Journal of Theoretical and Applied Information Technology
15th August 2024. Vol.102. No. 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ENHANCING VISUAL CONTENT QUALITY IN EDUCATION AND RELIGION: A FUSION APPROACH OF MIRNET-V2 AND AUTOENCODERS FOR IMAGE PREPROCESSING AND NOISE REDUCTION

SMITA RATH[1,*], SUSHREE BIBHUPRADA B. PRIYADARSHINI[1], PRABHAT KUMAR SAHU[1], MONALISA PANDA[2] DEEPAK KUMAR PATEL[3], NIBEDITA JAGADEV [4], SIPRA SAHOO[5], NARAYAN PATRA[6]

[1,3,6]Associate Professor, Siksha 'O' Anusandhan Deemed to be University , Department of CS&IT,India

[2,4,5]Associate Professor, Siksha 'O' Anusandhan Deemed to be University , Department of CSE,India

E-mail: [1,*]smitarath@soa.ac.in, [1]bimalabibhuprada@gmail.com, [1]prabhatsahu@soa.ac.in, [2]monalisapanda@soa.ac.in, [3]deepakpatel@soa.ac.in, [4]nibeditajagadev@soa.ac.in, [5]siprasahoo@soa.ac.in, [6]narayanpatra@soa.ac.in

ABSTRACT

Combining MIRNet-v2, an advanced image preprocessing technique, with autoencoders for noise reduction can have implications beyond traditional image processing applications. This study could be related to education and religion journals. High-quality images are essential for conveying complex information effectively. Research on advanced image processing techniques like MIRNet-v2 combined with autoencoders could focus on improving the quality of educational images used in textbooks, online courses, or presentations. Therefore, a study combining MIRNet-v2 and autoencoders for image preprocessing and noise reduction could find relevance in journals focusing on education technology, digital humanities, religious studies, or interdisciplinary research at the intersection of technology and society. MIRNetv2 can help print illustrations and book photos by color correction, which involves adjusting colors to improve their accuracy and vibrancy. It will enhance contrast to improve the visual attractiveness of photographs, and low-light enhancement will add brightness to very dark photos. Machine learning algorithms have shown promising results in noise reduction tasks by learning the statistical characteristics of the noise and the underlying image structures. The central component of our method is a multi-scale residual block that includes several crucial components: (a) parallel multi-resolution convolution flows for gathering multi-scale characteristics; (b) exchange of data across the multi-resolution streams; (c) spatial and channel attention processes for preserving contextual data; and (d) attention-based multi-scale feature accumulation. The proposed approach involves collecting relevant image datasets, preprocessing them, selecting suitable machine learning algorithms, optimizing the model parameters, validating and evaluating the models, and identifying potential areas for future research and improvements. Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Mean Squared Error (MSE) are the three famous performance metrics frequently used in image processing.In contrast to PSNR, which emphasizes pixel-level variations, SSIM considers the image data's structure and brightness. The luminance, contrast, and structure of the processed and original images are compared. The results clearly show that MIRNet-v2 outperforms previous approaches by a significant margin. MIRNet-v2 achieves a total performance gain of 3.44 dB on the LoL, SIDD, and DND datasets and some real-time images as training and testing datasets. In summary, this method trains an extended set of features that combines contextual data from many scales while maintaining high-resolution spatial aspects. Extensive tests on live picture benchmark datasets show that MIRNet-v2 produces cutting-edge performance for various image processing applications, comprising super-resolution, image denoising, and image enhancement.

Keywords: *Auto encoders; Convolutional Neural Network; MIRNet-v2; Noise Reduction; Deep Learning*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

1. INTRODUCTION

Our lives now revolve around digital images used for everything from personal photography to security and surveillance. However, various types of noise frequently contaminate the photos taken in real-world situations, lowering their quality and complicating subsequent image analysis tasks. A promising area of research that seeks to solve these problems is to enhance the image in conditions of low light and reduced noise. The underlying research has a significant potential to advance machine learning and image processing, with substantial social benefits. Success in this research could result in advancements in machine learning, security, healthcare, and accessibility. This text will thoroughly overview this fascinating study area and catalyze additional study and advancement. In this study, we explore the application of MIRNet-v2 and autoencoders, two deep learning-based methods for noise reduction and image preprocessing. MIRNet-v2, deep residual network architecture, has produced encouraging results in image-denoising tasks. It uses skip connections to spread information across layers and several residual blocks to extract features at various scales.

On the other hand, autoencoders are unsupervised learning models capable of discovering effective representations of input data. Our findings show that MIRNet-v2 and autoencoders perform better than conventional techniques at removing different kinds of noise from images. MIRNet-v2 and auto encoders use either technique separately to enhance denoising performance. We ran tests to examine how the network architecture and hyperparameters affected the denoising performance. The study can be related to the two main areas of our life, which are as follows.

Education:

Visual Learning Aids: In education, visual aids are crucial in facilitating learning. High-quality images are essential for conveying complex information effectively. Research on advanced image processing techniques like MIRNet-v2 combined with autoencoders could focus on improving the quality of educational images used in textbooks, online courses, or presentations. This ensures students can access explicit and visually appealing educational materials, enhancing their learning experience.

Accessible Learning Materials: Advanced image processing techniques can also contribute to creating accessible learning materials for students with visual impairments. Noise reduction and enhancement of images can make visual content more discernible through assistive technologies, promoting inclusivity in education.

Religion:

Digitization of Religious Texts and Artifacts: Many religious texts, manuscripts, and artifacts are preserved in images. Enhancing the quality of these images through noise reduction and preprocessing techniques can aid in digitization efforts, ensuring the Preservation and accessibility of religious heritage for future generations.

Online Religious Education: With the increasing popularity of online platforms for religious education and spiritual guidance, explicit and visually appealing content is essential. Applying advanced image processing techniques to improve the quality of images and multimedia content used in online religious courses or websites can enhance the learning experience for individuals seeking spiritual knowledge and guidance remotely.

In both education and religion, applying advanced image processing techniques improves the quality of visual content. It contributes to the broader goals of accessibility, Preservation, and effective communication of knowledge and beliefs.

1.1. Background Study

Image processing is fundamental to contemporary technology and scientific inquiry in visual information. Due to the inventiveness and commitment of the early pioneers who laid the groundwork for this field, it has undergone significant evolution. Image processing underwent a paradigm shift with the introduction of digital imaging, moving from conventional techniques to digital algorithms and computational methods. Image processing has become significantly more powerful thanks to developments in hardware and software, which have made it possible to manipulate and analyze digital images in more sophisticated and potent ways. Knowing how image processing has evolved historically can help one better understand how it is progressing and changing various fields.

Significant contributions to the theoretical underpinnings of image processing have been made by notable individuals like Harry Nyquist, Albert J. Ahumada Jr., and Claude Shannon, particularly in image representation, sampling, and quantization. Their efforts laid the groundwork for digital imaging and offered a solid basis for later

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

developments [1]. The development of digital imaging brought on a substantial period of change in image processing. Digital image acquisition, storage, and manipulation replaced earlier film and chemical production methods. Devices like digital cameras, scanners, and imaging sensors have enabled capturing and digitizing visual data directly. With the limitations of physical media gone, image processing is now more effective and flexible [1-5].

The use of digital algorithms and computational methods is one of the distinctive features of contemporary image processing. Mathematical operations are precisely manipulated and analyzed because digital images will be represented as arrays of discrete pixels. Convolution, filtering, and transformation made noise reduction [6], image enhancement[7], and feature extraction possible. Image processing has reached new heights thanks to developments in hardware and software. Moore's Law executes image processing algorithms more quickly and effectively thanks to, which clarifies the enhancement of computer power in an exponential way. Parallel computing architectures and Graphics processing units (GPUs) have significantly accelerated image processing tasks [8]in various ways.

Creating specialized software frameworks and libraries, like OpenCV and MATLAB, has also given researchers and professionals robust tools to study and use image processing methods [9]. The field of image processing has undergone a remarkable evolution. Knowing how image processing has evolved historically can help one better understand how it has transformed various fields. Image processing has evolved from its theoretical roots to become an essential tool for manipulating and analyzing visual data, thanks in part to the development of digital imaging and the application of digital algorithms [10-14]. Image processing will become more and more important as technology develops to utilize visual data in a variety of application areas fully.

1.2. Overview of the Model

This paper aims to assess MIRNet-v2 and auto-encoder performance in image preprocessing and noise reduction tasks and compares the results with those of conventional image processing techniques. It examines their advantages and disadvantages when dealing with different kinds of noise and evaluates their potential for real-time applications. A data set of clean and noisy images from various sources, including artificial, natural, and medical imaging, is gathered to form the data set. It has been observed that different noise types, such as Poisson, Gaussian, and salt-and-paper, corrupted the noisy images due to their varying degrees of intensity. It is then subsequently trained with MIRNet-v2 and auto encoder models on this data set using the PyTorch deep learning framework.

Structural similarity measurements index (SSIM) and peak signal-to-noise ratio (PSNR) within the denoised output are used to assess how well the trained models performed on the test dataset. The outcomes of the process have essential ramifications for several disciplines, including astronomy, industrial inspection, and medical imaging[15-18]. The project also offered insights into the exemplary network architecture and hyperparameters of deep learning-based methods for real-time applications. Enhancing book images for reading using models like MIRNetv2 is essential for several reasons:

1. Improved Readability

Enhanced Text Clarity based on as

- **Sharpening Text**: Blurry or fuzzy text can be challenging to read. Enhancement techniques can sharpen the text, making it more transparent and legible.
- **Noise Reduction**: Scanned images often contain noise, which can obscure text and make it hard to read. Reducing this noise improves text clarity.

2. Better Preservation and Archiving

Long-Term Preservation based on

- **Quality Restoration**: Older books and manuscripts often deteriorate over time. Digital enhancement helps restore and preserve these texts in better quality for future generations.
- **Digital Archiving**: Enhanced images provide higher-quality digital archives, which are valuable for libraries, museums, and researchers.

3. Accessibility

Support for Visually Impaired Readers based on

- **Readable Formats**: Enhanced images are converted into more readable formats for people with visual impairments, such as higher contrast text or larger fonts.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

- **Assistive Technologies**: High-quality images work better with assistive technologies like screen readers and text-to-speech software.

4. Enhanced Visual Appeal

Better Illustrations and Photos based on

- **Color Correction**: Improved color accuracy and brightness make illustrations and photos more visually appealing.
- **Detail Enhancement**: Enhancing details in images makes them more engaging and informative for readers.

5. Academic and Research Use

Improved Data Extraction based on

- **OCR Accuracy**: Optical Character Recognition (OCR) software, which transforms scanned photos into modifiable and searchable text, performs better on enhanced images.
- **Detailed Analysis**: High-quality images are essential for researchers who need to analyze fine details in texts and illustrations.

6. Convenience and Efficiency

Better Reading Experience based on

- **Ease of Reading**: Clearer and more detailed images make reading less straining and more enjoyable.
- **Efficient Study**: Students and researchers can study texts more efficiently when the quality of the images is high.

7. Support for Digitization Projects

Large-Scale Digitization based on

- **Mass Scanning**: Libraries and institutions undertaking large-scale digitization projects benefit from automated image enhancement to ensure consistent and high-quality scans.
- **Cost-Effective**: Automated enhancement reduces the need for manual intervention, saving time and costs.

Enhancing book images using advanced techniques like MIRNetv2 significantly improves the readability, accessibility, and Preservation of texts. It supports academic research, makes reading more enjoyable, and helps preserve valuable literary and historical works in the long term. By addressing issues like blurriness, noise, and low resolution, these enhancements ensure that digital copies of books are as valuable and accessible as possible.

2. MATERIALS AND METHODS

The current research aims to formulate an innovative low-light image enhancement model to improve the image quality captured under varying lighting conditions significantly. Low light conditions pose a considerable challenge in capturing high-quality images. In such situations, the available light is limited, resulting in decreased visibility, loss of details, and increased noise levels. The primary goal of enhancement of image techniques is to restore and enhance the visual quality of images captured under challenging lighting conditions. However, existing methods often need help to achieve optimal results, leading to various limitations we should consider.

Preserving important image details is another crucial aspect of low-light image enhancement. Low light conditions often lead to a loss of fine details due to noise reduction or contrast enhancement algorithms [19].

These details include textures, edges, and delicate structures essential for accurate representation and visual fidelity. Preserving such information while improving the overall image quality is a challenging task. Therefore, developing a technique that successfully preserves important image details is a fundamental requirement for an effective low-light image enhancement method. Color restoration is another significant challenge in low-light image enhancement [20].

2.1. Image Preprocessing and Noise Reduction using MIRNet-v2 Model

MIRNet-v2 (short for Multi-scale Input Reconstruction Network) is a deep learning-based image processing technique used for image super-resolution, denoising, and restoration tasks [21]. Enhancing the quality of low-light photographs without introducing a lot of noise or artifacts is one of the major issues in computer vision and image processing. It can be challenging to distinguish the crucial elements, and it impairs the efficiency of computer vision algorithms in low-light images due to the impact of lower visibility, low contrast, and high noise levels. Again, detecting the critical details in low-light images can be challenging and can reduce the effectiveness of computer vision algorithms due to the same conditions. Over the years, researchers have proposed various picture-improving techniques to address this issue, including tried-and-true methods like histogram equalization and more modern approaches based on deep learning. MIRNet-v2 employs a data-driven methodology in contrast to conventional techniques that rely on manually created features and heuristics to discover the optimum representation for upgrading low-light photos. MIRNet-v2 reconstructs images at various scales using a multi-resolution method to enhance image restoration. The model mimics the architectural design of a convolutional neural network that tries to map between low and high-resolution images. The network uses skip connections and residual learning to increase efficiency and decrease the parameters needed. Due to its capacity to learn from big datasets and generalize well to new images, deep learning has become prominent in the coming years. CNNs formed the foundation for early deep-learning algorithms for image improvement, mainly concentrating on single-image super-resolution [22-26].

Key Highlights of Using MIRNetv2 for Book Image Enhancement

Superior Image Quality

- **High Resolution**: MIRNetv2 enhances the resolution of scanned images, making text and illustrations sharper and more detailed.
- **Noise Reduction**: The model effectively removes noise from images, improving the clarity of both text and illustrations.

Improved Readability

- **Text Clarity**: Enhanced sharpness and contrast in text make it easier to read, especially for older and degraded scans.
- **Low-Light Enhancement**: Enhances images taken in poor lighting conditions, making the text and images more readable.

Preservation and Archiving

- **Restoration of Old Books**: Enhances the quality of images from old and deteriorated books, helping preserve them for future generations.
- **High-Quality Digital Archives**: Produces valuable digital copies for libraries, museums, and research institutions.

Enhanced Accessibility

- **Support for Assistive Technologies**: Improved image quality facilitates using screen readers and other assistive technologies, making books more accessible to visually impaired readers.
- **Readable Formats**: Enhanced images can be adapted into more readable formats with better contrast and clarity.

Better Illustrations and Photographs

- **Color Correction**: Corrects and enhances colors, making illustrations and photographs more vibrant and accurate.
- **Detail Enhancement**: This brings finer details in illustrations and photos and enriches the visual experience.

Increased OCR Accuracy

- **Optical Character Recognition (OCR)**: Improved image quality enhances the accuracy of OCR software, making text extraction more reliable and efficient.
- **Data Extraction for Research**: Facilitates detailed analysis and data extraction, essential for academic and research purposes.

Efficiency in Large-Scale Digitization

- **Automated Enhancement**: Automates the image enhancement process, reducing the need for manual correction and speeding up large-scale digitization projects.
- **Cost-Effective**: Reduces the costs associated with manual image processing and ensures consistent quality across large volumes of scanned books.

Enhanced Reader Experience

- **Visual Appeal**: Enhanced images with better resolution, clarity, and color accuracy make reading more enjoyable.
- **Ease of Reading**: Clearer and more detailed images reduce eye strain and improve the overall reading experience.

Using MIRNetv2 for enhancing images of books significantly improves the readability, Preservation, accessibility, and visual appeal of both text and illustrations. It supports large-scale digitization efforts and academic research and provides a better reading experience, ensuring that digital copies of books are of the highest possible quality.

2.1.1. Architecture

The foundation of MIRNet-v2 is a recurrent residual layout [27]. The core component of MIRNet-v2 is a Multi-Scale Residual Block (MRB), that is responsible for maintaining the whole system in terms of spatially accurate and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

high-resolution projections. Further, a supplementary set of simultaneous branches enhances the model performance by the introduction of various contextual features. SKFF fusion block is capable to combine low-resolution structures with high-resolution structures by sharing information across parallel streams.

2.1.2. Principal Components of the MIRNet-v2 Model

The underline feature extraction technology retains the original high resolution elements in order to preserve the fine spatial aspects of the output. Further an additional set of attributes are computed and evaluated using a variety of spatial scales. It is a novel technique for combining features at various scales while accurately preserving the original structural details at each spatial level. It also dynamically combines features with varied receptive fields. It is a recursive residual design as shown in Fig. 1 that allows for the rapid learning of very deep networks by gradually breaking down the input signal.

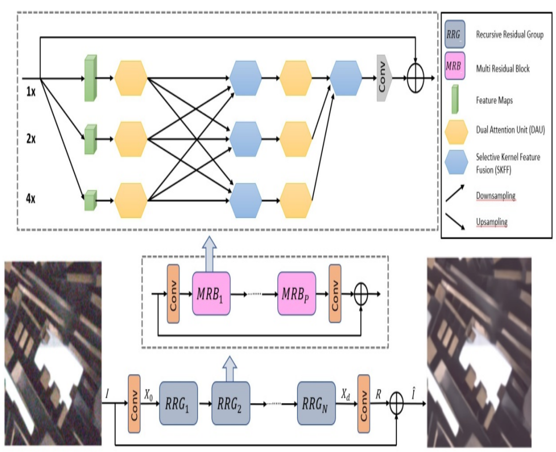

Fig.1 . Structure of MIRNet-v2

2.1.3. Fusion of Specific Kernel Features

The module of SKFF dynamically tunes receptive fields using two processes: fuse and select. The Fuse operator combines data from multiple-resolution streams to provide global feature descriptors.

Fuse: Three simultaneous convolution inputs with various informative scales are used to input the SKFF. We first use an element-wise sum to integrate these multi-scale properties, and then we use Global Average Pooling (GAP) throughout the space dimension. A condensed representation of features is created using a channel-downscaling convolution layer. The fuse operator combines data from many resolution streams to produce global feature descriptors. Two parallel convolution streams containing various informational scales are the inputs to the SKFF module as shown in Fig.2. Element-wise summation are used first to integrate these multi-scale properties. Global Average Pooling, or GAP, is used throughout the combined features' spatial dimension to compute channel-wise statistics. Downscaling operations are then applied to the channel-wise statistics to create a compact feature representation. The final step is to simultaneously use two channel-up scaling convolution layers, one for each resolution stream are applied to the compact feature vector, giving us two feature descriptors.

Select: The select operator uses the feature descriptors of the fuse operator to adjust the characteristic maps of various streams (Fig.2). To create attention activations; SoftMax is applied to the feature descriptors first. The original multi-scale feature maps are adaptively calibrated using a formula that describes the general process of feature recalibration and aggregation, where L1 and L2 are multi-scale feature maps and s1 and s2 are the related attention activations

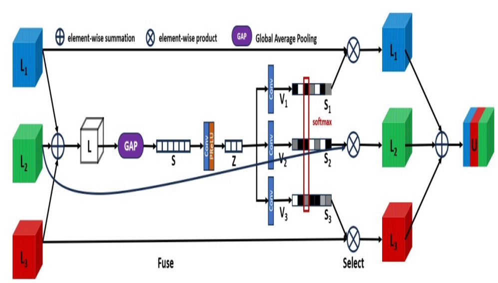

Fig. 2. Selective kernel feature fusion

2.1.4. Dual Attention Unit Module (DAU)

Features are taken out of the convolutional streams by the Dual Attention Unit, or DAU in Fig. 3. The DAU block provides a means for interaction inside a feature tensor across the spatial as well as channel dimensions, whereas the SKFF block merges data over numerous resolution branches. Building a spatial attention map and calibrating input data is the aim of spatial attention.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

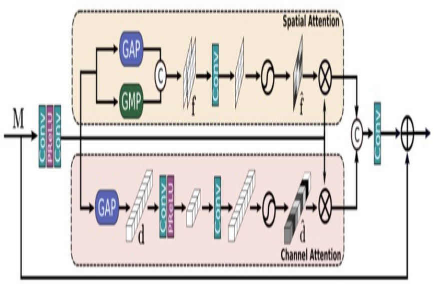

Fig.3 . Dual Attention Unit

2.1.5. Dual Attention Unit Module (DAU)

This block may generate a geographically precise outcome while providing high-resolution images and obtaining rich contextual data from low-resolutions. It allows data to go back and forth across parallel streams, combining low-resolution characteristics with high-resolution details, and vice versa [28-32]. MIRNet-v2 uses a recursive residual design with skip connections (Fig. 4), promoting information flow during learning. Additionally, several residual resizing modules are applied for various up-sampling and down-sampling operations in the MRB to preserve the proposed method's residual nature.

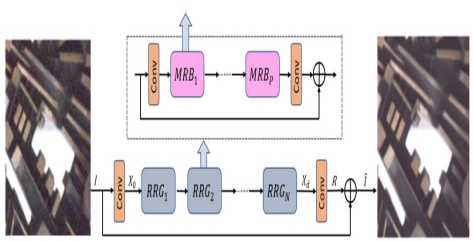

Fig.4 . Multi-Scale Residual Block

2.1.6. Residual Contextual Module

SKFF block fuses data from many branches with different resolutions and requires a distillation method to draw out relevant data from feature tensors. MIRNet-v2-v2 extracts feature from the convolutional streams using the residual contextual block (RCB)[27]. The RCB mutes less valuable features, and only the more informative ones can continue. The complete RCB procedure as shown in Equation 1:

$$FRCB = Fa + W\left(CM\left(Fb\right)\right) \quad (1)$$

Here;

Fa: the input feature maps

Fb: features are applied to two 3x3 group convolution layers to produce feature maps.

CM: represents a contextual module

W denotes the last convolutional layer (filter size =1 × 1).

The Contextual Module consists of 3 parts:

Context Modelling: Initiating with Fb (the initial feature maps), new features (Fc) are created by applying 11 convolutions. The reshaping and SoftMax procedures follow this procedure. Further, the reshaping of Fb is a matrix multiplied by Fc to generate the global feature descriptor Fd.

Feature Transformation: Fd (the descriptor) is routed through 2 numbers of 1111 convolutions to generate Fe, the new attention feature. It helps to figure out the inter-channel relationship.

Feature Fusion: During this fusion step, the contextual parameters (Fe) are combined with every instance of the original feature Fb (via element-wise addition), as shown in Fig. 5.

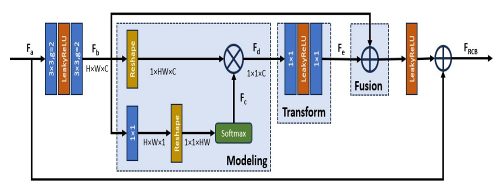

Fig.5 . Feature Fusion

2.1.7. Multi-Scale Residual Block

The primary advantages of MRB are that it is used for maintaining high-resolution representations, allowing it to produce an output that is spatially exact while obtaining rich contextual data from low-resolutions.

3. RESULTS

It is essential to many applications, including digital photography, surveillance systems, computer vision, and medical imaging. It is crucial to evaluate the performance of image processing models to gauge their efficiency and assess alternative methods.

There are three popular performance metrics frequently used in image processing. They are:

- Peak Signal-to-Noise Ratio (PSNR)[33]
- Structural Similarity Index (SSIM)[34]
- Mean Squared Error (MSE)

PSNR: One of the most used performance indicators for evaluating the effectiveness of image processing algorithms is the Peak Signal-to-Noise Ratio (PSNR). It calculates the noise or distortion difference between a treated image and its original version. PSNR is determined using the formula in Equation (2) and is represented in decibels (dB):

$$PSNR = 10\log_{10}\left(R^2 / MSE\right) \quad (2)$$

where R=255 expressed as db.

SSIM:An additional frequently used performance indicator for assessing image processing systems is the Structural Similarity Index (SSIM). In contrast to PSNR, which emphasizes pixel-level variations, SSIM considers both the pictures' structure and brightness data. It compares the luminance, contrast, and structure of the processed and original images to determine their similarity.

The formula used to calculate SSIM is shown in Equation (3):

$$SSIM(i,j) = L(i,j) \bullet C(i,j) \bullet S(i,j) \quad (3)$$

Here,

$L(i,j)$ =Luminance Similarity Index

$C(i,j)$ = Contrast Similarity Index

$S(i,j)$ = Structure Similarity Index

Where x and y are the unprocessed and processed versions of the images, stands for the standard deviation, represents the average of a parameter (such as brightness, contrast, or structure), and C1 and C2 are constants to maintain stability as the denominator approaches zero. A perfect match of the two photos is depicted by 1, an SSIM value, which ranges in [-1,1]. Higher SSIM values denote greater image quality, similar to PSNR. However, by taking into account the structural similarities of images, SSIM offers a more perceptually significant measure.

MSE:In image processing, MSE is an essential performance statistical measurement metric that estimates the mean squared difference between the pixel values of original and processed photos. Equation (4) is used to determine the same:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(y_i \pm y_t\right)^2 \quad (4)$$

Lower MSE values indicate greater performance because they measure the average error between the processed and original pictures.

3.1. Experimental Setup

For three separate restoration responsibilities, three distinct systems were trained. The training parameters are as follows:

There are three RRGs, each containing two MRBs. Each MRB is consisting of 3 parallel streams of varying channel dimensions (64, 128, and 256). These streams operate at different resolutions: 1, 1/2, and 1/4, respectively. Furthermore, each stream is equipped with two DAUs. To train the models, the Adam optimizer were used with particular parameters like; "β1=0.9" and "β2=0.999" over a duration of 7 X 10^5 iterations. During training, we extracted patches from the images, specifically sized at 128 x 128. To handle the training data, the batch size is fixed to16 along with the vertical and horizontal flips applied for data augmentation purposes as shown in Table 1.

Table 1. Evaluation of enhanced low-light image on the SIDD & DND dataset MIRNetv1.

SL. NO.	Dataset	PSNR ↑	SSIM ↑
1.	SIDD	39.72	0.959
2.	DND	39.88	0.956

Four different networks were trained on images, each dedicated to a specific restoration task. The two sub-aperture images (left and right) are merged together and are inputted to the network for smooth handling of dual-pixel defocus deblurring. The training parameters used are as follows:

For each restoration task, four RRGs were implemented, each containing two MRBs. Further, the MRBs are consisting of 3 streams (parallel to each other) with varying channel dimensions: 80, 120, and 180. The corresponding resolutions are 1, 1/2, and ¼ respectively. In each stream of MRB, there are two RCBs with shared parameters. The models use Adam optimizer with specific values (β1=0.9 and β2=0.999) for 3 X 10^5 iterations. The images with patch sizes 128, 144, 192 and 224 are utilized during the process of progressive training. Further, the batch size is fixed with a value of 64, and horizontal and vertical flips are applied for data augmentation purposes..

Table 2. Evaluation of enhanced low-light image on the DPDD dataset on MIRNetv2.

Techniques Used	Indoor Scenes		Outdoor Scenes		Combined	
Metrics	PSNR	SSIM	PSNR	SSIM	PSNR	SSID
JNB[6]	26.73	0.828	21.10	0.608	23.84	0.715
DMENet [20]	25.50	0.788	21.43	0.644	23.41	0.714
MIRnet V2	28.96	0.881	23.59	0.753	26.20	0.816

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

4. DISCUSSION

The network underwent training exclusively on the SIDD training set and was subsequently evaluated on both the test dataset (SIDD and DND) images. The result has been summarized with quantitative comparisons and tabulated sequentially in Tables 1 and 2, respectively. Table 1 includes the output data to evaluate enhanced low-light images on the SIDD & DND dataset with PSNR and SSIM metrics. Table II summarizes the output for the DND dataset. The results from both tables indicate that MIRNet outperforms both data-driven and conventional denoising algorithms. It effectively removes real noise, resulting in visually pleasing and sharp images. Moreover, MIRNet successfully maintains spatial smoothness in homogeneous regions without introducing artifacts. Table 1 and Table 2 present the PSNR/SSIM values of this method and several other techniques, specifically for the DPDD dataset. The results clearly show that MIRNet-v2 outperforms previous approaches by a significant margin. Also, when compared to other methods, MIRNet V2 achieves a total performance gain of 3.44 dB on the LoL dataset. Compared to other techniques, this method produces enhanced visually appealing and natural images and exhibit improved global and local contrast.
The Input and output images of the MIRNet model are given in Fig. 5 – 8.

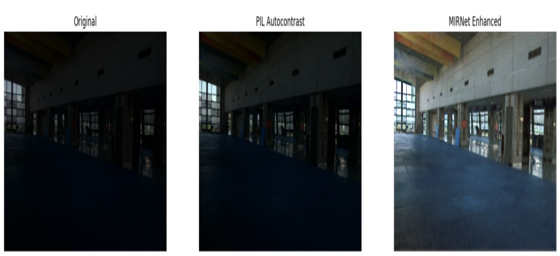

Fig.5. Image Preprocessing of Original to PIL Autocontrast to MIRNet enhanced Images

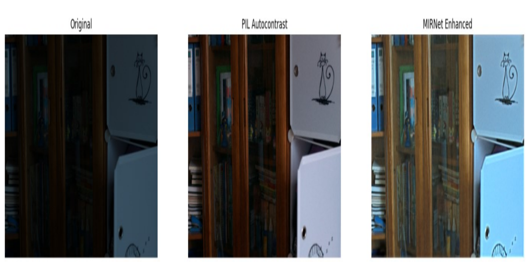

Fig.6. Image Preprocessing of Original to PIL Autocontrast to MIRNet enhanced Images

Fig.7. Image Preprocessing of Original using MIRNet and MIRNetv2 enhanced Images

Fig.8. Image Preprocessing of Original to MIRNet and MIRNetv2 enhanced Images

5. CONCLUSION

Image preprocessing and noise reduction are essential tasks in computer vision and image processing. The emergence of advanced techniques such as MIRNet, MIRNet v2, and autoencoders has significantly improved the quality of images by effectively reducing noise and enhancing visual details. These models leverage deep learning architectures and iterative processes to achieve impressive results. MIRNet, the pioneering model in this area, introduced a novel approach that combines iterative residual learning and multi-scale feature extraction. It aims to enhance noise reduction capabilities by exploring advanced denoising algorithms and incorporating contextual information through attention mechanisms. These models leverage deep learning techniques to denoise images while preserving important details. However, further research and innovation are needed to explore advanced denoising algorithms, improve attention mechanisms, and optimize architectures for real-time performance. With continued advancements, these models have the potential to significantly impact various applications in computer vision, including medical imaging, surveillance systems, and image restoration..

AUTHOR CONTRIBUTIONS

Smita Rath, Sushree Bibhuprada B. Priyadarshini: Conceptualization, Methodology, Software, Field study Deepak Kumar Patel, Prabhat Kumar Sahu, Nibedita Jagadev: Data curation, Writing-Original

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

draft preparation, Software, Validation., Field study Monalisa Panda, Narayan Patra, Sipra Sahoo: Visualization, Investigation, Writing-Reviewing and Editing. All authors have read and agreed to the published version of the manuscript.

ACKNOWLEDGMENTS
The authors are highly grateful to Department of Computer Science & Information Technology, Siksha 'O' Anusandhan Deemed to be University for making this investigation successful.

CONFLICT OF INTEREST
The authors declare no conflict of interest.

REFERENCES:

[1] Mohammed Abd-Alsalam ,Selami, Ameen, and Ahmed Freidoon Fadhil, "A study of the effects of gaussian noise on image features", *Kirkuk Journal of Science*, 11, no. 3 (2016), pp.152-169.

[2] Azzeh, J., Zahran, B. and Alqadi, Z., "Salt and pepper noise: Effects and removal", *JOIV: International Journal on Informatics Visualization*, 2(4),2018, pp.252-256..

[3] Le, T., Chartrand, R. and Asaki, T.J., "A variational approach to reconstructing images corrupted by Poisson noise",*Journal of mathematical imaging and vision,* 27(3),2007, pp.257-263..

[4] Wang, R., Zhang, Q., Fu, C.W., Shen, X., Zheng, W.S. and Jia, J., "Underexposed photo enhancement using deep illumination estimation.", *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*,2019 (pp. 6849-6857)..

[5] Guan, J., Lai, R., Xiong, A., Liu, Z. and Gu, L., " Fixed pattern noise reduction for infrared images based on cascade residual attention CNN", *Neurocomputing*, 377,2020, pp.301-313..

[6] Shi, J., Xu, L. and Jia, J., " Just noticeable defocus blur detection and estimation", *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,*2015, (pp. 657-665)..

[7] I.Ashraf, Y.Park, S.Hur, S.W.Kim, R. Alroobaea, Y.B. Zikria, S. Nosheen, " A survey on cyber security threats in IoT-enabled maritime industry",IEEE Transactions on Intelligent Transportation Systems,2022,pp. 280-288.

[8] Kirchner, M. and Fridrich, J., "On detection of median filtering in digital images", *Media forensics and security II* , 7541, 2010, pp. 371-382..

[9] Kumar, S. and Kumar, B., "Image denoising based on non-local means filter and its method noise thresholding", *Signal, image and video processing ,7, 2013, pp.1211-1227.*

[10] Huo, Y., and Yoon, S. E., "A survey on deep learning-based Monte Carlo denoising". *Computational visual media*,7, 2021, pp.169-185.

[11] Lan, R., Zou, H., Pang, C., Zhong, Y., Liu, Z., and Luo, X., "Image denoising via deep residual convolutional neural networks", *Signal, Image and Video Processing*, 15, 2021, pp.1-8.

[12] Mahapatra, D., Bozorgtabar, B., Thiran, J. P., and Reyes, M., "Efficient active learning for image classification and segmentation using a sample selection and conditional generative adversarial network", *International Conference on Medical Image Computing and Computer-Assisted Intervention* , Cham: Springer International Publishing; 2018, pp. 580-588.

[13] Wang, L., Li, Y., Huang, J., and Lazebnik, S., "Learning two-branch neural networks for image-text matching tasks", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41(2), 2018, pp. 394-407.

[14] Bonnet, B., Furon T., Bas P., "Generating adversarial images in quantized domains", *IEEE Transactions on Information Forensics and Security*, 17, 2021,pp.373-385.

[15] Li. Q., Li C., Yan C., Li X., Li H., Zhang T., Song H., Schaffert R., Yu W., Fan Y., Ye J., Ultra-low Dose CT Image Denoising based on Conditional Denoising Diffusion Probabilistic model, *International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC),* IEEE; 2022, pp. 198-205.

[16] Sharnish,, G, Mishra P, Kohli, A., "Image Denoising for an Efficient Fake Image Identification", *International Conference on Edge Computing and Applications (ICECAA), IEEE, 2022,* pp. 984-989.

[17] Zeng. GL., "Examples of Machine Learning Models from Classic to Modern", *Intermountain Engineering, Technology and Computing (IETC,* IEEE; 2022 May 13, pp. 1-6.

[18] Alam, S. H., Hasan M. N., Sarker M. A., Suvo M. R. and Hasan M. M., "Comparative Analysis Between Gaussian Filter & Its Modified Version for High-Resolution Image Compression", *International Conference on*

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Science & Contemporary Technologies (ICSCT) , IEEE, 2021, pp. 1-5.

[19] Bertalmío, M., *Denoising of photographic images and video: fundamentals, open challenges and new trends.* Springer; 2018.

[20] Muthusamy, T., Aruna, S. P., Sonti, K., and Chakravarthi, P. S., "Underwater Image Enhancement Using MIRNet", Journal of Electronic & Information Systems, 5(1), 2023, pp. 36-44.

[21] Lee, J., Lee, S., Cho, S., and Lee, S., "Deep defocus map estimation using domain adaptation", Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2019, pp. 12222-12230.

[22] Chang, L., Xia, J., "MicroRNA regulatory network analysis using miRNet 2.0", *In Transcription Factor Regulatory Networks,* New York, NY: Springer US, 2022, pp. 185-204.

[23] Zhang, K., Zuo, W, Zhang, L., "FFDNet: Toward a fast and flexible solution for CNN-based image denoising", IEEE Transactions on Image Processing, 27(9), 2018,pp 4608-22.

[24] Rajashekar, U, Simoncelli, EP, "Multiscale denoising of photographic images", *The Essential Guide to Image Processing*, 2009, pp. 241-261.

[25] Fan, L., Li, X., Fan, H., Feng, Y., and Zhang, C., "Adaptive texture-preserving denoising method using gradient histogram and nonlocal self-similarity priors", *IEEE Transactions on Circuits and Systems for Video Technolog,* 2018, 29(11), pp. 3222-3235.

[26] Fan, L, Li, X, Guo, Q, Zhang, C., "Nonlocal image denoising using edge-based similarity metric and adaptive parameter selection", *Science China Information Sciences*, 61, 2018, pp 1-3.

[27] Diwakar, M., Kumar, M., "A review on CT image noise and its denoising", *Biomedical Signal Processing and Control*, 42, 2018, pp.73-88.

[28] Zhang, Z, Li, G, Wang, L., "Seismic Random Noise Suppression Based on MIRNet With Dense Feature Fusion", *IEEE Geoscience and Remote Sensing Letters*, 19, 2022, pp.1-5.

[29] Pinaya, W. H., Vieira, S, Garcia-Dias, R., Mechelli, A., "Autoencoders", *In Machine learning,* Academic Press, 2020, pp. 193-208.

[30]Cheng, S., Wang, Y., Huang, H., Liu, D., Fan, H., Liu, S., "Nbnet: Noise basis learning for image denoising with subspace projection", *In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition,* 2021, pp. 4896-4906.

[31] Dally, W. J., Keckler, S. W., Kirk, D. B., "Evolution of the graphics processing unit (GPU)", IEEE Micro, 41(6), 2021, pp.42-51.

[32]Setiadi, D. R., "PSNR vs SSIM: imperceptibility quality assessment for image steganography", Multimedia Tools and Applications, 80(6), 2021, pp.8423-44.

[33] Mateo, C, Talavera, J. A., "Short-time Fourier transform with the window size fixed in the frequency domain", Digital Signal Processing, 77, 2018, pp.13-21.

[34] Hore, A., Ziou, D., "Image quality metrics: PSNR vs. SSIM", *20th international conference on pattern recognition*, IEEE, 2010, pp. 2366-2369.

[35] Huynh-Thu, Q., Ghanbari, M., "Scope of validity of PSNR in image/video quality assessment", Electronics letters, 44(13), 2008, pp.800-1.