Starlight: An Image Enhancement System By Reducing Noise Using Deep Learning

Nihal B
Department of CSE
MES College of Engineering, Kuttippuram
APJ Abdul Kalam Technological University
Kerala, India
nihalbilal616@gmail.com

Nithin S Varrier
Department of CSE
MES College of Engineering, Kuttippuram
APJ Abdul Kalam Technological University
Kerala, India
nithin486@hotmail.com

Shabeeb M
Department of CSE
MES College of Engineering, Kuttippuram
APJ Abdul Kalam Technological University
Kerala, India
shabeebmattummal@gmail.com

Sreedev K
Department of CSE
MES College of Engineering, Kuttippuram
APJ Abdul Kalam Technological University
Kerala, India
sreedevz346@gmail.com

Badharudheen P
Department of CSE
MES College of Engineering, Kuttippuram
APJ Abdul Kalam Technological University
Kerala, India
badharu@gmail.com

***Abstract*—Digital images are often plagued by noise, degrading visual quality and hindering accurate analysis. Traditional noise removal methods were tedious, impractical, and often ineffective, especially for complex noise patterns. This report presents a deep learning-based solution using convolutional neural networks (CNNs) to address these limitations. Our approach leverages the power of deep learning by training specialized CNN models on diverse datasets containing various noise types. These models can effectively distinguish noise from meaningful image features, removing noise while preserving important details. For maximum accessibility and practical usability, we integrated the trained models into a user-friendly web application. This application provides an intuitive interface for image capture, noise detection, and seamless denoising. By automating the noise removal process and offering an accessible interface, our solution has the potential to significantly improve digital image quality and facilitate accurate image analysis across various applications, including photography, medical imaging, scientific research, and robotics.**

***Index Terms*—Image denoising, deep learning, convolutional neural networks, noise removal, attention mechanisms, web application, real-time denoising**

I. Introduction

In today's digital era, the demand for efficient image processing techniques has surged significantly. With the widespread use of high-resolution imaging devices, the sheer volume of image data has skyrocketed. However, this wealth of visual information often gets marred by noise during the process of acquisition and transmission, severely compromising image quality and utility. Hence, there's an current need for robust image denoising methods capable of handling both synthetic and real noise.

Addressing this challenge isn't straightforward. Real-world noise comes with uncertainties regarding its level and type, alongside its spatially variant and signal-dependent nature. Moreover, preserving intricate details and textures while denoising adds another layer of complexity. Unfortunately, existing methods often fall short in adequately addressing these hurdles, leading to subpar results. This results in a need for more better methods to solve these issues

Our proposed approach revolves around leveraging deep learning for real image denoising. This means using a deep-learning based CNN model that could make use of techniques like feature attention.

To put our solution into action, we modify a existing model to improve it's use for this particular usecase. We train our modified model using a mix of synthetic and real noisy images, employing an L1 loss function.

Considering the prevalent utilization of web-based applications and the escalating demand for online image processing services, the integration of a web application manifestation of our system appears pragmatically sound. Such implementation would afford users access to our image denoising service from any internet-enabled device, furnishing a convenient and user-centric solution. The merits of our proposed solution are manifold. Firstly, our methodology adeptly manages both synthetic and real noise within a single-blind model, thus surpassing the versatility of extant methodologies. Secondly, our model attains superior performance in terms of quantitative metrics and perceptual quality, thereby ensuring the delivery of high-fidelity denoising outcomes. Lastly, through the deployment of our solution as a web application, we endeavor to furnish a convenient and universally accessible service to a diverse user base. In summation, our proposed solution epitomizes a robust and efficacious approach to real image denoising, possessing significant potential for practical deployment across various domains.

II. Literature Survey

Image denoising is a fundamental problem in image processing and computer vision, aiming to recover a clean image from its noisy counterpart. Over the years, various techniques have been proposed to address this challenging task. One of the earliest and most influential works in this field is the total variation (TV) regularization method introduced by Rudin et al. [1]. This approach exploits the statistical property that natural images are locally smooth, with pixel intensities gradually varying in most regions. The TV regularization model can effectively calculate the optimal solution while preserving sharp edges. However, it suffers from several drawbacks, including over-smoothing of textures, stair-casing artifacts, and loss of contrast.

To overcome the limitations of local methods, researchers have explored the non-local self-similarity (NSS) prior in natural images. The pioneering work in this direction is the non-local means (NLM) algorithm proposed by Buades et al. [2]. NLM calculates the denoised value of a pixel as a weighted average of pixels within a search window, with weights determined by the similarity between neighboring patches. This non-local approach can effectively reduce noise while preserving edges. However, it may struggle to recover fine details and tiny structures, especially in the presence of high noise levels.

Another notable approach is the sparse representation model, exemplified by the K-SVD algorithm proposed by Aharon et al. [3]. These methods encode an image over an over-complete dictionary with L1-norm sparsity regularization. By learning the dictionary from the noisy image itself, the sparse representation model can flexibly represent image structures. However, these methods are primarily local and may not perform well under high noise levels.

The low-rank minimization approach, such as the weighted nuclear norm minimization (WNNM) method proposed by Gu et al. [4], exploits the low-rank property of similar image patches stacked as a matrix. WNNM can adaptively assign weights to singular values and denoise them using a soft thresholding technique. While WNNM and other low-rank methods demonstrate advanced denoising performance, the iterative boosting step can be computationally expensive .

In the transform domain, the Block-matching and 3D filtering (BM3D) algorithm proposed by Dabov et al. [5] is a powerful extension of the NLM approach. BM3D groups similar patches into 3D arrays, transforms them into the wavelet domain, and performs collaborative filtering. Despite its effectiveness, BM3D's denoising performance may degrade as the noise level increases, and it can introduce artifacts in flat areas.

With the recent advances in deep learning, convolutional neural networks (CNNs) have emerged as a powerful tool for image denoising. The DnCNN model, introduced by Zhang et al. [6], is a pioneering work in this area. It introduces residual learning and batch normalization to train a feed-forward denoising CNN. DnCNN achieves state-of-the-art performance and can handle various types of noise. However, its trained model is specific to a fixed noise level, limiting its flexibility.

To address this limitation, Zhang et al. [7] proposed the Fast and Flexible Denoising Convolutional Neural Network (FFDNet). FFDNet treats the noise level map as an additional input, making it flexible and adaptive to different noise levels. It also employs a down-sampling strategy to increase the receptive field and speed up the training and testing processes. While FFDNet demonstrates impressive denoising performance and flexibility, its training process can be computationally demanding.

Another notable CNN-based denoising method is the Residual Dense Network (RDN) proposed by Zhang et al. [8]. RDN combines the strengths of residual learning and dense connections, exploiting the persistent memory of deep neural networks to better capture multi-scale information. By effectively modeling the hierarchical features of natural images, RDN achieves superior denoising performance compared to previous CNN-based methods.

Recently, RIDNet, an advanced deep learning architecture proposed by Zhang et al. [9], stands out among convolutional neural network-based denoising methods for its remarkable performance in removing noise while preserving essential image details. By integrating residual learning and dense connections, RIDNet effectively captures multi-scale information, enabling it to handle various noise levels and types. By incorporating residual blocks within residual blocks, RIDNet demonstrates superior denoising capabilities compared to previous approaches. This model represents a significant advancement in image denoising, addressing the limitations of traditional techniques and showcasing the potential of deep learning in tackling complex real-world challenges.

III. Methodologies

A. RIDNet

RIDNet employs an efficient architecture combining residual learning and inception blocks with dilated convolutions. The inception blocks facilitate the extraction of multi-scale features, while dilated convolutions expand the receptive field efficiently. It achieves state-of-the-art performance in recovering fine textures and details from high noise levels compared to previous CNN denoisers [9]. Its architecture comprises multiple inception-residual blocks, each containing parallel paths with dilated convolutions of varying rates. These paths are then concatenated and passed through residual connections, allowing RIDNet to capture contextual information effectively for denoising.

An analysis of the applied method showed that it was effective in denoising several kinds of noisy images, producing significant gains in quantitative metrics like PSNR and SSIM scores as well as qualitative evaluations. The fact that crucial image elements and structures were successfully preserved in spite of difficult real-world circumstances highlights the applicability of our method in a variety of fields, such as robotics, scientific research, medical imaging, and photography [10]. Our method offers a substantial leap in picture denoising, promising improved image quality and more precise image analysis by effectively reducing noise while maintaining critical image properties [6].

B. Peak Signal-to-Noise Ratio (PSNR)

To quantitatively evaluate the performance of the proposed system, we employed two widely-used metrics in the field of image denoising and restoration: Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM).

PSNR (Peak Signal-to-Noise Ratio) is a metric used to assess the quality of a reconstructed image, particularly after noise reduction. It estimates the peak error between the denoised image and the original, noise-free image (ground truth) by measuring the ratio between the maximum possible signal power and the corrupting noise power [11]. Mathematically, PSNR is calculated using the mean squared error (MSE) as shown below:

$$\mathrm{PSNR} = 10 \times \log_{10}\left(\frac{\mathrm{MAX\_I}^2}{\mathrm{MSE}}\right) \tag{1}$$

Where *MAX_I* is the maximum possible pixel value of the image (e.g., 255 for 8-bit images), and *MSE* is the mean squared error between the denoised and ground truth images, calculated as:

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=0}^{m}\sum_{j=0}^{n}[I(i,j) - K(i,j)]^2 \tag{2}$$

Here, m and n are the image dimensions, $I(i,j)$ is the denoised image, and $K(i,j)$ is the ground truth clean image.

Higher PSNR values indicate better denoising quality. This is because a higher PSNR signifies a larger peak signal value compared to the noise power, meaning the reconstructed image has a stronger signal relative to the noise. Due to its ease of interpretation, PSNR is widely used as a measure of reconstruction quality in image denoising tasks [12].

However, PSNR alone may not fully capture perceptual similarity, as it is based solely on pixel-wise differences. Therefore, we also employ the Structural Similarity Index (SSIM) as an additional evaluation metric.

C. Structural Similarity Index (SSIM)

While PSNR provides a measure of pixel-wise differences, the Structural Similarity Index (SSIM) [13] is a perception-based metric that considers image degradation as perceived change in structural information. SSIM evaluates the similarity between two images in terms of luminance, contrast, and structure, providing a more comprehensive assessment of image quality [14].

The SSIM index between two image signals x and y is defined as:

$$\mathrm{SSIM}(x,y) = [l(x,y)]^{\alpha} * [c(x,y)]^{\beta} * [s(x,y)]^{\gamma} \tag{3}$$

Where: $l(x,y)$ is the luminance comparison function, $c(x,y)$ is the contrast comparison function, $s(x,y)$ is the structure comparison function, α, β, and γ are parameters that determine the relative importance of the three components.

These comparison functions are calculated locally within a window around each pixel, and then combined using a weighted averaging approach to yield the overall SSIM index [13].

SSIM values range from -1 to 1, with a value of 1 indicating perfect structural similarity between the two images. Higher SSIM scores represent better denoising performance, as they signify greater preservation of structural information and perceptual quality [15].

By incorporating SSIM alongside PSNR, we can comprehensively evaluate the performance of RIDNet in terms of both pixel-wise accuracy and perceptual similarity, ensuring a thorough assessment of the proposed denoising approach across various image characteristics and noise types.

The complementary use of PSNR and SSIM allows us to capture different aspects of image quality, providing a more reliable and holistic evaluation of our denoising system's effectiveness in real-world scenarios.

IV. Proposed System

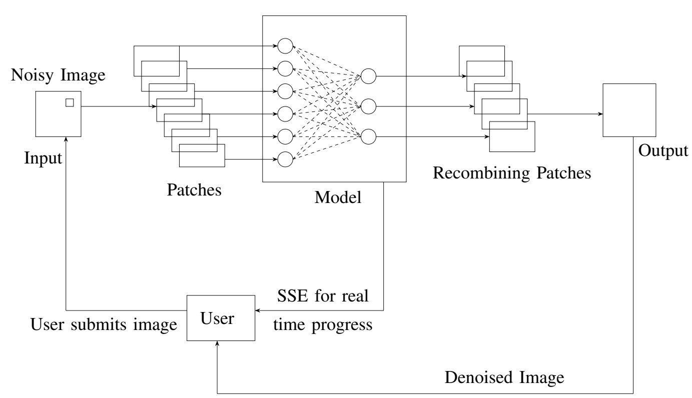

Fig. 1. Proposed System

To address the limitations of existing image denoising methods and provide a robust solution for real-world scenarios, we propose a novel deep learning-based system leveraging the state-of-the-art RIDNet architecture [9]. Fig. 1 shows our system proposed which incorporates several key components to achieve effective and efficient image denoising:

1) Reactive user interface
2) RIDNet Model with modifications
3) Realtime Progress using Server Sent Events

A. RIDNet Architecture

The core of the proposed system is the RIDNet (Real Image Denoising Network with Feature Attention) architecture, a convolutional neural network specifically designed for image-denoising tasks. The RIDNet architecture consists of the following modules:

1) Feature Extraction Module: The Feature Extraction Module extracts initial low-level features such as edges, textures, and patterns from the noisy input image using a convolutional layer [9].

2) Feature Learning Residual-on-Residual Module: The Feature Learning Residual-on-Residual Module is the core component of RIDNet, consisting of Enhancement Attention Modules (EAMs) cascaded together. Unlike the original architecture, which utilized four EAMs, our modified RIDNet incorporates an additional EAM to enhance the denoising performance further. Each EAM incorporates a novel residual-on-residual structure with local and global skip connections, facilitating the flow of low-frequency information and enabling the network to learn residual representations more effectively [9]. Additionally, to optimize GPU memory usage, we introduced a modification in the input processing stage, where we divide the input image into overlapping patches of size 128x128. These patches are then denoised separately and recombined to generate the final output, allowing for efficient utilization of GPU resources while maintaining denoising efficacy.

3) Reconstruction Module: The Reconstruction Module consists of a single convolutional layer that maps the enhanced features from the residual-on-residual module to produce the denoised output image [9].

B. Training Procedure

Our System's RIDNet model is trained in an end-to-end fashion on a combination of synthetic and real noisy image datasets using the following techniques:

- Optimizer: Adam optimizer [16] for efficient training convergence, dynamically adapting the learning rates for each parameter.
- Loss Function: Structure Similarity Index (SSIM) loss [13] to optimize the model for perceptual similarity between the denoised output and the ground truth clean image.In this context the Structure Similarity is used as a loss function. The loss function is calculated as :

$$\mathrm{LSSIM}(y_{\text{true}}, y_{\text{pred}}) = 1 - \frac{1}{N}\sum_{i=1}^{N}\mathrm{SSIM}(y_{i,\text{true}}, y_{i,\text{pred}}) \tag{4}$$

  Where: In Eq. 4 , y_{pred} is the predicted value by the model and y_{true} represents the ground truth.

C. Evaluation Metrics

To quantitatively evaluate the performance of our system's proposed RIDNet model, we employ two widely-used metrics in the field of image denoising and restoration:

1) Peak Signal-to-Noise Ratio (PSNR): Peak Signal-to-Noise Ratio (PSNR) measures the ratio between the maximum possible signal power and the noise power, providing an estimate of the peak error between the denoised image and the ground truth clean image. Higher PSNR values indicate better denoising quality [12], [17].

2) Structural Similarity Index (SSIM): Structural Similarity Index (SSIM) is a perception-based metric that evaluates the similarity between two images in terms of luminance, contrast, and structure. Higher SSIM scores represent better denoising performance [14], [15].

By leveraging the RIDNet architecture, optimized training procedure, and comprehensive evaluation metrics, Our System aims to provide a robust and effective solution for real-world image denoising tasks, addressing the limitations of existing methods while preserving crucial image details and structures.

D. Objectives of Proposed System

The objectives of our proposed image denoising system are crafted with precision to tackle the core challenges in the domain, ensuring both practicality and effectiveness. These objectives serve as the guiding principles for the development process and delineate the desired outcomes of the system. The main objectives of Starlight are as follows:

1) **Effective Noise Reduction:** The main goal of the model is to develop an image denoising system based on deep learning methods that can effectively reduce noise in digital images. Through the use of advanced convolutional neural network (CNN) designs and training techniques, the system aims to improve images' visual clarity and interpretability.
2) **Robust Performance:** The system aims to achieve robust performance across various types and levels of noise commonly encountered in real-world scenarios. By training the model on diverse datasets containing different noise patterns, the system endeavors to ensure consistent and reliable denoising performance under diverse conditions.
3) **Optimized Computational Efficiency:** In addition to performance, the system places emphasis on optimizing computational efficiency to enable rapid processing of large volumes of image data. Through the implementation of efficient algorithms and leveraging parallel processing capabilities, the system aims to achieve real-time or near-real-time denoising performance, facilitating seamless integration into existing workflows.
4) **User-Friendly Interface:** It places importance on usability and accessibility by offering an interface designed for ease of use, simplifying the denoising process. The system aims to enable users to effortlessly upload noisy images, adjust denoising parameters as needed, and download the cleaned images with minimal effort, thereby enhancing user experience and adoption.
5) **Integration with Existing Workflows:** It aims to seamlessly integrate into existing image processing workflows

and applications, enabling users to incorporate the denoising functionality without significant modifications. This compatibility with commonly used file formats and platforms streamlines the adoption process and maximizes utility.
6) **Scalability and Adaptability:** The system prioritizes a modular architecture with well-defined interfaces. This enables seamless integration of new functionalities and algorithms, promoting code reusability and maintainability. This flexible design ensures the system's long-term adaptability to address emerging challenges in image denoising.
7) **Evaluation and Validation:** Finally, our system prioritizes rigorous evaluation and validation in order to analyze its performance and effectiveness. In order to verify the system's utility in real-world circumstances, it will be subjected to a rigorous testing process leveraging standard image denoising metrics including Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), in along with qualitative evaluations.

By pursuing these objectives, Our system aims to provide a versatile, reliable, and practical solution for image denoising, empowering users across various domains to enhance the quality and usability of their digital images.

E. Advantages of Proposed System

1) **Superior Performance:** The proposed system exhibits superior performance in terms of quantitative metrics and visual quality, as demonstrated through rigorous evaluations conducted in the study.
2) **Accessibility and Practicality:** The integration of trained models into a user-friendly web application ensures maximum accessibility and practical usability. Users can effortlessly upload images, analyze noise patterns, and obtain denoised results, facilitating a broad range of applications.
3) **Effective Noise Removal:** The deep learning model developed in the system adeptly removes noise from various image types while preserving crucial details. This achievement is realized through the utilization of specialized CNN models trained on diverse noise types, effectively discerning noise from meaningful image features.
4) **Convenient Service:** Deploying the solution as a web application offers a convenient and accessible service to users, meeting the escalating demand for online image processing services. This enhances the usability and applicability of the system across different fields.

V. Results and Discussions

Fig. 2 shows the fluctuation of PSNR values over training epochs. The PSNR across epoch's peaked at 32.55, after which our hardware limits were expired. Reaching a plateu on this graph meant that the model could no longer learn anything new from the dataset given and training any longer would a waste of time. Higher PSNR values indicate better

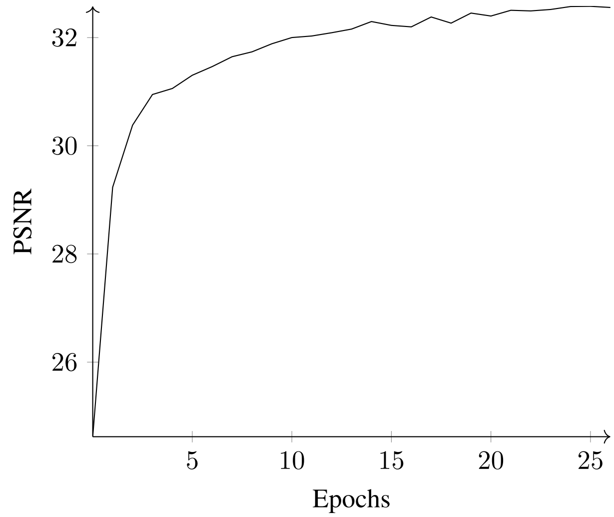

Fig. 2. Plot showing the Peak Signal-to-Noise Ratio (PSNR) values for each epoch during model training.

visual detail preservation and lower noise levels. PSNR is a regularly used metric to evaluate the quality of denoising algorithms. Plotting the denoising performance increase over several epochs gives insights into the model's convergence behavior throughout training. Understanding the efficiency and optimization trajectory of the denoising model during training is made easier with the help of this visualization.

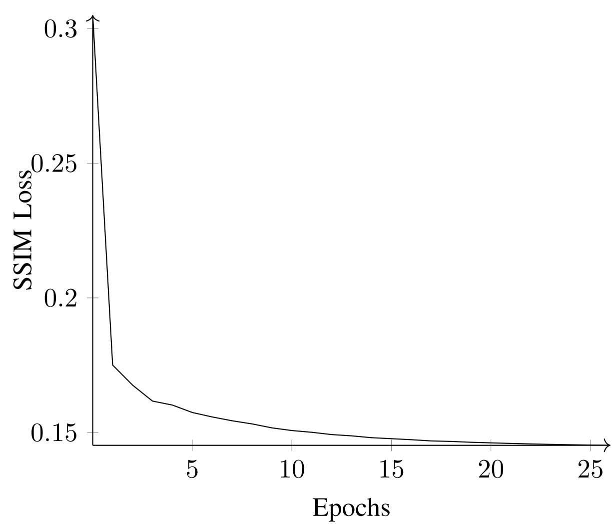

Fig. 3. SSIM loss in each epoch when training the model

Fig. 3 illustrates the SSIM loss variation across training epochs. SSIM loss quantifies the perceptual similarity between denoised and clean images, with lower values indicating improved alignment. The plot tracks the model's optimization progress, showcasing iterative improvement in denoising performance. This graphical representation offers insights into the convergence behavior and efficacy of the denoising model during training. The lowest SSIM loss was 0.145 at epoch 26.

Due to lack of hardware resources, the training was stopped at epoch 26.

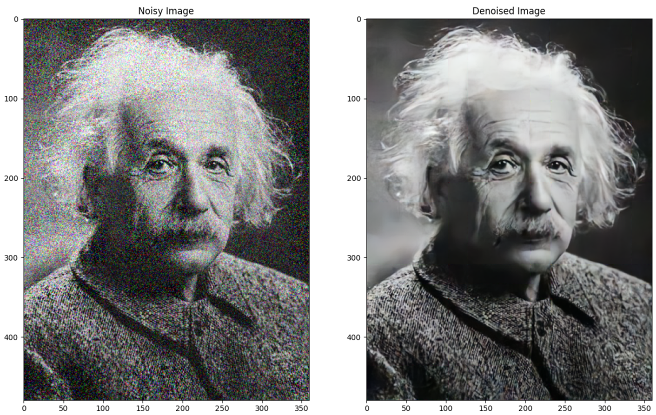

Fig. 4. Noisy and Denoised Image of Albert Einstein

Fig. 4 represents the result of the model when given an image with synthetic noise added to it. The output is generated within seconds on a GPU, showcasing the efficiency of the model. Table I shows the PSNR and SSIM values of the output with respect to the original images.

TABLE I
EVALUATION METRICS

Image Type	Metrics w.r.t ground truth	
	PSNR	***SSIM***
Noisy	16.669977	0.3225038
Denoised	26.90759	0.8338222

Calculated using Tensorflow's PSNR and SSIM [13] implementation

CONCLUSION

In this project, we modified an advanced single-stage blind real image denoising network, alongside evaluation metrics like PSNR and SSIM. The model was developed utilizing Keras, providing a resilient and straightforward training framework, ultimately contributing to a substantial enhancement in the model's performance. The modifications done on the network had better performance and hence used as the core model in the system. Moreover, integration with TensorBoard and Flask enabled efficient model training and deployment. The system's modular design, along with a powerful and efficient deep-learning based model proved instrumental in enhancing denoising performance, while the utilization of SSE ensured reactivity and a better user interaction. Our proposed system not only integrates easily with present systems, but also presents a promising solution for real world image enhancement applications.

REFERENCES

[1] L. I. Rudin, S. Osher, and E. Fatemi, "Nonlinear total variation based noise removal algorithms," *Physica D: Nonlinear Phenomena*, vol. 60, no. 1-4, pp. 259–268, Nov. 1992. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/016727899290242F

[2] A. Buades, B. Coll, and J.-M. Morel, "A Non-Local Algorithm for Image Denoising," in *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, vol. 2. San Diego, CA, USA: IEEE, 2005, pp. 60–65. [Online]. Available: http://ieeexplore.ieee.org/document/1467423/

[3] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311–4322, Nov. 2006. [Online]. Available: https://ieeexplore.ieee.org/document/1710377

[4] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted Nuclear Norm Minimization with Application to Image Denoising," in *2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, OH, USA: IEEE, Jun. 2014, pp. 2862–2869. [Online]. Available: https://ieeexplore.ieee.org/document/6909762

[5] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007. [Online]. Available: https://ieeexplore.ieee.org/document/4271520/

[6] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017. [Online]. Available: https://ieeexplore.ieee.org/document/7839189/

[7] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a Fast and Flexible Solution for CNN-Based Image Denoising," *IEEE Transactions on Image Processing*, vol. 27, no. 9, pp. 4608–4622, Sep. 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8365806/

[8] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual Dense Network for Image Super-Resolution," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Salt Lake City, UT: IEEE, Jun. 2018, pp. 2472–2481. [Online]. Available: https://ieeexplore.ieee.org/document/8578360/

[9] S. Anwar and N. Barnes, "Real Image Denoising with Feature Attention," Mar. 2020, arXiv:1904.07396 [cs]. [Online]. Available: http://arxiv.org/abs/1904.07396

[10] S. Lefkimmiatis, "Non-local color image denoising with convolutional neural networks," 2017.

[11] "Peak signal-to-noise ratio - Wikipedia — en.wikipedia.org," https://en.wikipedia.org/wiki/Peak_signal-to-noise_ratio, [Accessed 04-05-2024].

[12] A. Horé and D. Ziou, "Image quality metrics: Psnr vs. ssim," in *2010 20th International Conference on Pattern Recognition*, Aug 2010, pp. 2366–2369.

[13] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, Apr. 2004. [Online]. Available: http://ieeexplore.ieee.org/document/1284395/

[14] R. Dosselmann and X. Yang, "A comprehensive assessment of the structural similarity index," *Signal, image and video processing*, vol. 5, no. 1, pp. 81–91, 11 2009. [Online]. Available: https://doi.org/10.1007/s11760-009-0144-1

[15] Z. Wang and A. C. Bovik, "Mean squared error: Love it or leave it? a new look at signal fidelity measures," *IEEE Signal Processing Magazine*, vol. 26, no. 1, pp. 98–117, Jan 2009.

[16] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2017.

[17] A. C. Bovik, "Automatic prediction of perceptual image and video quality," *Proceedings of the IEEE*, vol. 101, no. 9, pp. 2008–2024, Sep. 2013.