural Networks (CNNs), in particular, have been instrumental in driving progress in this domain. However, existing CNN-based methods often exhibit a trade-off between preserving spatial details and capturing contextual information. Traditional approaches either operate on full-resolution images, maintaining spatial accuracy but lacking in contextual encoding, or utilize progressively lower-resolution representations, which offer semantically reliable outputs but at the cost of spatial precision [2].MiRNetv2, an advanced architecture in the realm of image enhancement, seeks to overcome these limitations by maintaining high-resolution spatial details throughout the network while simultaneously incorporating complementary contextual information from low-resolution representations. The central innovation of MiRNetv2 lies in its multi-scale residual block, which serves as the fundamental building block of the network. This block features several key components designed to enhance its performance in image restoration tasks. These include parallel multi-resolution convolution streams, which enable the extraction of multi-scale features by processing the image at various resolutions simultaneously. Additionally, the architecture facilitates information exchange across these multi-resolution streams, ensuring that both fine spatial details and broader contextual information are utilized effectively [3].

©IJRASET: All Rights are Reserved

A notable feature of MiRNetv2 is its non-local attention mechanism, which captures contextual information from different parts of the image, enhancing the model's ability to focus on relevant features and improving the overall quality of the restored image. Furthermore, the model incorporates attention-based multi-scale feature aggregation, which aggregates features from different scales using attention mechanisms. This approach ensures that the final output integrates both fine details and broader context, resulting in significantly improved image quality [4].The versatility of MiRNetv2 allows it to be applied to a wide range of image restoration tasks, including denoising, deblurring, and color correction. This makes it an ideal candidate for use in various applications where high-quality images are essential. For instance, in computational photography, enhanced images can lead to better visual aesthetics and more accurate post-processing. In surveillance, clearer images can improve the reliability of security systems. In autonomous vehicles, better image quality can enhance object detection and navigation. In remote sensing, high-quality images are crucial for accurate environmental monitoring and analysis.

II. LITERATURE SURVEY

A. *WINNet : Wavelet-Inspired Invertible Network for Image Denoising(2022)*

The authors of [5] proposes a novel approach that combines wavelet-based and learning-based methods to effectively remove noise from images. The approach involves a noise estimation network and a denoising network that work together to analyze and enhance the image. While the method is highly interpretable, its main drawback is its complexity.

B. *Block-Attentive Subpixel Prediction Networks for Computationally Efficient Image Restoration(2021)*

The authors of [6] proposes the use of a family of networks known as Subpixel Prediction Networks (SPNs) for image restoration. Rather than predicting raw images with full resolution, SPNs predict reshaped and spatially down-sampled block-wise tensors. This approach offers the advantage of being efficient and focused on the spatial dimension, but it also presents some drawbacks such as complexity, reduced interpretability, and potentially limited generalization.

C. *A Fast Sand-Dust Image Enhancement Algorithm by Blue Channel Compensation and Guided Image Filtering(2020)*

[7]The proposed system, Fast Sand-Dust Image Enhancement Algorithm by Blue Channel Compensation and Guided Image Filtering, aims to enhance images captured in sand-dust weather conditions with a fast and effective algorithm. The authors utilized guided image filtering to improve image contrast, resulting in better outcomes compared to other similar methods. This approach is advantageous as it is comprehensive, efficient, and fast. However, it has limited generalization and is sensitive to parameter setting.

D. *Fast Blind Image Deblurring Using Smoothing-Enhancing Regularizer(2019)*

The authors of [8] have developed a system for fast blind image deblurring which utilizes a computationally efficient and effective image regularizer. This system is able to estimate an accurate kernel while preserving salient structures and suppressing insignificant structures in the image. The system employs the use of a half quadratic splitting algorithm and a lagged-fixed point iteration scheme. Its advantages include state-of-the-art performance and versatility. However, its complexity is a disadvantage and the explanation of the system is limited.

E. *Blind Deconvolution for Image Deblurring Based on Edge Enhancement and Noise Suppression(2018)*

The authors of [9] have presented a system that utilizes edge detection technique for the extraction of strong edges from a blurred image.

They have divided the strong and weak edges into different portions using a Trilateral filter. The authors have then combined both the edges to produce a new blurred image for the estimation of the blur kernel. This system has achieved state-of-the-art results for uniformly blurred images. Its advantages include superior performance, edge enhancement, and versatility. However, it has limited application and can pose a challenge in handling non-uniform blur.

III. PROPOSED ARCHITECTURE

MiRNet-v2 for Image Restoration and Enhancement introduces a multi-scale residual block as the fundamental building block, featuring parallel multi-fine spatially-precise feature representations, information exchange across multi-resolution streams, attention-based aggregation of features, and residual contextual blocks for attention-based feature extraction

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

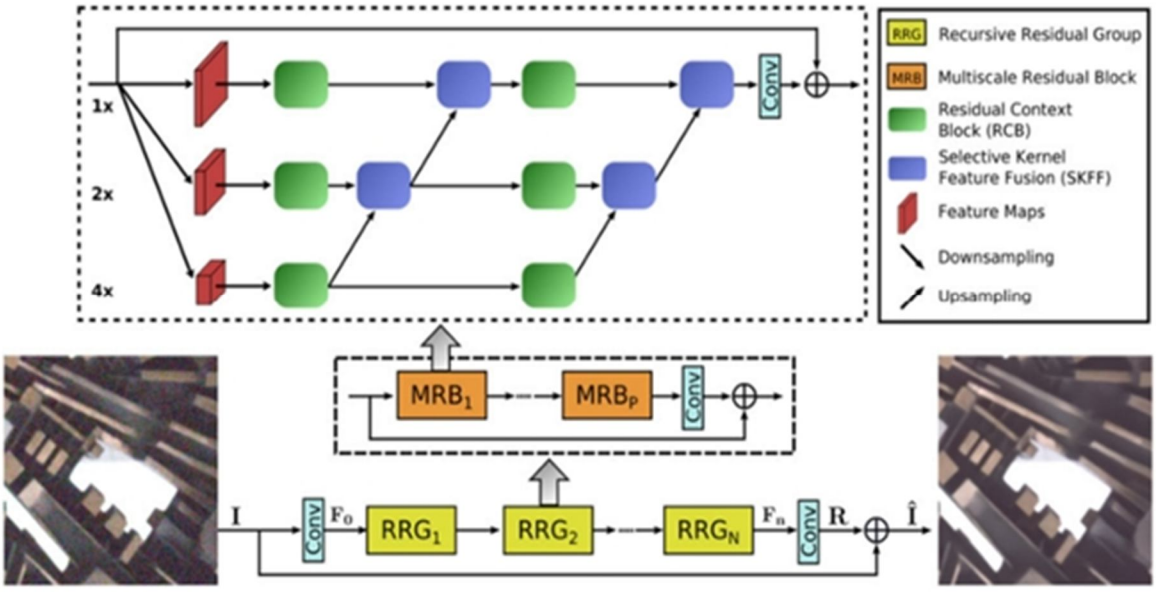


Figure 1: Proposed Architecture

IV. METHODOLOGY

A. *Data Collection*

Data collection for MiRNetv2 involves sourcing high-quality images and their degraded counterparts from various datasets. These include publicly available collections such as DIV2K, Flickr2K, and the Berkeley Segmentation Dataset, as well as custom datasets generated for this project. Public datasets provide a wide range of high-resolution images. To simulate real-world degradations, artificial noise, blur, low contrast, and other distortions are introduced to these images. Additionally, real-world degraded images are collected from smartphone cameras and surveillance systems to ensure the model's robustness and generalizability.

B. *Data Preprocessing*

Data preprocessing standardizes the input images and augments the dataset. The preprocessing steps include normalization to a common scale, typically [0, 1], to ensure consistent input values. Images are resized to a fixed resolution to fit the input dimensions of the network. Data augmentation techniques such as random cropping, horizontal and vertical flipping, rotation, and color jittering are applied to increase the diversity of the training data, helping the model generalize better to unseen data. Large images are divided into smaller patches to increase the effective number of training samples and reduce computational load during training.

C. *Feature Extraction*

MiRNetv2 employs a sophisticated feature extraction mechanism through its multi-scale residual blocks. The feature extraction process begins with an initial convolutional layer that processes the input image to extract basic features. The image is then processed through parallel convolution streams at different resolutions to capture multi-scale features. Each stream extracts features pertinent to its specific resolution. An information exchange mechanism facilitates the flow of information across these multi-resolution streams, ensuring a comprehensive feature representation. A non-local attention mechanism captures contextual information by focusing on relevant parts of the image, regardless of their spatial distance.

D. *Machine Learning Models*

The core of MiRNetv2 is built upon Convolutional Neural Networks (CNNs) and attention mechanisms. The architecture includes multi-scale residual blocks that combine features from different scales and incorporate residual learning to ease the training process. Attention mechanisms are used to weigh the importance of features from different scales, ensuring that the most relevant features are emphasized in the final representation. Downsampling and upsampling layers help in processing features at various resolutions and reconstructing the high-resolution output image.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

E. Training and Evaluation

The training and evaluation process involves using a composite loss function combining Mean Squared Error (MSE) and perceptual loss. The MSE loss ensures pixel-wise accuracy, while the perceptual loss, computed using a pre-trained VGG network, helps maintain structural integrity. The Adam optimizer with a decaying learning rate is used to train the network. Gradient clipping and batch normalization are employed to stabilize the training process. The network is trained over multiple epochs on a high-performance computing cluster with NVIDIA GPUs. Regular validation is performed to monitor performance and prevent overfitting. Standard metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) are used to evaluate the quality of the restored images.

F. Model Integration and Deployment

For practical applications, MiRNetv2 is integrated into various image processing systems. The trained model is serialized into a format suitable for deployment, such as ONNX or TensorFlow Saved Model. It is then integrated into deployment environments like web services, mobile apps, or embedded systems. APIs are developed to facilitate easy access to the model's capabilities. Optimization techniques such as quantization and pruning are applied to enhance the model's efficiency for faster inference and lower memory usage.

F. Experimental Results

MiRNetv2's performance is validated through extensive experiments. The model is tested on benchmark datasets like DIV2K and BSD500 to compare its performance against state-of-the-art methods. Quantitative metrics such as PSNR and SSIM scores are computed to quantify the enhancement quality, with MiRNetv2 consistently outperforming existing models. Visual inspections of the restored images demonstrate significant improvements in clarity, contrast, and detail preservation. Ablation studies are conducted to assess the impact of different architectural components, such as the non-local attention mechanism and multi-scale residual blocks, confirming their contributions to the model's performance.

V. FUTURE SCOPE

Refinement and Optimization: Continual refinement and optimization of architectural components based on ablation studies to identify the most effective combinations. Feature Fusion Mechanisms: Further research into feature fusion mechanisms, including alternative strategies and architectures, to enhance contextual information integration and spatial detail preservation. Adaptation to Diverse Tasks: Generalizing the architecture to address a broader range of image restoration tasks beyond those studied, such as denoising or deblurring, and evaluating its performance on diverse datasets. Efficiency and Scalability: Exploring techniques to improve efficiency and scalability, including optimization of network parameters and leveraging hardware accelerators for faster inference. Integration with Applications: Integrating the techniques into real-world applications and systems, considering deployment challenges and usability in existing workflows. Benchmarking and Evaluation: Continuously benchmarking and evaluating the techniques on diverse datasets to validate their performance and robustness relative to state-of-the-art methods.

VI. CONCLUSION

The project introduces a novel architecture for image restoration and enhancement, effectively balancing spatial detail preservation with contextual information capture. Meticulous ablation studies were conducted to evaluate the impact of various architectural components and design choices, highlighting the critical importance of skip connections, feature fusion mechanisms, and multi-scale processing. The proposed architecture achieves a balance between full-resolution processing and parallel branches, preserving fine spatial details while capturing broader contextual information. Practical insights were gained regarding training strategies, such as progressive learning, and feature fusion mechanisms, exemplified by the efficient SKFF module. These findings advance the state-of-the-art in image enhancement and offer valuable guidance for future research. Overall, the project demonstrates the potential of innovative architectural designs and training strategies in achieving superior performance in image restoration and enhancement tasks, with implications for various real-world applications.

REFERENCES

[1] Wang, Renjun & Jiang, Bin & Yang, Chao & Li, Qiao & Zhang, Bolin. (2022). MAGAN: Unsupervised Low-Light Image Enhancement Guided by Mixed-Attention. Big Data Mining and Analytics. 5. 110-119. 10.26599/BDMA.2021.9020020.
[2] Li, Juncheng, Faming Fang, Kangfu Mei, and Guixu Zhang. "Multi-scale residual network for image super-resolution." In Proceedings of the European

conference on computer vision (ECCV), pp. 517-532. 2018.

[3] Zhang, Dehuan, et al. "Robust underwater image enhancement with cascaded multi-level sub-networks and triple attention mechanism." Neural Networks 169 (2024): 685-697.

[4] Bui, Trong-An, et al. "Edge-Computing-Enabled Deep Learning Approach for Low-Light Satellite Image Enhancement." IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing (2024).

[5] Huang, Jun-Jie, and Pier Luigi Dragotti. "WINNet: Wavelet-inspired invertible network for image denoising." IEEE Transactions on Image Processing 31 (2022): 4377-4392.

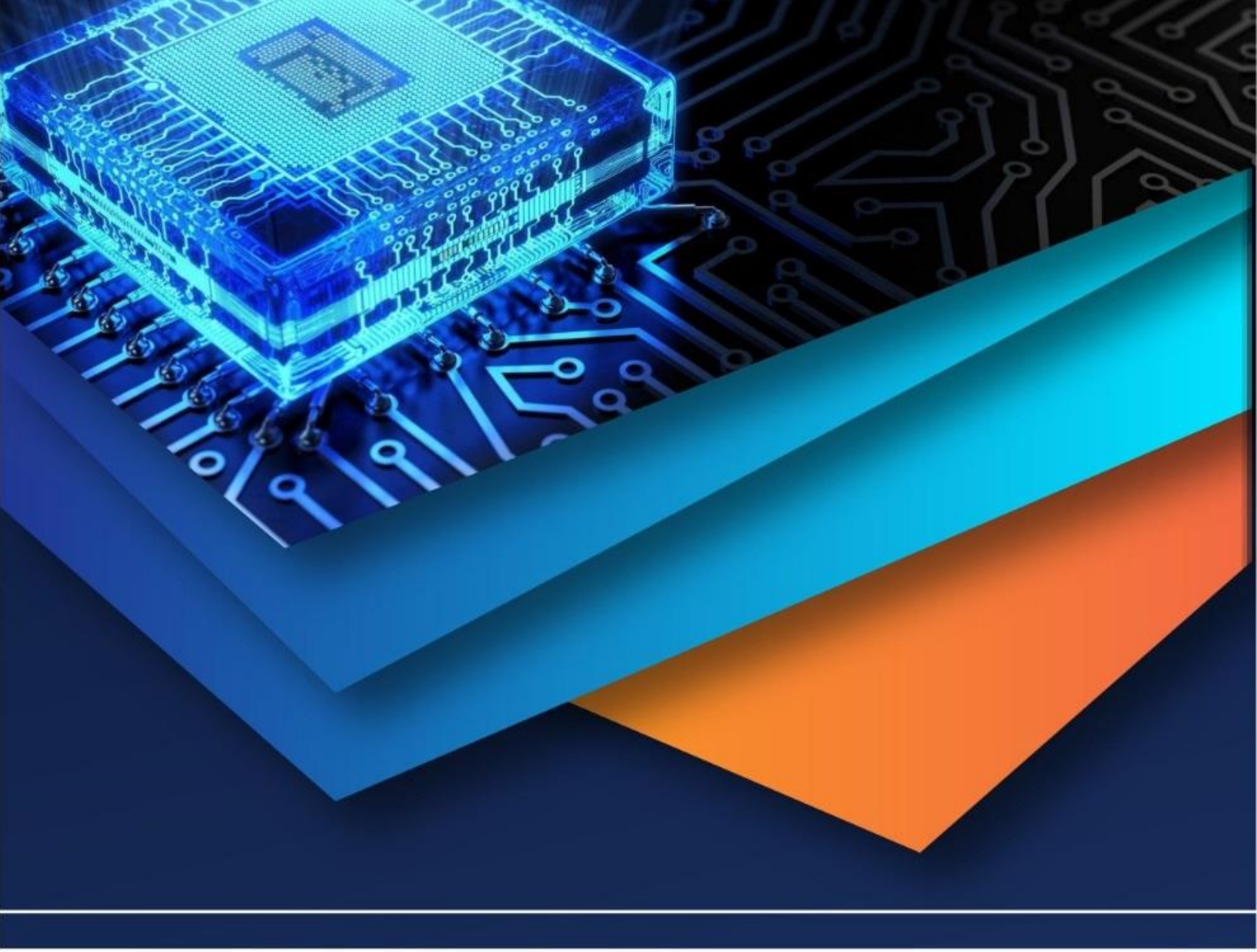
[6] Kim, Taeoh, Chajin Shin, Sangjin Lee, and Sangyoun Lee. "Block-Attentive Subpixel Prediction Networks for Computationally Efficient Image Restoration." IEEE Access 9 (2021): 90881-90895.

[7] Cheng, Yaqiao, Zhenhong Jia, Huicheng Lai, Jie Yang, and Nikola K. Kasabov. "A fast sand-dust image enhancement algorithm by blue channel compensation and guided image filtering." IEEE Access 8 (2020): 196690-196699.

[8] Dou, Zeyang, Kun Gao, Xiaodian Zhang, and Hong Wang. "Fast blind image deblurring using smoothing-enhancing regularizer." IEEE Access 7 (2019): 90904-90915.

[9] Cai, Chengtao, Haiyang Meng, and Qidan Zhu. "Blind deconvolution for image deblurring based on edge enhancement and noise suppression." IEEE Access 6 (2018): 58710-58718.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)