International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Jul-24)

http://dx.doi.org/10.12785/ijcds/160117

Enhancing Image Clarity with the Combined Use of REDNet and Attention Channel Module

Rico Halim[1] **and Gede Putra Kusuma**[1]

[1]*Computer Science Department, BINUS Graduate Program - Master of Computer Science Bina Nusantara University, Jakarta, 11480, Indonesia*

Received 7 Feb. 2024, Revised 12 Apr. 2024, Accepted 19 Apr. 2024, Published 1 Jul. 2024

Abstract: The primary aim of our study is to improve the efficacy of image denoising, specifically in situations when there is a limited availability of data, such as the BSD68 dataset. Insufficient data presents a challenge in achieving optimal outcomes due to the complexity involved in constructing models. In order to tackle this difficulty, we provide a method that incorporates Channel Attention, Batch Normalization, and Dropout approaches into the current REDNet framework. Our investigation indicates enhancements in performance parameters, such as PSNR (Peak Signal to Noise Ratio) and SSIM (Structural Similarity Index), across various levels of noise. With a noise level of 15, we obtained a Peak Signal-to-Noise Ratio (PSNR) of 34.9858 dB and a Structural Similarity Index (SSIM) of 0.9371. At a noise level of 25, our tests yielded a PSNR of 31.7886 decibels and an SSIM of 0.8876. In addition, at a noise level of 50, we achieved a Peak Signal-to-Noise Ratio (PSNR) of 27.9063 decibels and a Structural Similarity Index (SSIM) of 0.7754. The incorporation of Channel Attention, Batch Normalization, and Dropout has been demonstrated to be a crucial element in enhancing the efficacy of image denoising. The Channel Attention approach enables the model to choose and concentrate on crucial information inside the image, while Batch Normalization and Dropout techniques provide stability and mitigate overfitting issues throughout the training process. Our research highlights the effectiveness of these three strategies and emphasizes their integration as a novel way to address the constraints presented by the scarcity of data in image denoising jobs. This emphasizes the significant potential in creating dependable and effective image denoising methods when dealing with circumstances when there is a limited dataset.

Keywords: : Image Denoising, Deep Learning, Channel Attention Module, REDNet Model, Image Processing

1. INTRODUCTION

Image denoising is a long-standing problem that has been intensively studied for a significant amount of time. However, it remains a challenging and unresolved task. The main reason for this is that image denoising, when analyses from a mathematical perspective, is an inverse problem and its solution is not unique [1]. Advancements in the field of Image Denoising have made major advances in tackling this specific challenge. The task of eliminating noise from images has been thoroughly examined over a considerable period, establishing it as a well-known issue. Nevertheless, despite the extensive research conducted, the task remains unresolved. The complexity of image denoising arises from its classification as an inverse problem, which poses challenges in finding a straightforward solution.

Image noise can occur as a result of sensor defects, improper memory allocation, image compression, and transmission, leading to a degradation of the accurate intensity distribution of the scene. Denoising is an intricate process that aims to remove noise from an image while maintaining its quality and intricate details. The primary classifications of noise in images include Impulse Noise (IN) [2][3][4], Additive White Gaussian Noise (AWGN) [3][4], and Speckle Noise [5][6][7], and Speckle Noise. There are two types of impulse noise: Salt and Pepper Noise (SPN) [8][9][10][11] and Random Valued Impulse Noise (RVIN) [12][13].

Gaussian noise, as described in the references [1][3][14][15], follows a Gaussian distribution and exhibits the characteristics of additivity and independence. The cause of this problem can be traced to three main factors: amplifier noise, sensor noise, and grain noise. (1) denotes the presence of noise in an image.

$$I_n(i, j) = I(i, j) + n \tag{1}$$

In equation (1), I_n represents the image with noise, where I represents the original image and n represents the noise present at each pixel [16]. To reduce this noise, one could utilise several strategies such as spatial filtering, frequency filtering, denoising algorithms, and the application

E-mail address: rico.halim@binus.ac.id, inegara@binus.edu

of deep learning. The complex structure of deep learning has resulted in significant advancements by facilitating the acquisition of complicated features and extracting a larger volume of information from images [17]. Image denoising involves many methods that aim to differentiate between noise and authentic image content. The methodologies can be broadly categorised into spatial domain methods and frequency domain methods. The primary goal of spatial domain techniques is to modify the pixel values of an image in order to reduce noise. Filters, such as median filters and Gaussian filters, are commonly employed for this purpose. These filters analyse the intensity levels of pixels in a neighbouring region and replace the central pixel with a computed value. While spatial domain techniques can successfully reduce certain types of noise, they may struggle to preserve the complex details of the image.

On the other hand, the frequency domain approach necessitates the conversion of the image into its individual frequencies, sometimes employing techniques like the Fourier transform. By manipulating the frequency spectrum of the image, it is feasible to extract and eliminate noise. Nevertheless, this method may necessitate intricate mathematical calculations and may lack the user-friendly characteristics of the spatial domain approach. Furthermore, nearly all techniques necessitate substantial quantities of data. This becomes a significant worry if there is an inadequate amount of computational resources.

This study presents novel techniques for enhancing the performance of REDNet by incorporating channel attention, dropout, and batch normalisation. The primary objectives are to dynamically recalibrate feature importance, mitigate overfitting, and stabilise the training process. The contribution of this research is the utilisation of a novel methodology to address the issue of restricted computing resources and datasets. The effectiveness and resilience of this approach are demonstrated by the usage of BSD68 [18].The improved REDNet [19] not only produces greater noise reduction outcomes, but also provides useful knowledge about the deep learning structure for practical image processing assignments, hence promoting additional exploration and advancement in the domain. This paper is structured to offer a thorough examination of relevant research, outline the methodology, present findings and analysis, and summarise future research prospects.

2. RELATED WORKS

REDNet [19] is widely recognised for its unique architecture designed for image restoration applications. It operates using a convolutional neural network structure. The model utilises convolutional and deconvolutional layers to address tiny visual disturbances and preserve detailed features. The utilisation of Encoder-Decoder convolution layers, in addition to establishing direct connections between matching layers, enables the seamless transmission of information from convolutional levels to deconvolutional layers. This intentional interconnection tackles the inherent difficulty of obtaining visual features from simplified representations while simultaneously reducing the decline in performance. This model intentionally utilises downsampling in convolution layers and deconvolution in symmetric deconvolution layers to improve testing efficiency while maintaining performance.

MWCNN [20] is distinguished by its innovative design approach, aiming to provide a seamless integration of receptive field size and computational efficiency in low-level vision applications. This model incorporates wavelet transformations into the U-Net architecture, offering an innovative approach to expand and contract subnetworks. The model comprises contracting and expanding subnetworks, utilising discrete wavelet transform (DWT) to diminish the dimensions of the contracting subnetworks. Moreover, it employs additional convolution layers to reduce the quantity of feature map channels. The model distinguishes itself from standard U-Net designs by employing element-wise addition to combine feature maps from contracting and expanding subnetworks.

PRIDNet [21] is a sophisticated deep learning model specifically created to eliminate intricate noise from images taken in real-world settings. It is highly efficient in circumstances where denoising is required for blind scenarios and the characteristics of the noise are unknown. The architecture comprises three meticulously designed stages: noise estimation, multi-scale denoising, and feature fusion. During the noise estimation process, a compact but powerful five-layer convolutional subnetwork is used to determine the amount of noise in the input image. This model incorporates a vital channel attention module to enhance its capacity to differentiate noise estimates. The multi-scale denoising stage employs a pyramid denoising structure to handle features of various sizes simultaneously. The last phase of feature fusion involves merging denoised features from various scales using a kernel selection module, hence augmenting the model's denoising efficacy.

SUNet [22] is a distinctive architectural design created exclusively for the purpose of image denoising. It seamlessly integrates the Swin Transformer with the U-Net infrastructure. The system comprises three primary modules: Shallow Feature Extraction, U-Net Feature Extraction, and the Reconstruction Module. The integration of these components yields highly effective denoising capabilities. The Shallow Feature Extraction method efficiently handles the input noisy image by employing 3x3 convolution layers to merge shallow features that encompass low-frequency information. The U-Net Feature Extraction use Swin Transformer Blocks instead of conventional convolutions to capture high-level, multi-scale data. The implementation of Swin Transformer Blocks enhances the extraction of intricate and meaningful data. The Reconstruction Module employs 3x3 convolutional layers to reconstruct the image by using deep features, resulting in an output that is free

from noise.

NSTBNet [23] model utilises two modified U-net networks in parallel to enhance breadth instead of depth, efficiently maintaining image texture and details. The image decomposition and reconstruction process in this system utilises NSST (Non-Subsampled Shearlet Transform) and its inverse, effectively preventing any loss of information that may occur from pooling and deconvolution layers. Dilation convolutions increase the size of the receptive field to enhance the extraction of fine details. The model is trained using Mean Squared Error (MSE) loss for denoising tasks, with a specific emphasis on preserving details and removing noise.

DBDIP [24] method presents a new technique for reducing noise in images by utilizing additional information from two fundamental images created using a Y-Net structure. At first, the model divides the contracting path of a U-Net network into two branches to separately extract features using convolution and pooling techniques. The two primary images are produced by these branches and are subsequently combined using a Y-Net dual-branch fusion structure. The fusion procedure entails merging characteristics from the two fundamental images using convolution and up-convolution processes to generate the ultimate denoised image. The D-DIP model notably employs initial denoised images obtained via supervised denoising techniques, such as DAGL and Restormer, to generate the fundamental images. The D-DIP model intends to efficiently improve denoising performance by successfully utilizing complimentary information from these images.

DDT [25] is an advanced neural network specifically developed for the purpose of image denoising. It utilises a distinctive architecture that effectively combines local and global data in a simultaneous manner. The model utilises deformable attention methods to selectively concentrate on important parts in the image, hence minimising superfluous computations and enhancing overall efficiency. The model effectively manages high-resolution images with linear computational complexity by partitioning incoming images into patches and processing them through dual-branch deformable attention blocks. The utilisation of this approach enables efficient reduction of noise and improves the ability to restore images, establishing the model as a highly promising option for jobs involving the removal of image noise. The results obtained from analysing the model under study are presented in Table I.

3. METHODS

A. Dataset

We utilise the BSD68 [18] dataset, which is a subset of the BSD 100 (Berkeley Segmentation Dataset and Benchmark). This collection comprises 68 images encompassing architectural structures, wildlife, persons, and natural settings. For the objectives of this research, the dataset was divided into 40 training, 14 testing, and 14 validation sets. The study utilised images with noise levels of 15, 25, and 50. With the scarcity of image data, our objective was to find a model that can successfully address this issue and deliver images of superior quality.

B. Data Preprocessing

During this stage, we undertake the preparation phase, which consists of three steps: storing the image in memory, producing TensorFlow data, and forming the append list. Initially, we extract and process a collection of image files from the training, testing, and validation directories, transforming them into the RGB colour format and resizing them to a resolution of 256x256 pixels.

Next, we construct a TensorFlow Dataset by utilising input images and labels. We then arrange the data in a specific order and implement augmentation techniques. The training data is diversified by incorporating various augmentation techniques such as flipping in different directions (up, down, left, right), rotating from 0 to 360 degrees, and adjusting hue, brightness, saturation, and contrast.

After that, a specific group of the augmentation functions is employed to create three data generators for the training, testing, and validation datasets. This generator incorporates augmentation techniques during the training process to enhance the model's ability to handle variations in input data.

C. Proposed Model

The ReLU activation function is utilised in both the convolution and deconvolution layers of the proposed model. The choice of ReLU (Rectified Linear Unit) was based on its simplicity and effectiveness in accelerating training and mitigating the vanishing gradient problem. Furthermore, we opted for the MSE (Mean Squared Error) loss function to minimise the disparity between the restored image and the initial image. MSE is frequently selected for image restoration jobs due to its ability to enable the model to focus on the disparity in pixel intensity between the restored image and the original image.

During the training process, we employ the Adam optimizer technique to modify the weights and minimise the loss value. Adam's optimizer is frequently successful in locating a local minimum point within the parameter space. The sequence of stages in this model, starting with the input and ending at the final stage, is as follows:

- Input Layer: The input layer receives a image with specified dimensions.
- Convolution Layer: We employ a sequence of convolutional layers enhanced by Channel Attention, Batch Normalisation, and Dropout to extract significant features from the image. Batch Normalisation is applied after each convolution layer to accelerate convergence, whereas ReLU is used as the activation function.

TABLE I. Analysing the constraints of model variation in Image Denoising.

Methods	Limitation
REDNet[19]	Overfitting could be a concern with such a deep architecture, necessitating careful regularization techniques.
MWCNN[20]	Slower in terms of speed compared to some methods due to model complexity.
PRIDNet[21]	The incorporation of multiple stages and mechanisms may increase the computational complexity of PRIDNet, potentially impacting its efficiency during inference.
SUNet[22]	Implementing and training a model like SUNet may require substantial computational resources and expertise due to its architecture complexity.
NSTBNet[23]	The use of very large datasets results in huge computational efficiency and resources.
DBDIP[24]	The Dual-branch Deep Image Prior model has a higher implementation complexity compared to the original DIP method, requiring more computational resources due to its dual-branch structure. The running time of the DBDIP method is longer than the original DIP method because of its iterative nature and more complex architecture.
DDT[25]	Despite efforts to reduce computational costs, the model may still require significant resources for training and inference.

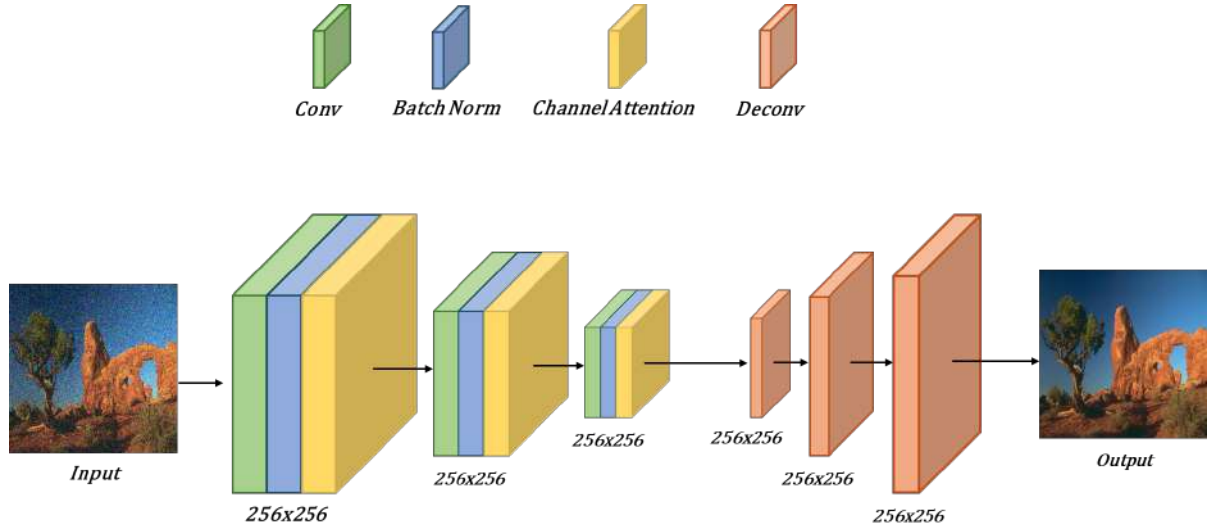

Figure 1. Architecture

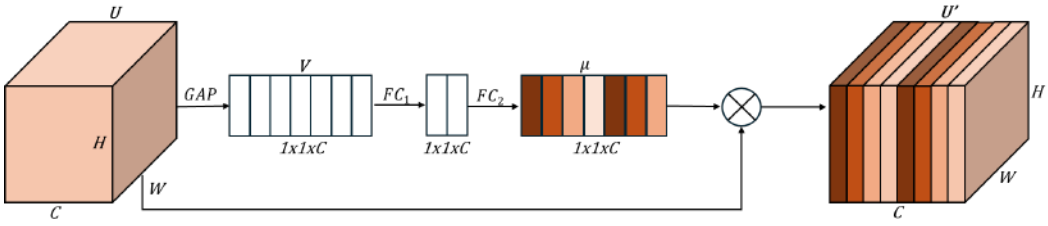

Figure 2. Channel Attention Module

- Deconvolution Layer: After the convolution layer, we use a sequence of deconvolution layers to generate an output image that has the same dimensions as the input. The purpose of the deconvolution layers is to restore the information that was lost during the convolution process.
- Residual Connections: Residual connections are employed to expedite the training process and mitigate

TABLE II. ARCHITECTURE DETAIL

Layer	Output Shape	Description
Input Layer	(256, 256, 3)	X1
Convolution Layer	(256, 256, 256)	X5
Batch Normalization	(256, 256, 256)	
Activation	(256, 256, 256)	
Channel Attention	(256, 256, 256)	
Dropout	(256, 256, 256)	
Deconvolution Layer	(256, 256, 256)	X5
Final Deconvolution Layer	(256, 256, 3)	X1
Final Add	(256, 256, 3)	X1

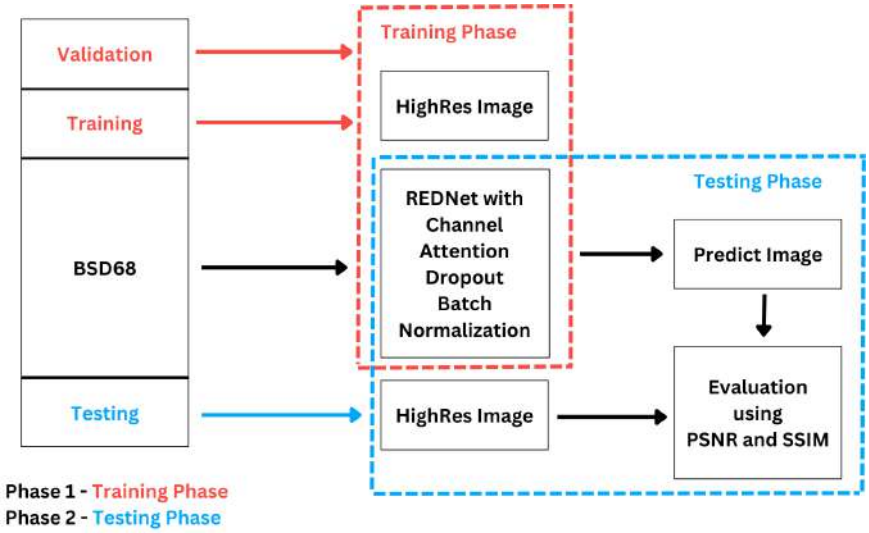

Figure 3. Block Diagram

TABLE III. THE RESULT OF NEW METHODS PERFORMANCE IN PSNR AND SSIM

Method	σ = 15		σ = 25		σ = 50	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
REDNet [19]	32.6284	0.8843	29.6832	0.8483	26.6769	0.7470
MWCNN [20]	30.5043	0.8836	27.8058	0.7678	25.5968	0.6323
PRIDNet [21]	32.5955	0.8951	28.4842	0.8308	25.6192	0.7157
SUNet [22]	31.9932	0.9050	29.4074	0.809	25.3035	0.681
ours	**34.9858**	**0.9371**	**31.7886**	**0.8876**	**27.9063**	**0.7754**

substantial information loss during backward propagation. Residual connections link the output of the convolution layer to the output of the deconvolution layer.

- Output Layer: The output layer merges the input image and the deconvolved image using an addition function. The outcome is the ultimate restored image, which is a fusion of the original image and the restored image.

The proposed model, as shown in Table II and depicted in Figure 1, showcases the effectiveness of our method in tackling the difficulties associated with image denoising.

This can be attributed to the impact of the Channel Attention Module, which recalibrates the importance of the input noise characteristics based on channel significance. The first phase consists of obtaining global information from the input feature map by employing Global Average Pooling (GAP). This technique enables the creation of channel descriptors by aggregating information from the full feature map. Afterwards, the channel descriptors are inputted into two fully connected (FC) layers and then processed with a sigmoid activation function. This process yields recalibration weights that are then employed to precisely scale the input feature map. Multiplying the input feature map with recalibration weights discretely prioritises key channels and eliminates less relevant channels. This enables the network to adaptively modify the connections between feature channels, hence enhancing the efficiency of noise reduction in the process of removing image noise. Figure 2 provides a visualisation of the Channel attention layer.The block diagram of the process can be seen in Figure 3.

D. Evaluation

This study employed well-established criteria, namely Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), to objectively evaluate the efficacy of denoising algorithms. PSNR is a widely used metric in image processing that assesses the quality of denoised image by comparing the maximum power of the signal to the size of the noise impacting it. The PSNR formula, denoted by (2), computes the logarithmic ratio between the squared maximum pixel value (MAX) and the Mean Squared Error (MSE) of the original and denoised image. PSNR measures the level of noise relative to the target signal, where higher PSNR values indicate superior image fidelity.

$$PSNR = 10log_{10}(\frac{MAX^2}{MSE}) \tag{2}$$

The Mean Squared Error (MSE) is a crucial factor in the computation of the Peak Signal-to-Noise Ratio (PSNR). It quantifies the average squared difference between matching pixels in the original and denoised image. It quantifies the reconstruction error, which reflects the degree of similarity between the denoised image and the original image. The Mean Squared Error (MSE) formula, denoted by (3), calculates the sum of the squared differences between corresponding pixels in the original image $I_{original}$ and the denoised image $I_{denoised}$. This sum is then divided by the total number of pixels (N).

$$MSE = \frac{1}{n}\sum_{i=1}^{N}(I_{original} - I_{denoised})^2 \tag{3}$$

Conversely, SSIM offers a perceptual measure for evaluating the similarity of images by considering factors such as brightness, contrast, and structure. The SSIM formula, represented by (4), computes the similarity index by comparing the local pixel intensity patterns of the original (x) and denoised (y) images. The variables μ_x, μ_y, σ_x, and σ_y denote the average values and standard deviations of pixel intensities and their cross-covariance, respectively. The SSIM formula incorporates constants C_1 and C_2 to provide numerical stability and account for variations in signal scale.

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{4}$$

These measurements give vital insights into the success of denoising algorithms, providing a thorough assessment framework for analyzing performance and directing future improvements. PSNR and SSIM not only measure the objective quality of a image, but also assess its perceptual similarity. This helps in comprehensively evaluating the effectiveness of denoising techniques and in advancing image processing technology.

4. RESULT AND DISCUSSION

This section provides the outcomes of assessing a denoising model specifically designed to tackle image degradation caused by Gaussian noise. The model we propose integrates the REDNet architecture with an Attention Channel Module and incorporates Batch Normalization and Dropout methods. The evaluation centers on the model's performance across different degrees of noise intensity, represented by the parameter σ (standard deviation).

Table II presents a thorough overview of the model's performance metrics, encompassing Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM). Significantly, the suggested approach consistently surpasses previous denoising algorithms across various degrees of noise. At a standard deviation (σ) of 15, the model has outstanding performance in terms of Peak Signal-to-Noise Ratio (PSNR), suggesting its capability to maintain visual quality even when subjected to relatively low levels of noise. As the noise levels climb to σ = 25 and σ = 50, the model consistently demonstrates its superiority by achieving higher PSNR values. This highlights its ability to effectively manage increasing noise intensities and generating images that are cleaner and more detailed.

The suggested model obtains a Peak Signal-to-Noise Ratio (PSNR) of 34.9858 decibels and a Structural Similarity Index (SSIM) of 0.9371 at a standard deviation (σ) of 15. When the value of σ is 25, the Peak Signal-to-Noise Ratio (PSNR) is 31.7886 decibels, and the Structural Similarity Index (SSIM) is 0.8876. At a standard deviation (σ) of 50, the Peak Signal-to-Noise Ratio (PSNR) is 27.9063 decibels (dB), and the Structural Similarity Index (SSIM) is 0.7754. The findings demonstrate the exceptional efficacy of our suggested approach in removing noise from images

Figure 4. Result on σ 15

Figure 5. Result on σ 25

Figure 6. Result on σ 50

at different degrees of noise intensity.

The significant improvement observed in the performance of our model may be due to the intentional integration of three crucial components: the Attention Channel Module, Batch Normalization, and Dropout techniques. The Attention Channel Module is essential for preserving image quality at different levels of noise intensity, as seen by the constant rise in PSNR values. This module enhances

Figure 7. Result on σ 15

Figure 8. Result on σ 25

Figure 9. Result on σ 50

the model's ability to prioritize crucial information in the image, reducing the impact of noise on less essential channels and enhancing the clarity of depicted features. In addition, the use of Batch Normalization promotes stability during the training process, speeds up the rate at which the model learns, and assures a dependable and consistent representation of features even when the level of noise varies. Similarly, the incorporation of Dropout

strategies greatly decreases overfitting, allowing the model to efficiently generalize to test data, even in situations with varying amounts of noise.

Figures 4, 5, 6, 7, 8, and 9 give visual evidence of the model's exceptional ability to maintain detailed features and clarity in the presence of noise, thereby enhancing understanding. The results clearly demonstrate the efficacy of our proposed model in comparison to previous denoising techniques, highlighting its capacity to greatly enhance image quality across all levels of noise. Our model demonstrates significant enhancements in denoising effectiveness when compared to baseline models like REDNet, MWCNN, PRIDNet, and SUNet. This is evident by the increased PSNR and SSIM values attained across various noise intensities.

Specifically, Figure 4 demonstrates the enhanced visual quality of our model, achieving a PSNR (Peak Signal-to-Noise Ratio) of 34.3748 and SSIM (Structural Similarity Index) of 0.94 when subjected to a noise level of 15. Figure 5 shows that our model is effective, achieving a PSNR of 31.1618 and SSIM of 0.8852 at a noise level of 25. Figure 6 exhibits its performance with a PSNR of 27.3008 and SSIM of 0.7857 at a noise level of 50. Furthermore, Figures 7,8, and 9 provide additional evidence of the exceptional performance of our model across a broader spectrum of noise levels, as indicated by the corresponding PSNR values of 36.1657, 33.1994, and 30.0704, and SSIM values of 0.9189, 0.87, and 0.7877.

The thorough assessment, supported by visual proof and quantitative measurements, confirms the substantial progress made by our suggested denoising method. The combination of the Attention Channel Module, Batch Normalization, and Dropout techniques results in a very effective model for various denoising image processing tasks. This section highlights not only the quantitative performance evaluated by PSNR and SSIM, but also underlines the crucial significance of the Attention Channel Module in generating excellent denoising results. In summary, the findings demonstrate a significant improvement in image quality using the suggested denoising model, which further confirms its effectiveness in real-world scenarios.

5. CONCLUSION AND FUTURE WORKS

Our research aims to address the complex task of regulating different levels of noise severity in photographic images using sophisticated image denoising algorithms. In order to tackle this difficulty, we suggest a new denoising model that incorporates advanced approaches such as the Channel Attention Module, Batch Normalization, and Dropout procedures. The combination of these components works together to improve the model's performance and durability, resulting in better denoising results.

The Channel Attention Module significantly improves the model's capacity to adapt to varying levels of noise by selectively emphasizing important visual attributes and reducing the influence of noise on less relevant elements. The selective attention method efficiently amplifies the distinctness of characteristics in denoised images, therefore enhancing the overall quality of the image. Furthermore, Batch Normalization guarantees robust model training by standardizing activations across the network, promoting quicker convergence and maintaining consistent feature representations across different levels of noise. Moreover, the use of Dropout regularization improves the model's resilience by reducing overfitting and boosting its ability to generalize, especially in situations with varying degrees of noise.

The model we present has exceptional performance and robustness over a wide range of noise levels, outperforming earlier methods for denoising. At a standard deviation (σ) of 15, the model demonstrates a peak signal-to-noise ratio (PSNR) of 34.9858 decibels (dB) and a structural similarity index (SSIM) of 0.9371. These results emphasize the model's remarkable capability to maintain image quality, even when subjected to relatively low levels of noise. As the noise levels increase to $\sigma = 25$ and $\sigma = 50$, the model continually surpasses other methods by reaching higher PSNR values, demonstrating its ability to efficiently handle increasing noise intensities and provide clearer and more detailed images.

This thorough assessment highlights the effectiveness and importance of our proposed denoising approach, providing significant insights into its superiority compared to existing techniques. The careful incorporation of sophisticated approaches and procedures not only improves the performance of denoising but also establishes the foundation for future improvements in image denoising research and applications.

Recommendations for future research include acquiring a more profound comprehension of this obstacle through the examination of alternate approaches apart from Channel Attention. Furthermore, it is important to take into account the utilisation of pruning or quantization methods to decrease the intricacy of the model while maintaining the quality of denoising. Enhancing computing efficiency in denoising models is a significant research objective, which necessitates the advancement of more effective learning algorithms. Therefore, it is anticipated that this endeavour can facilitate the progress of denoising models that are more efficient and can be feasibly deployed in diverse settings.

6. ACKNOWLEDGMENT

The authors express their gratitude to the individuals who collected the BSD dataset and developed the REDNet, MWCNN, PRIDNet, SUNet, NSTBNet, DBDIP, and DDT algorithms. These contributions were crucial in supporting our study on Image Denoising.

References

[1] L. Fan, F. Zhang, H. Fan, and C. Zhang, "Brief review of image denoising techniques," *Visual Computing for Industry, Biomedicine, and Art*, vol. 2, no. 1, Jul. 2019. [Online]. Available: http://dx.doi.org/10.1186/s42492-019-0016-7

[2] V. Ananthi and P. Balasubramaniam, "A new image denoising method using interval-valued intuitionistic fuzzy sets for the removal of impulse noise," *Signal Processing*, vol. 121, p. 81–93, Apr. 2016. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2015.10.030

[3] A. Awad, "Denoising images corrupted with impulse, gaussian, or a mixture of impulse and gaussian noise," *Engineering Science and Technology, an International Journal*, vol. 22, no. 3, p. 746–753, Jun. 2019. [Online]. Available: http://dx.doi.org/10.1016/j.jestch.2019.01.012

[4] B. Smolka, D. Kusnik, and K. Radlak, "On the reduction of mixed gaussian and impulsive noise in heavily corrupted color images," *Scientific Reports*, vol. 13, no. 1, Nov. 2023. [Online]. Available: http://dx.doi.org/10.1038/s41598-023-48036-1

[5] E. Goceri, "Evaluation of denoising techniques to remove speckle and gaussian noise from dermoscopy images," *Computers in Biology and Medicine*, vol. 152, p. 106474, Jan. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.compbiomed.2022.106474

[6] Z. Chen, Z. Zeng, H. Shen, X. Zheng, P. Dai, and P. Ouyang, "Dn-gan: Denoising generative adversarial networks for speckle noise reduction in optical coherence tomography images," *Biomedical Signal Processing and Control*, vol. 55, p. 101632, Jan. 2020. [Online]. Available: http://dx.doi.org/10.1016/j.bspc.2019.101632

[7] J. Liu, C. Li, L. Liu, H. Chen, H. Han, B. Zhang, and Q. Zhang, "Speckle noise reduction for medical ultrasound images based on cycle-consistent generative adversarial network," *Biomedical Signal Processing and Control*, vol. 86, p. 105150, Sep. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.bspc.2023.105150

[8] C.-T. Lu and T.-C. Chou, "Denoising of salt-and-pepper noise corrupted image using modified directional-weighted-median filter," *Pattern Recognition Letters*, vol. 33, no. 10, p. 1287–1295, Jul. 2012. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2012.03.025

[9] D. N. H. Thanh, L. T. Thanh, N. N. Hien, and S. Prasath, "Adaptive total variation l1 regularization for salt and pepper image denoising," *Optik*, vol. 208, p. 163677, Apr. 2020. [Online]. Available: http://dx.doi.org/10.1016/j.ijleo.2019.163677

[10] D. N. Thanh, N. N. Hien, P. Kalavathi, and V. S. Prasath, "Adaptive switching weight mean filter for salt and pepper image denoising," *Procedia Computer Science*, vol. 171, p. 292–301, 2020. [Online]. Available: http://dx.doi.org/10.1016/j.procs.2020.04.031

[11] A. A. Rafiee and M. Farhang, "A deep convolutional neural network for salt-and-pepper noise removal using selective convolutional blocks," *Applied Soft Computing*, vol. 145, p. 110535, Sep. 2023. [Online]. Available: http://dx.doi.org/10.1016/j.asoc.2023.110535

[12] N. Singh and O. U. Maheswari, "A new denoising algorithm for random valued impulse noise in images using measures of dispersion," Mar. 2017. [Online]. Available: http://dx.doi.org/10.1109/ICSCN.2017.8085737

[13] M. Xing and G. Gao, "An efficient method to remove mixed gaussian and random-valued impulse noise," *PLOS ONE*, vol. 17, no. 3, p. e0264793, Mar. 2022. [Online]. Available: http://dx.doi.org/10.1371/journal.pone.0264793

[14] D.-H. Shin, R.-H. Park, S. Yang, and J.-H. Jung, "Block-based noise estimation using adaptive gaussian filtering," *IEEE Transactions on Consumer Electronics*, vol. 51, no. 1, p. 218–226, Feb. 2005. [Online]. Available: http://dx.doi.org/10.1109/TCE.2005.1405723

[15] F. Russo, "A method for estimation and filtering of gaussian noise in images," *IEEE Transactions on Instrumentation and Measurement*, vol. 52, no. 4, p. 1148–1154, Aug. 2003. [Online]. Available: http://dx.doi.org/10.1109/TIM.2003.815989

[16] M. Mafi, H. Martin, M. Cabrerizo, J. Andrian, A. Barreto, and M. Adjouadi, "A comprehensive survey on impulse and gaussian denoising filters for digital images," *Signal Processing*, vol. 157, p. 236–260, Apr. 2019. [Online]. Available: http://dx.doi.org/10.1016/j.sigpro.2018.12.006

[17] C. Tian, Y. Xu, L. Fei, and K. Yan, *Deep Learning for Image Denoising: A Survey*. Springer Singapore, 2019, p. 563–572. [Online]. Available: http://dx.doi.org/10.1007/978-981-13-5841-8_59

[18] D. Martin, C. Fowlkes, D. Tal, and J. Malik, "A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, ser. ICCV-01. IEEE Comput. Soc.

[19] X.-J. Mao, C. Shen, and Y.-B. Yang, "Image restoration using convolutional auto-encoders with symmetric skip connections," 2016. [Online]. Available: https://arxiv.org/abs/1606.08921

[20] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-cnn for image restoration," Jun. 2018. [Online]. Available: http://dx.doi.org/10.1109/CVPRW.2018.00121

[21] Y. Zhao, Z. Jiang, A. Men, and G. Ju, "Pyramid real image denoising network," Dec. 2019. [Online]. Available: http://dx.doi.org/10.1109/VCIP47243.2019.8965754

[22] C.-M. Fan, T.-J. Liu, and K.-H. Liu, "Sunet: Swin transformer unet for image denoising," May 2022. [Online]. Available: http://dx.doi.org/10.1109/ISCAS48785.2022.9937486

[23] Z. Lyu, Y. Chen, Y. Hou, and C. Zhang, "Nstbnet: Toward a nonsubsampled shearlet transform for broad convolutional neural network image denoising," *Digital Signal Processing*, vol. 123, p. 103407, Apr. 2022. [Online]. Available: http://dx.doi.org/10.1016/j.dsp.2022.103407

[24] S. Xu, X. Cheng, J. Luo, N. Xiao, M. Xiong, and C. Zhou, "Dual-branch deep image prior for image denoising," *Journal of Visual Communication and Image Representation*, vol. 93, p. 103821, May 2023. [Online]. Available: http://dx.doi.org/10.1016/j.jvcir.2023.103821

[25] K. Liu, X. Du, S. Liu, Y. Zheng, X. Wu, and C. Jin, "Ddt: Dual-branch deformable transformer for image denoising," in *2023 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, Jul. 2023.

Rico Halim is a Graduate Computer Science Student at Bina Nusantara University. His research interests include Data Science, Machine Learning, Computer Vision, and Image Processing.

Gede Putra Kusuma received PhD degree in Electrical and Electronic Engineering from Nanyang Technological University (NTU), Singapore, in 2013. He is currently working as a Lecturer and Head of Department of Master of Computer Science, Bina Nusantara University, Indonesia. Before joining Bina Nusantara University, he was working as a Research Scientist in I2R – A*STAR, Singapore. His research interests include computer vision, deep learning, face recognition, appearance-based object recognition, gamification of learning, and indoor positioning system.