Improving Diabetic Retinopathy Fundus Image Quality Using Dynamic Sparse Representation and Dilated Convolutional Networks

Veerendra Basreddy
Department of Computer Science and Engineering, Faculty of Engineering and Technology (Coed), Sharnbasva University, Kalaburagi, India
veerendrabasreddy@gmail.com (corresponding author)

Sachinkumar Veerashetty
Department of Computer Science and Design, Faculty of Engineering and Technology (Coed), Sharnbasva University, Kalaburagi, India
sveerashetty123@gmail.com

Received: 23 September 2025 | Revised: 17 December 2025 | Accepted: 29 December 2025

 DOI: https://doi.org/10.48084/etasr.15050

ABSTRACT

In recent years, there has been an increase in Diabetic Retinopathy (DR). In addition, the images used for the detection of DR are not of good quality. This study focuses on improving the quality of DR fundus images by proposing an image enhancement approach. Specifically, this work presents a Dynamic Sparse Representation (DSR) approach that combines Convolution layers (Conv), Rectified-Linear-Unit (ReLU), and Dilated-Convolution methods for preprocessing. To further improve image quality, DSR utilizes a dictionary approach during the training phase. Before testing, Additive White Gaussian Noise (AWGN) is induced in the input images to simulate real-world noise conditions. In the testing phase, noisy images are processed to remove noise and enhance image quality. The proposed DSR method was evaluated on the MESSIDOR and EyeQ datasets using the following metrics: Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Mean Squared Error (MSE). For the EyeQ and MESSIDOR datasets, DSR achieved 37.82 PSNR, 0.94 SSIM, 0.0316 MSE, and 36.85 PSNR, 0.93 SSIM, and 0.0431 MSE, respectively. The findings show that the proposed DSR approach achieved better performance compared to other approaches.

Keywords-diabetic retinopathy; fundus images; image enhancement; convolutional neural network; rectified linear unit

I. INTRODUCTION

Diabetes is a chronic disease with harmful consequences throughout the body, but its most severe impact is often observed in the visual system, specifically in the retina. High blood sugar levels damage the blood vessels that supply the retina, leading to Diabetic Retinopathy (DR), a condition that can cause severe vision loss. From 2012 to 2021, the number of adults diagnosed with diabetes increased dramatically, and by 2040 is projected to reach 642 million individuals worldwide [1]. Approximately one-third of these patients are expected to develop DR, highlighting the urgent need for effective screening and timely intervention. Hence, early identification of DR is critical in preventing vision loss, but current diagnostic methods face multiple limitations. Traditional techniques such as Fluorescein Angiography (FA) are expensive, require specialized instruments, and are not always feasible in resource-constrained settings.

Automated DR screening powered by AI, particularly Deep Learning (DL), offers a cost-effective and faster alternative to conventional methods [2]. In recent years, DL-based models have been extensively explored for enhancing, segmenting, and classifying DR images [3-13]. For example, the Clinically-Oriented-Fundus-Enhancement-Network (COFE-Net) [3] enhances retinal images while preserving structural details, achieving Peak Signal-to-Noise Ratio (PSNR) values of 20.51 (EyeQ dataset) and 21.24 (DRIVE dataset). Similarly, a Generative Adversarial Network (GAN) model incorporating convolutional attention reported a PSNR of 24.73 and a Structural Similarity Index Measure (SSIM) of 0.8103 on EyePACS, STARE, CHASEDB1, and proprietary datasets [4]. In [5], neural networks for image enhancement in both Color Fundus (CF) and FA images achieved 93.0% and 89.5% accuracy, respectively. In [6], a Fundus-GAN model enhanced DR images by segmenting and extracting retinal vessels, achieving 98.5% segmentation accuracy on the HRF, STARE, DRIVE, and CHASE-DB1 datasets.

Other studies focused on denoising and classification. In [7], a preprocessing and feature-extraction algorithm was developed to remove noise in CF images, achieving classification accuracies of 92% (EyePACS) and 93.6% (MESSIDOR). In [8], IoT and DL were combined, using May-Fly-Optimization-Region-Growing (MFORG) to enhance images, achieving 96.87% accuracy using DenseNet with LSTM-Honey-Bee-Optimization (HBO) on MESSIDOR. In [9], an approach for enhancement and grading of DR fundus images reported 88.5% and 73.6% accuracy on EyeQ and MESSIDOR, respectively. In [10], an optimized Convolutional Neural Network (CNN) was used to transform low-quality DR images into high-quality ones, achieving PSNR values of 30.57 and 27.28 on datasets from two eye hospitals. In [11], GANs were employed for fundus image enhancement, reporting PSNR of 19.67 and SSIM of 0.76 on EyeQ. The ESDiff framework [12] integrates masking and U-Net for enhancement and segmentation, achieving PSNR of 19.27 and SSIM of 0.864 on EyeQ, along with segmentation accuracies above 97% on DRIVE, STARE, and CHASE-DB1. In [13], Sharpening-Varying-Vessels (SVV) with a variational autoencoder and GAN was used for vessel-level feature retention, achieving SSIM values of 0.9175 (Synthetic-Real) and 0.8398 (Synthetic-Synthetic) on the DRIVE and MESSIDOR-1 datasets.

From these studies, it is clear that most existing works focus on segmentation, classification, and grading, while comparatively fewer emphasize preprocessing, image enhancement, and noise removal, despite being critical for improving downstream DR analysis. Noisy or low-quality images can compromise segmentation and classification accuracy, leading to unreliable diagnostic outcomes. To address this gap, this paper presents a Dynamic Sparse Representation (DSR) approach for preprocessing, noise removal, and enhancement of DR fundus images. The proposed model combines Convolutional layers (Conv), Rectified Linear Units (ReLU), and dilated convolutions for improved feature representation, and employs a dictionary-learning method during training to effectively remove noise. The main contributions of this work are as follows:

- An approach for enhancing the quality of DR fundus images, i.e., the introduction of a DSR technique, combining Conv, ReLU, and dilated convolution for effective image enhancement.
- A dictionary-learning strategy for noise removal during the training phase.
- Comprehensive evaluation using MESSIDOR and EyeQ datasets, where DSR is compared against existing approaches using PSNR, SSIM, and Mean Squared Error (MSE), demonstrating better performance.

II. METHODOLOGY

A. Architecture

Figure 1 presents the flow of the proposed method. The input image is divided into different sections/patches for quality enhancement. The input image patches go through the DSR approach, as shown in Figure 2.

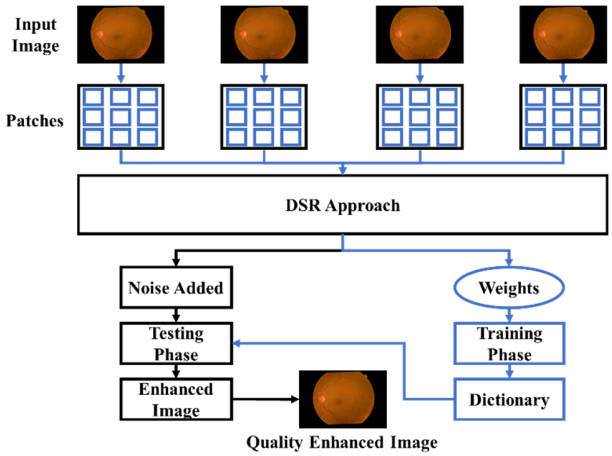


Fig. 1. Flow of the proposed method.

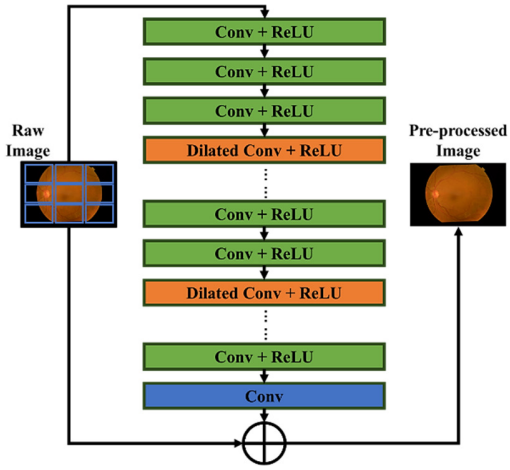


Fig. 2. DSR approach.

In DSR, the image goes through multiple layers of a CNN network, which consists of Conv and ReLU layers. Dilated Conv and ReLU layers help preprocess the image, i.e., removing noise and enhancing its quality. After the last Conv and ReLU layer, the Conv layer extracts important features. After the DSR process, the weights obtained are used in the training phase, and a dictionary is created. The output images from the DSR are added with additional Additive White Gaussian Noise (AWGN) prior to the testing phase. In the testing phase, the dictionary is utilized for image enhancement and, finally, the Quality Enhanced Image (QEI) is achieved.

B. Preprocessing

Among the most important methods for processing massive medical imaging datasets, CNNs are particularly well-suited for training on medical imaging datasets. The CNN architecture used in this work provides high-resolution quality for medical images. However, there are only a few techniques that work efficiently and faster in coordination with CNN architectures to process and fully reconstruct images. To increase the quality of high-resolution images, noise, blurriness, and other artifacts need to be eliminated. Hence, this work introduces the DSR technique, which incorporates CNN and ReLU to remove noise for DR images, helping eliminate existing challenges and

providing a faster implementation of medical imaging. Integrating DSR allows for rapid computing, accurate implementation, simplified training, and linking of spatial and temporal features. The use of a dictionary-learning approach makes the implementation of the DSR approach even faster and more precise due to its effective patch extraction method, which helps reconstruct high-resolution images by eliminating redundancy, computational complexity, and different kinds of noise. This makes the DSR approach more robust and efficient.

CNN and ReLU are utilized in DSR. The CNN architecture helps to denoise and reconstruct low-resolution images in better quality [14]. CNNs are typically described through the convolution operation, which is the core component of their architecture. The convolution operation is expressed as:

$$(I * K)(x, y) = \sum_{i=0}^{m-1} \sum_{j=o}^{n-1} I(x + i, y + j)K(i, j) \quad (1)$$

where I represents the input image, K represents a Conv kernel of size $m \times n$, (x, y) are the coordinates of the output pixel, and i, j represent indexes that repeat over kernel dimensions. The Conv-operation $(I * K)(x, y)$ at position (x, y) is computed by taking the element-wise product of the kernel K with the corresponding sub-region of the input image I, and then summing up all these products. ReLU is a widely used activation function in neural networks, especially in CNNs, defined as in (2), where $f(x)$ represents the ReLU function output and x represents the input. For positive inputs, the ReLU function returns the value of the argument, while for negative ones, it returns zero (3). The overall process in a CNN can be described through a series of such Conv operations, followed by non-linear activation functions, pooling layers, and Fully-Connected Layers (FCL). The Conv layer's operation, including ReLU, is described (4), where a_{ij}^k is activation at position (i, j) in k^{th} input image, σ represents ReLU, b^k represents a bias term for the k^{th} input image, and m and n iterate over the spatial dimension of the input image and the kernel.

$$f(x) = \max(0, x) \quad (2)$$

$$ReLu = \begin{cases} x > 0, \ then \ f(x) = x \\ x \geq 0, \ then \ f(x) = 0 \end{cases} \quad (3)$$

$$a_{ij}^k = \sigma\left(\sum_{i=0}^{m-1} \sum_{j=o}^{n-1} I_{mn} \cdot K_{mn}^k + b^k\right) \quad (4)$$

Then, the process of Dilated Conv+ReLU comes into place. The Dilated Conv technique is particularly useful in tasks like semantic segmentation, where capturing multi-scale context information is crucial. When combined with the ReLU activation function, it forms a powerful tool to capture features on multiple scales while maintaining computational efficiency. The Dilated Conv operation for I and K with a dilation rate d is represented in (5). After the Dilated Conv, the ReLU activation function introduces non-linearity. The ReLU function is represented in (6), where z is the output from the dilated convolution operation at each pixel location. The combined operation of Dilated Conv followed by ReLU is expressed in (7), where $a(x, y)$ is the activation at position (x, y) and $ReLU$ is the activation function applied element-wise.

$$(I *_d K)(x, y) = \sum_{i=0}^{m-1} \sum_{j=o}^{n-1} I(x + d_{i,y} + d_j)K(i, j) \quad (5)$$

$$ReLU(z) = \max(0, z) \quad (6)$$

$$a(x, y) = ReLU((I *_d K)(x, y) + b) \quad (7)$$

C. Training Phase

Training is an essential phase for image reconstruction and denoising. An efficient technique should easily denoise an image regardless of the kind of noise. In the proposed DSR, a Graphical Processing Unit (GPU) was used for parallel computing, expediting execution and saving time. Parameter regularization was performed through pre-training, and small-norm vectors became more adaptable to modification. Grayscale training was performed, and DSR was adapted such that switching on different noises was easy with a slight change in training. For training DSR, this work considered the differences among each layer feeding the respective dictionary, helping to achieve quality denoising.

A dictionary-learning approach was employed to achieve better denoising of the image patches. Assume that noise k is added to the input patched image i, which can be represented using (9). The patch size considered was $m \times m$, where k often was between 8 and 12. Each patch was separately denoised, and then the image was reconstructed by using all frames. An averaging approach was considered for controlling the overlapping of patches. Moreover, denoising was dependent on the dictionary $\mathbb{D}$, which had a size of $m^2 \times n$ where $n > m^2$. In the dictionary-learning approach, an atom-set was utilized as the basis-function. All noisy patches were denoised using the dictionary-based denoising technique with sparse linear reconstruction methods. Additive White Gaussian Noise (AWGN) was minimized using (9), where $\mathbb{Y}$ is considered a noisy-image patch, a pre-defined factor is utilized as ε, and l_0 pseudo-normalization is defined using $\|\cdot\|$

$$\mathbb{Y} = i + k \quad (8)$$

$$\min\|\alpha\|_0 \ \ s.t. \ \|\mathbb{D}_\alpha - \mathbb{Y}\|_2 \leq \varepsilon \quad (9)$$

In this work, sparsity was imposed using l_0. The l_1 norm was not suitable due to its convex normalization. The variable ε was selected in accordance with the approximation-error $\|\mathbb{D}_\alpha - \mathbb{Y}\|_2$. For performing quality denoising internally, a sliding-window approach was utilized. $\mathbb{D}$ was updated according to $\|\mathbb{D}_\alpha - \mathbb{Y}\|_2$ as shown in (10), where η belonged to the step-size. The overlapping patches were minimized using (11).

$$\mathbb{D} \leftarrow \eta \triangle \frac{\triangle\|i - \mathbb{D}_\alpha\|_2}{\triangle \mathbb{D}} \quad (10)$$

$$\min\|i - \mathbb{D}_\alpha\|_2 \quad s.t. \quad -1 \leq \alpha \leq 1 \quad (11)$$

In DSR, dictionary-learning helps to reconstruct the low-quality DR fundus image with high-quality. This approach utilizes the dictionary by training globally on the original and noise-induced datasets, helping to achieve better results. The features learned using the dictionary-learning approach include image patches, utilized as an incoherent inter-class dictionary and helping reduce the sparse-optimization error. Moreover, dictionary-learning and patch classification schemes are very fast, using less than 15% of the training time in the denoising

process. Thus, DSR creates a dictionary by training on the dataset, which is used in testing for evaluating the reconstructed image.

To evaluate the pixel means of various patch sizes of $m \times k$, where both m and k represent hyperparameters, the local mean estimation $\bar{i}$ for the image is given by (12). This work further estimates whether the input image is corrupted or not due to noise (AWGN), which is evaluated using (13), where r and n represent each position of the input image. The minimization function is defined using (14), where β represents the decision-factor, such that if β is near 1, it represents noisy images, and β near 0 represents a denoised image.

$$\bar{i} = \mathbb{F}\{\mathbb{Y} - k\} = \mathbb{F}\{\mathbb{Y}\} - \mathbb{F}\{k\} = \mathbb{F}\{\mathbb{Y}\} = \overline{\mathbb{Y}} \quad (12)$$

$$\log q(\mathbb{Y}|i) = \frac{1}{2\sigma^2}\sum_{r,n}\left(\mathbb{Y}_{r,n} - i_{r,n}\right)^2 \quad (13)$$

$$\min_{\beta} \mathcal{P} = \mathbb{F}\left\{\left(i - \bar{i} - \beta\left(\mathbb{Y} - \bar{i}\right)\right)\right\} \quad (14)$$

D. Testing Phase

This section describes the testing phase. Variability among patches is eliminated to prevent overfitting. The same dictionary utilized during the training phase is also utilized for testing in order to remove noise. This is done using (15), which requires the computation of $\bar{i}$ and $\mathbb{F}\left\{\left(\mathbb{Y} - \overline{\mathbb{Y}}\right)^2\right\}$, where σ denotes the noise level. Moreover, the computation of $\bar{i}$ is evaluated efficiently by noise-filtering by utilizing the dictionary for the $m \times k$ patch size. The final probability-distribution function for de-noising is evaluated using (16), where N_r^- defines the central pixel, $*$ defines the convolutional operation, η defines the step size, and λ represents an a priori function for controlling the probability, which is reliant upon the σ noise-level.

$$\hat{\imath} = \bar{i} + \frac{\left(\mathbb{F}\{(\mathbb{Y}-\mathbb{Y})^2\}-\sigma^2\right)(\mathbb{Y}-\bar{i})}{\mathbb{F}\left\{\left(\mathbb{Y}-\overline{\mathbb{Y}}\right)^2\right\}} \quad (15)$$

$$i^{t+1} = i^t + \eta\left[\sum_{r=1}^{K} N_r^- \psi_r(N_r i^t) + \frac{\lambda}{\sigma^2}(\mathbb{Y} - i^t)\right] \quad (16)$$

E. Quality Evaluation

The proposed DSR approach was evaluated using PSNR, SSIM, and MSE. PSNR is defined as the ratio between the highest level of signal and the degree of fluctuating noise that affects the image. PSNR is determined with MSE, which is calculated using two images. MSE quantifies the total mean squared error between the initial and reconstructed image, with a lower MSE indicating less error rate. PSNR is evaluated using (17), and MSE is evaluated using (18). SSIM is used to estimate the similarity between the input and the reconstructed image. SSIM is evaluated using (19), where μ_1 represents input image similarity, μ_2 represents reconstructed image similarity, σ_1 and σ_2 are the variances of the input and reconstructed images, σ_{12} denotes the covariance between the input and the reconstructed images, and U_1 and U_2 are variables that help stabilize weak denominators.

$$PSNR = \frac{10\log_{10}(L-1)^2}{MSE} \quad (17)$$

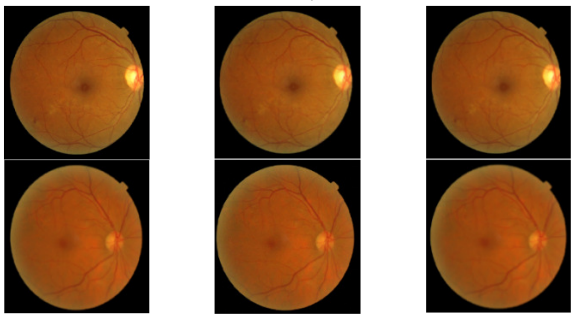
$$MSE = \frac{\sum_{M,N}[I_1(m,n)-I_2(m,n)]^2}{M*N} \quad (18)$$

$$SSIM(x_1, x_2) = \frac{(2\mu_1\mu_2+U_1)(2\sigma_{12}+U_2)}{(\mu_1+\mu_2+U_1)(\sigma_1+\sigma_2+U_2)} \quad (19)$$

III. RESULTS AND DISCUSSION

In this work, two publicly available benchmark datasets, EyeQ [15] and MESSIDOR [16, 17], were utilized to evaluate the effectiveness of the proposed DSR approach. The EyeQ dataset consists of color fundus images with varying quality levels, including good, usable, and reject-quality images. These images exhibit diverse real-world degradations such as uneven illumination, blur, noise, and low contrast, making the dataset well-suited for evaluating image enhancement and denoising algorithms. EyeQ images are widely used in retinal image quality assessment and enhancement studies due to their realistic imaging conditions. The MESSIDOR dataset contains high-resolution color fundus images acquired from multiple clinical centers using different fundus cameras. It includes images with varying degrees of DR severity and image quality variations caused by illumination differences and acquisition settings. The dataset is commonly used for benchmarking DR analysis techniques and provides a reliable platform to evaluate the generalizability of image enhancement methods.

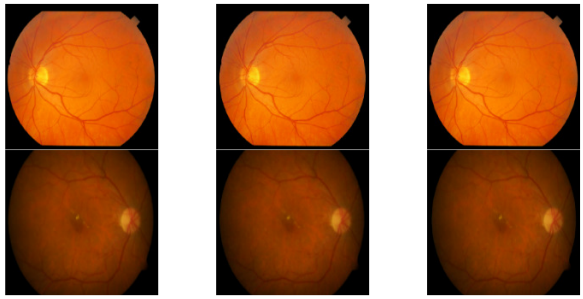
For this study, images from both datasets were resized and preprocessed before being divided into overlapping patches for training and testing. Artificial AWGN was introduced to simulate real-world noise conditions. The proposed DSR approach was then applied to remove noise and enhance image quality, and its performance was quantitatively evaluated using PSNR, SSIM, and MSE metrics. Figure 3 presents two results for the EyeQ dataset achieved by the proposed DRS approach. Figure 3(a) presents the input image, Figure 3 (b) presents the noise-induced image, and Figure 3(c) presents the output image.



(a) (b) (c)

Fig. 3. Results for the EyeQ dataset (a) Input image, (b) Images with induced noise, (c) enhanced image.

Similarly, Figure 4 presents the results achieved by the proposed DRS approach for two images from the MESSIDOR dataset. Figure 4(a) presents the input image of the MESSIDOR dataset, Figure 4(b) presents the noise-induced image, and Figure 4(c) presents the enhanced image.



(a) (b) (c)

Fig. 4. Results for the MESSIDOR dataset (a) Input image, (b) Images with induced noise, (c) enhanced image.

The results achieved by the DSR approach were compared with existing approaches in terms of PSNR, SSIM, and MSE for the EyeQ dataset, as presented in Table I. Among the listed models, DSR stood out with exceptional performance, achieving a PSNR of 37.82, an SSIM of 0.94, and an MSE of 0.0316. Notably, CLEAQ-DR and DASQE also demonstrated strong performance, with PSNR values of 28.53 and 28.47 and SSIM values of 0.917 and 0.914, showing significant improvements over earlier models such as cGAN and CuGAN. Models such as SCR-NET and OTT-GAN exhibited relatively lower performance, with PSNRs of 18.39 and 18.77 and SSIMs of 0.773 and 0.734, suggesting less effective image reconstruction capabilities.

TABLE I. COMPARISON OF DSR WITH CURRENT STATE-OF-THE-ART WORKS ON THE EYEQ DATASET

Models	PSNR (dB)	SSIM	MSE
cGAN [9], 2017	26.35	0.894	-
CuGAN [9], 2020	22.76	0.872	-
I-SECRET [9], 2021	27.36	0.908	-
IQE [9], 2023	22.36	0.813	-
CLEAQ [9], 2023	25.43	0.872	-
CLEAQ-DR [9], 2023	28.53	0.917	-
SCR-NET [12], 2022	18.39	0.773	-
OTT-GAN [12], 2023	18.77	0.734	-
ESDiff [12], 2023	19.27	0.864	-
DASQE [12], 2024	28.47	0.914	-
DSR (Proposed)	**37.82**	**0.94**	**0.0316**

In addition, the results achieved by the DSR approach were compared with current state-of-the-art works in terms of PSNR, SSIM, and MSE for the MESSIDOR dataset, as presented in Table II. Among the models, DSR was the most outstanding, achieving a PSNR of 36.85, an SSIM of 0.93, and an MSE of 0.0431, indicating superior image reconstruction quality, maintaining high fidelity and SSIM to the original images while minimizing error. In contrast, the IQE, CLEAQ, and CLEAQ-DR models showed progressively improved performance. IQE had a PSNR of 17.85 and an SSIM of 0.758, indicating a relatively lower image quality. CLEAQ improved with a PSNR of 19.36 and an SSIM of 0.805, and CLEAQ-DR further improved performance with a PSNR of 21.97 and an SSIM of 0.852. This comparison highlights the exceptional capability of DSR in producing high-quality images, far exceeding the other models.

TABLE II. COMPARISON OF DSR WITH CURRENT STATE-OF-THE-ART WORKS USING THE MESSIDOR DATASET

Models	PSNR	SSIM	MSE
IQE [9], 2023	17.85	0.758	-
CLEAQ [9], 2023	19.36	0.805	-
CLEAQ-DR [9], 2023	21.97	0.852	-
DSR (Proposed)	**36.85**	**0.93**	**0.0431**

IV. CONCLUSION

This work focused on improving the quality of DR fundus images by proposing an approach to image enhancement. The proposed DSR approach combines CNN, ReLU, and Dilated Conv methods for preprocessing. To further improve image quality, DSR utilized a dictionary approach during the training phase. Before testing, AWGN was induced in the input images to simulate real-world conditions. In the testing stage, noisy images were processed to remove noise and enhance image quality. Images from the MESSIDOR and EyeQ datasets were evaluated using SSIM, PSNR, and MSE. The proposed DSR method achieved 37.82 PSNR, 0.94 SSIM, 0.0316 MSE and 36.85 PSNR, 0.93 SSIM, and 0.0431 MSE on the EyeQ and MESSIDOR datasets, respectively. These results demonstrate improved performance compared to existing approaches, particularly in terms of PSNR and SSIM, indicating the effectiveness of the DSR approach in enhancing the quality of DR fundus images. Future work will extend to segmentation and classification using the DSR approach and the U-Net architecture.

REFERENCES

[1] K. Ogurtsova *et al.*, "IDF Diabetes Atlas: Global estimates for the prevalence of diabetes for 2015 and 2040," *Diabetes Research and Clinical Practice*, vol. 128, pp. 40–50, June 2017, https://doi.org/10.1016/j.diabres.2017.03.024.

[2] V. T. H. Tuyet, N. T. Binh, and D. T. Tin, "Improving the Curvelet Saliency and Deep Convolutional Neural Networks for Diabetic Retinopathy Classification in Fundus Images," *Engineering, Technology & Applied Science Research*, vol. 12, no. 1, pp. 8204–8209, Feb. 2022, https://doi.org/10.48084/etasr.4679.

[3] Z. Shen, H. Fu, J. Shen, and L. Shao, "Modeling and Enhancing Low-Quality Retinal Fundus Images," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 996–1006, Mar. 2021, https://doi.org/10.1109/TMI.2020.3043495.

[4] C. Wan *et al.*, "Retinal Image Enhancement Using Cycle-Constraint Adversarial Network," *Frontiers in Medicine*, vol. 8, Jan. 2022, Art. no. 793726, https://doi.org/10.3389/fmed.2021.793726.

[5] M. König *et al.*, "Quality assessment of colour fundus and fluorescein angiography images using deep learning," *British Journal of Ophthalmology*, vol. 108, no. 1, pp. 98–104, Jan. 2024, https://doi.org/10.1136/bjo-2022-321963.

[6] D. Shenkut and V. Bhagavatula, "Fundus GAN - GAN-based Fundus Image Synthesis for Training Retinal Image Classifiers," in *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, July 2022, pp. 2185–2189, https://doi.org/10.1109/EMBC48229.2022.9871771.

[7] S. H. Abbood, H. N. A. Hamed, M. S. M. Rahim, A. Rehman, T. Saba, and S. A. Bahaj, "Hybrid Retinal Image Enhancement Algorithm for Diabetic Retinopathy Diagnostic Using Deep Learning Model," *IEEE Access*, vol. 10, pp. 73079–73086, 2022, https://doi.org/10.1109/ACCESS.2022.3189374.

[8] T. Palaniswamy and M. Vellingiri, "Internet of Things and Deep Learning Enabled Diabetic Retinopathy Diagnosis Using Retinal Fundus

Images," *IEEE Access*, vol. 11, pp. 27590–27601, 2023, https://doi.org/10.1109/ACCESS.2023.3257988.

[9] Q. Hou, P. Cao, L. Jia, L. Chen, J. Yang, and O. R. Zaiane, "Image Quality Assessment Guided Collaborative Learning of Image Enhancement and Classification for Diabetic Retinopathy Grading," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1455–1466, Mar. 2023, https://doi.org/10.1109/JBHI.2022.3231276.

[10] K. G. Lee, S. J. Song, S. Lee, H. G. Yu, D. I. Kim, and K. M. Lee, "A deep learning-based framework for retinal fundus image enhancement," *PLOS ONE*, vol. 18, no. 3, Mar. 2023, Art. no. e0282416, https://doi.org/10.1371/journal.pone.0282416.

[11] W. Zhu, P. Qiu, M. Farazi, K. Nandakumar, O. M. Dumitrascu, and Y. Wang, "Optimal Transport Guided Unsupervised Learning for Enhancing Low-Quality Retinal Images," in *2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI)*, Apr. 2023, pp. 1–5, https://doi.org/10.1109/ISBI53787.2023.10230719.

[12] F. Liu and W. Huang, "ESDiff: a joint model for low-quality retinal image enhancement and vessel segmentation using a diffusion model," *Biomedical Optics Express*, vol. 14, no. 12, Dec. 2023, Art. no. 6563, https://doi.org/10.1364/BOE.506205.

[13] A. Q. Saeed, S. N. H. Sheikh Abdullah, J. Che-Hamzah, A. T. Abdul Ghani, and W. A. K. Abu-ain, "Synthesizing Retinal Images using End-To-End VAEs-GAN Pipeline-Based Sharpening and Varying Layer," *Multimedia Tools and Applications*, vol. 83, no. 1, pp. 1283–1307, Jan. 2024, https://doi.org/10.1007/s11042-023-17058-2.

[14] A. E. Ilesanmi and T. O. Ilesanmi, "Methods for image denoising using convolutional neural network: a review," *Complex & Intelligent Systems*, vol. 7, no. 5, pp. 2179–2198, Oct. 2021, https://doi.org/10.1007/s40747-021-00428-4.

[15] H. Fu *et al.*, "Evaluation of Retinal Image Quality Assessment Networks in Different Color-Spaces," in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*, vol. 11764, D. Shen, T. Liu, T. M. Peters, L. H. Staib, C. Essert, S. Zhou, P. T. Yap, and A. Khan, Eds. Springer International Publishing, 2019, pp. 48–56.

[16] "Messidor," *ADCIS*. https://www.adcis.net/en/third-party/messidor/.

[17] E. Decencière *et al.*, "Feedback on a Publicly Distributed Image Database: The MESSIDOR Database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231-234, Aug. 2014, https://doi.org/10.5566/ias.1155.