CHAPTER 4

A TWO-STAGE IMAGE PREPROCESSING FRAMEWORK FOR DENOISING AND EXPANDING RETINAL FUNDUS DATASETS

4.1 Introduction

The retinal fundus images consist of Gaussian, salt and pepper noises that disturb the image features which are essential for retinal diseases diagnosis. The noise removal process is challenging for researchers. The objective is to diminish the Gaussian noise effect on retinal images. The wavelet transform method is implemented in image processing, especially in medical images which have multi resolutions.

Preprocessing the retinal image prior to further analysis is a vital process that aids in removing the unwanted components from the original image. The noise removal, image enhancement and correction procedures provide positive effects on the image analysis results. In this study, a two-stage image preprocessing framework is developed that includes image denoising and augmentation. Different denoising techniques, such as discrete wavelet transform and filter-based techniques, such as Median and Wiener filters, are applied to retinal fundus images. Another preprocessing task is to improve the image contrast by implementing histogram-based image enhancement. These techniques not only enhance the quality of the image but also serve as a crucial foundation for subsequent classification tasks. Additionally, data augmentation strategies were implemented to expand the dataset by synthesizing retinal fundus images. This expansion enhanced the number of samples, enabling the model to provide unbiased results by balancing out the image samples present in the dataset classes.

This chapter is structured as follows. The image preprocessing techniques such as median, Wiener, DWT and DWT_K-SVD techniques for image denoising are defined in the subsequent section. For contrast enhancement, HE, AHE and CLAHE are presented. Experimental results of these techniques are discussed, the performance analysis is carried out, and the outcome assessments of the proposed techniques are presented conclusively. The augmentation methods are applied to the pre-processed retinal images, and the training dataset after static augmentation for APTOS and IDRiD datasets is finally presented.

4.2 Retinal Fundus Image Preprocessing Techniques

The preprocessing techniques enhance the image quality, which aids in analyzing the image effectively. The preprocessing technique neglects unwanted misrepresentations and enhances the precise qualities that are vital for the proposed model to work efficiently. The fundus image consists of several types of noises such as dots, triangles, circles, squares and so on. Incorrect focus leads to blurring and over-exposure, which can obstruct correctly identifying the stages of DR disease. One of the vital challenges in the preprocessing stage is to preserve and retain the essential image information post-processing to make effective image classification. Different kinds of preprocessing like image denoising and contrast enhancement methods, are proposed in this study.

Image denoising is the primary process before any image processing steps like as pattern recognition, remote sensing, object identification, image classification and feature selection. The image noise leads to various intensity variations that generate grain like occurrence in the image. The noise in the digital image arises due to reasons like capturing instrument fault, heat sensor, scattering, lens misalignment and surrounding noise etc. The distorted digital image noise is categorized as Gaussian, speckle, Poisson, salt and pepper noises. Mallat S (2008) came up with a wavelet transform noise elimination method. The input signal transforms into diverse scales consisting of unique components. The threshold value is fixed, and the operations are carried out to minimize the noise (Chang S. G. *et al.* (2000); Donoho D. L. and Johnstone I. M. *et al*. (1995); Donoho D. L. (1995)). The denoised image output is reconstructed back to the wavelet coefficients.

The filtering techniques can be broadly categorized as spatial domain filtering and frequency domain filtering. The linear and non-linear filtering methods, a wavelet-based denoising technique, and contrasts are discussed in detail. The subsequent sub-sections discuss in detail the median, Wiener, DWT, Hybrid DWT_K-SVD, HE, AHE and CLAHE techniques. In this study, these filtering processes are performed prior to giving them as inputs to the proposed deep learning models.

4.2.1 Median Filter

To remove the noise a median filter is implemented, which is simple and a non-linear filtering method. Median filter is used for minimizing the strength differences between pixels. Median value is substituted with the pixel value. The pixels are sorted in ascending order and the pixels are replaced by estimating the mid pixel value. The median filter enhances the

image quality, enhances vital details, and preserves the edges before performing deep image operations. The median filter replaces each entry with an adjacent value, a simple process is required if the value is odd, and a complex process is required if the value is even (James C *et al.*, 2008). Figure 4.1 shows the denoised images using Median filter on APTOS and IDRiD datasets.

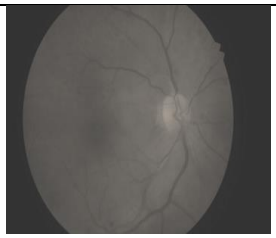
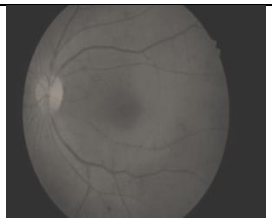
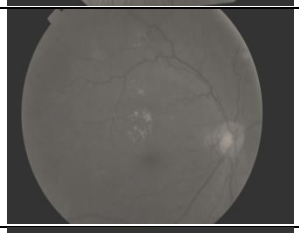
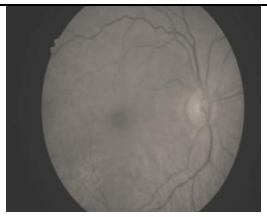
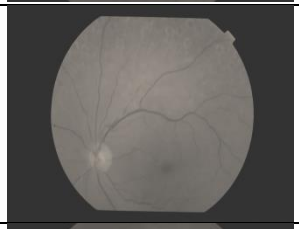
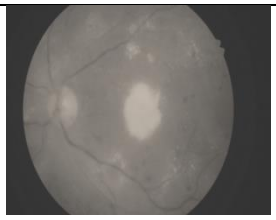
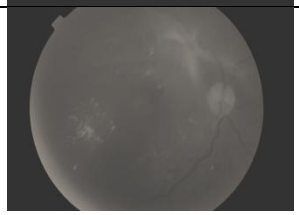
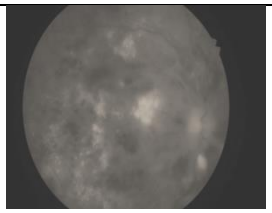
Class	Median Filter on APTOS dataset image	Median Filter on IDRiD dataset image
0		
1		
2		
3		
4		

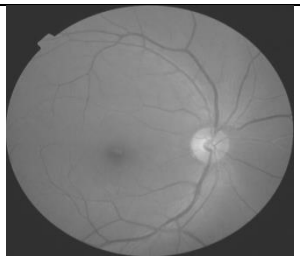
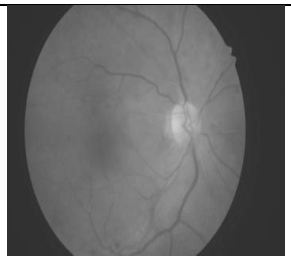
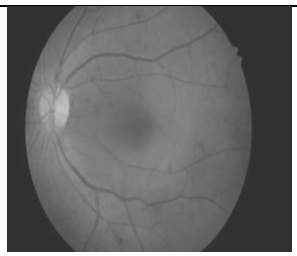
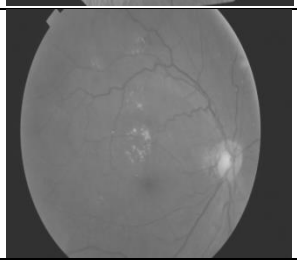
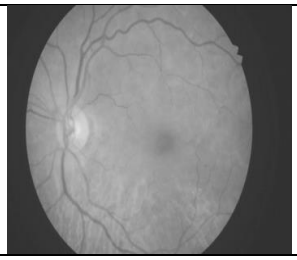
Figure 4.1: Denoised fundus images using the Median filter from the APTOS and IDRiD Datasets

The median filter is effective in reducing the noise by preserving image edges and important features on the images. In medical images, the edge and image features are more important for the classification of the different DR stages. Based on the result and analysis,

the unwanted image components are cleared, and the nerves around the images are clearly seen, which helps in further analysis. However, in certain cases, the median filter tends to leave some white patches around the retinal region, making it challenging for further analysis. Retaining important features such as blood vessels, lesions, and microaneurysms is crucial in diabetic retinopathy diagnosis, so these white patches around the vital region make it less efficient when the affected images undergo classification tasks.

4.2.2 Wiener Filter

To remove or reduce the noises in the signal a Wiener filter is applied to the image which is a linear filtering method. Norbert Wiener proposed the Wiener filter (WF) in 1940. A low-pass filter blurs the image and it's recovered by the inverse filtering process. This filter is sensitized to additive noises. WF is a combination of inverse filter and noise leveling process. WF eliminates additive noises and reverses image blurring. The processing speed and the simplicity of WF makes it effective in handling Gaussian noises (Anilet B *et al.*, 2014). Figure 4.2 shows the denoised images using Wiener filter on APTOS and IDRiD datasets.

Class	Wiener Filter on APTOS dataset image	Wiener Filter on IDRiD dataset image
0		
1		
2		

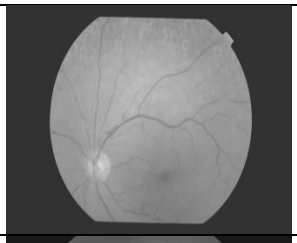
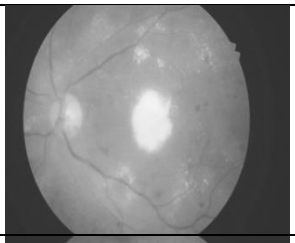
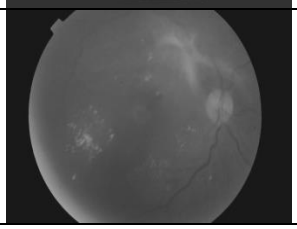
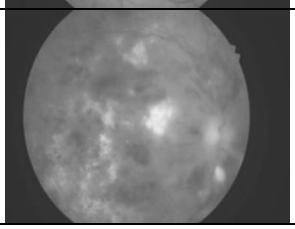
Class	Wiener Filter on APTOS dataset image	Wiener Filter on IDRiD dataset image
3		
4		

Figure 4.2: Denoised images using the Wiener filter from the APTOS and IDRiD Datasets

Like the median filter, the Wiener filter is effective in reducing image noise while preserving features in medical images. This filter helps remove blurriness caused by motion artifacts during image capture. Based on the results and analysis, applying the Wiener filter improves image sharpness, particularly on edges, and increases the image histogram, making nerves more prominent to the naked eye. Like the median filter, this WF also affects the image quality by creating white patches around the retinal region.

4.2.3 Discrete Wavelet Transform

Mallat S. G. (1989) proposed a multi-resolution analysis method using discrete wavelet transformation. The non-redundant image representation formed by the DWT image output is accurate in spectral and spatial image localization. Image denoising using DWT uses multi-scale methods which consists of multi-scale approximation, sparse representation, image quality preservation etc. In DWT method, the image signal decomposition is performed, and the signals are provided with two filters that produces two outputs with different coefficients approximately in the decomposition process. The image quality is retained by reassembling the image components in the reconstruction process. The decomposition is referred as DWT and reconstruction is referred as inverse DWT. Mohideen S. *et al.* (2008) identified wavelet technique is effective for Gaussian noise. Figure 4.3 presents the framework of DWT based image denoising method.

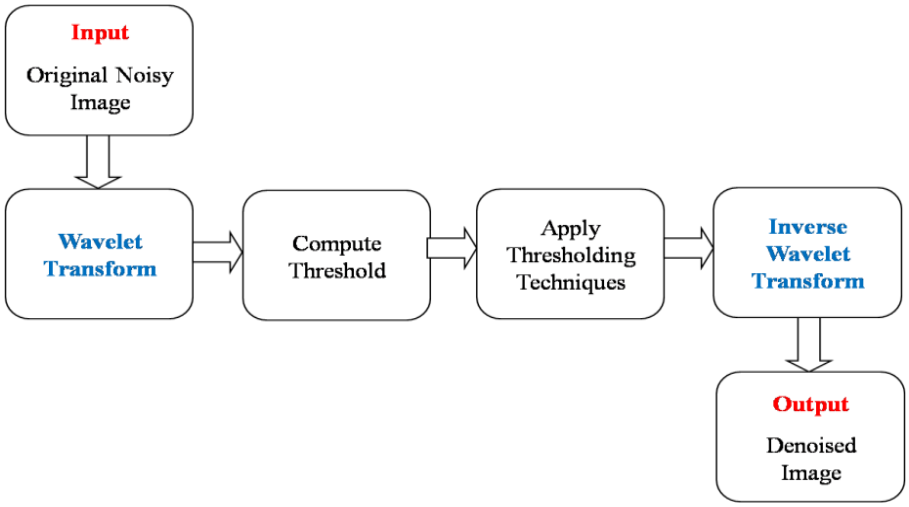
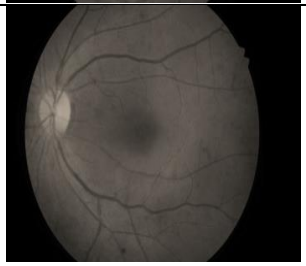


Figure 4.3: Framework of DWT-based image denoising

The typical steps followed in the DWT-based image denoising method are given below:

- DWT converts noisy input images to an orthogonal domain.
- The signals are separated to create wavelet coefficients.
- Wavelet threshold value is set to de-noise wavelet coefficients.
- Inverse DWT is made to obtain the denoised image.

Figure 4.4 shows the denoised images (class 0-4) using the DWT on the APTOS and IDRiD datasets.

Class	Discrete Wavelet Transform on APTOS dataset image	Discrete Wavelet Transform on IDRiD dataset image
0		
1		

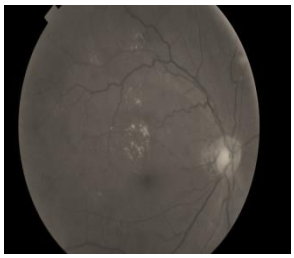
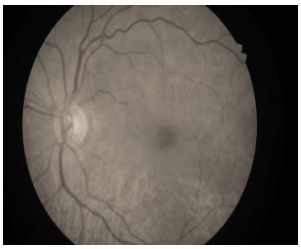
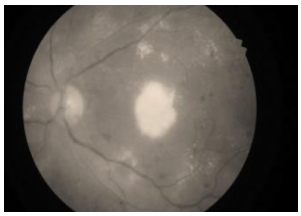
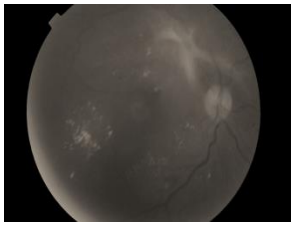
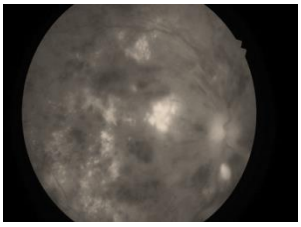
Class	Discrete Wavelet Transform on APTOS dataset image	Discrete Wavelet Transform on IDRiD dataset image
2		
3		
4		

Figure 4.4: Denoised images using the DWT on the APTOS and IDRiD datasets

The DWT technique leverages multi-resolution representation to eliminate image noise. Based on the result and analysis, it is observed that the formation of white patches during the denoising process is minimized, attributable to the segmentation of an image into distinct frequency bands. This segmentation enhances the sharpness of image edges, consequently refining image quality for subsequent analysis. Notably, regions surrounding the retina and optic disc become more discernible for further examination. While DWT surpasses both median and Wiener filters in performance, some images still exhibit residual patches. Nonetheless, the overall prevalence of affected images is diminished compared to other filters.

4.2.4 Hybrid DWT_K-SVD

As Gaussian noise and Salt & Pepper noise are the predominant noises present in Retinal fundus images, it is required to eliminate such noises to achieve better DR diagnosis. However, eliminating image noise is a difficult task to all researchers. The objective of this research work is to minimize Gaussian noise, especially on retinal images. Wavelet transform

has gained significant popularity in various domains of image processing, with a notable emphasis on its application in medical image denoising. This widespread adoption is primarily attributed to its valuable attributes, including multi-resolution capabilities and the ability to exploit sparsity. The proposed DWT_k-SVD method is a hybrid technique that uses the wavelet coefficients to de-noise the distorted image from the sparse representation, which is learned from the redundant dictionary. The k-SVD (Aharon M *et al.* 2006) denoising technique initiates k-means clustering and decomposition. A redundant dictionary that creates a sparse representation is a dictionary-based learning method. The signal and dictionary are used earlier for learning strategies. Let 'x' and 'y' denote noisy and noiseless image with additive noise and 'z' denotes the Gaussian noise with σ. 'x' input image is represented as

$$x = y + z \tag{4.1}$$

In the DWT_K-SVD method, image noise 'z' is removed from 'x' to create a denoised image 'y1' similar to unique image 'y'. DWT_K-SVD based image denoising method are given below:

- DWT method is implemented on the original image 'x' to produce wavelet co-efficient and approximation. Adaptive Gaussian threshold is applied on the wavelet coefficients.
- K-SVD is a learning algorithm to de-noise wavelet co-efficient. The GOMP algorithm is applied to find the sparse codes for the threshold coefficients. The algorithm learns from the wavelet illustration of the removed patches.
- Inverse DWT is implemented to produce denoised output image 'y1'.

'y' in equation (4.1), consists of √M ×√M pixels and DWT is implemented on 'x'. Equation (4.2) is produced when wavelet property is applied on the equation (4.1).

$$W_x = W_y + W_z \tag{4.2}$$

The x, y, and z wavelet transforms are represented as, W_x, W_y *and* W_z respectively. The dictionary learns from the small image patches that are extracted $W_{yjj} = K_{jj} W_y$ of size √m ×√m is (i,j) of W_y. Using matrix $K_{ij,}$ the W_y image's block size is extracted. The dictionary (redundant) 'D' and sparse illustrations ($\hat{\alpha}$) of all mined patch W_{yjj} with restricted error is represented as,

$$\hat{\alpha}_{ij} = \arg\min_{\alpha_{ij}} \| \alpha_{ij} \| 0 \; such \; that \; \| Wy_{ij} - D\alpha_{ij} \|_2^2 \leq (C\sigma)^2 \quad (4.3)$$

The Lagrange form of the equation (4.3) is represented in equation (4.4) as follows,

$$\hat{\alpha}_{ij} = \arg\min_{\alpha_{ij}} \| Wy_{ij} - D\alpha_{ij} \|_2^2 + \mu_{ij} \| \alpha_{ij} \| 0 \quad (4.4)$$

Figure 4.5 presents the framework of DWT_K-SVD-based image denoising, and the pseudo-code is also given.

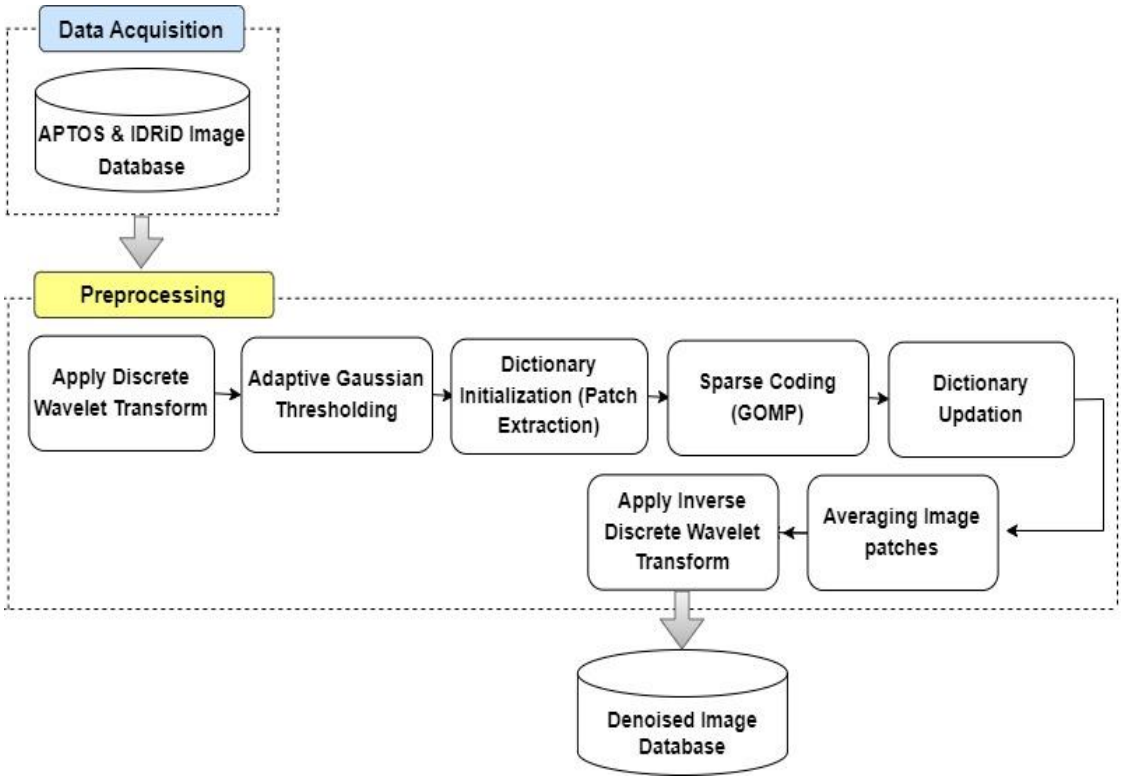


Figure 4.5: The Proposed DWT_K-SVD Framework for Image Denoising

Pseudo Code: A Hybrid DWT_K-SVD technique for image denoising

Input: *Image data from APTOS & IDRiD dataset*

Output: *Denoised images*

Begin

Step 1: Data acquired from the APTOS & IDRiD image dataset.

Step 2: Apply Discrete Wavelet Transform (DWT) that uses Daubechies 2 (db2) wavelet to decompose the input image into its wavelet coefficients.

Step 3: Use adaptive Gaussian thresholding to wavelet coefficients.

Step 3.1: Calculate the standard deviation of coefficients.

Step 3.2: Calculate the threshold based on the standard deviation.

Step 3.3: Apply the threshold to coefficients.

Step 4: Sparse Coding using Greedy Orthogonal Matching Pursuit (GOMP)

Step 4.1: The dictionary is initialized, and the size can be adjusted.

Step 4.2: Sparsity level is defined, indicating the number of non-zero coefficients that represent each thresholded coefficient.

Step 4.3: The GOMP algorithm is applied to find the sparse codes for the thresholded coefficients. This iteratively selects the most promising atoms from the dictionary to represent the data.

Step 5: Apply Inverse Discrete Wavelet Transform (IDWT) on denoised wavelet coefficients to reconstruct the denoised image. The same wavelet type (db2) is applied.

End

In this approach, adaptive Gaussian thresholding is used to address the issue of variable noise levels within an image. Images often contain regions with varying levels of noise, and this adaptive thresholding approach helps preserve image details while effectively reducing noise. The Gaussian thresholding method adapts the threshold to the local characteristics of the coefficients, making it a competent denoising technique.

Also, the GOMP algorithm is applied to find the sparse codes for the thresholded coefficients. GOMP is a greedy algorithm that iteratively selects the most promising atoms from the dictionary to represent the data. This process efficiently captures the significant features in the data while reducing the number of coefficients used for representation. It efficiently finds a sparse representation of the data, which is crucial for denoising. By using a

sparse representation, it reduces the impact of noise in the coefficients while retaining essential image information. The error is minimized as a result of the operation, and the image patches are averaged. Finally, the inverse DWT is executed on the wavelet coefficients to obtain the denoised image.

The computational complexity of the DWT_K-SVD method is analyzed based on the execution time, the type and nature of the images, the complexity of the algorithm, hardware used such as the type of processor, RAM size, and free space available in the system. In this comparison, average execution time is computed using MATLAB 2020b tool on an Intel Core i7 processor having 16 GB RAM. The average execution time (in seconds) of the proposed DWT_K-SVD method is 1.64 for APTOS dataset and 1.89 for IDRiD dataset.

4.2.4.1 Performance Evaluation of DWT_K-SVD method

The proposed DWT_K-SVD method is analyzed and compared with Median, Wiener and DWT. Figure 4.6 (a) and (b) show the original noisy image and the denoised image obtained by the DWT_K-SVD method for the APTOS dataset. Figure 4.7 (a) and (b) show the original noisy and the denoised image obtained by the DWT_K-SVD method for the IDRiD.

(a) **(b)**

Figure 4.6: (a) Original noisy image (b) Denoised image by DWT_K-SVD (Proposed method) on the APTOS Dataset

Upon implementing the proposed methodology, a noticeable enhancement is observed in the border regions of the images, where sharpness is notably improved. This sharpening effect

extends to reveal intricate details, such as nervous lines, with remarkable clarity. Importantly, this enhancement does not compromise the overall integrity of the image; rather, it accentuates the visual quality without introducing any visible damage. The methodology appears to effectively balance between sharpening and preserving the natural appearance of the image, resulting in a refined and aesthetically pleasing output. Overall, it demonstrates promising potential for enhancing image quality in various applications.

(a) **(b)**

Figure 4.7: (a) Original noisy image (b) Denoised image by DWT_K-SVD (Proposed method) on the IDRiD Dataset

The PSNR, MSE, and SSIM are the performance metrics used in this study to assess the performance of the proposed DWT_K-SVD. The mathematical formulation for calculating the MSE and PSNR values is represented in equations (4.5) and (4.6).

$$MSE = \frac{1}{M_1 x M_2} \sum_{i=1}^{M_2} \sum_{j=1}^{M_1} [I(i,j) - I'(i,j)]^2 \tag{4.5}$$

$$PSNR = 10\ x \log(\frac{255^2}{MSE}) \tag{4.6}$$

where I and I’ denotes two images M1×M2 size. The pixel value of the i^{th} row and j^{th} column of I and I’ is represented as I(i,j) and I’(i,j). The estimation of SSIM values are shown in equation (4.7).

$$SSIM(I, I') = \frac{(2\mu_I \mu_{I'} + c_1)(2\sigma_{II'} + c_2)}{(\mu_I^2 + \mu_{I'}^2 + c_1)(\sigma_I^2 + \sigma_{I'}^2 + c_2)} \tag{4.7}$$

where $\mu_I = \sum_{i=1}^{N} w_i x_i$ and $\sigma_I = \left(\sum_{i=1}^{N} w_i (x_i - \mu_I) \right)^{1/2}$ denotes average grey value and I variance. $\sigma_{II'} = \sum_{i=1}^{N} w_i (x_i - \mu_I)(y_i - \mu_{I'})$ denotes the covariance between I and I'. $C_1 = (K_1 L)^2$ and $C_2 = (K_2 L)^2$ denotes the two constants.

The MSE, PSNR and SSIM are the three-performance metrics used in this study to evaluate DWT_K-SVD. Figure 4.8 represents the MSE performance metrics analysis of APTOS and IDRiD datasets obtained by median, Wiener, DWT and DWT_K-SVD techniques. The proposed DWT_K-SVD technique outperforms the other techniques on APTOS and IDRiD. Unlike PSNR and SSIM, the MSE values should be lesser. The MSE value of 0.17 is obtained on both APTOS and IDRiD datasets.

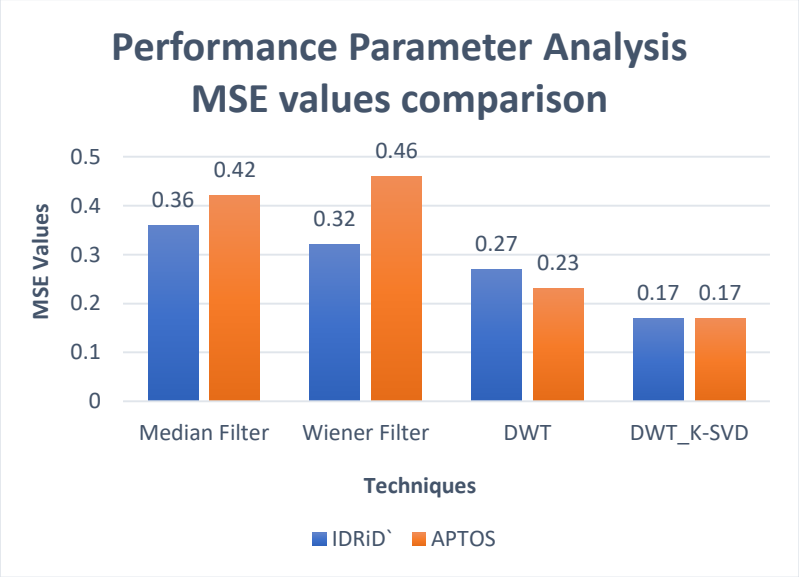


Figure 4.8: MSE Performance metric analysis for APTOS and IDRiD datasets

Figure 4.9 represents the PSNR performance metric analysis of APTOS and IDRiD datasets obtained by the median, Wiener, DWT and DWT_K-SVD techniques. The PSNR values obtained by the DWT_K-SVD technique on APTOS and IDRiD datasets are 44 dB and 45 dB, respectively. The DWT_K-SVD technique beats the other techniques on APTOS and IDRiD by obtaining the maximum PSNR value compared to the other techniques.

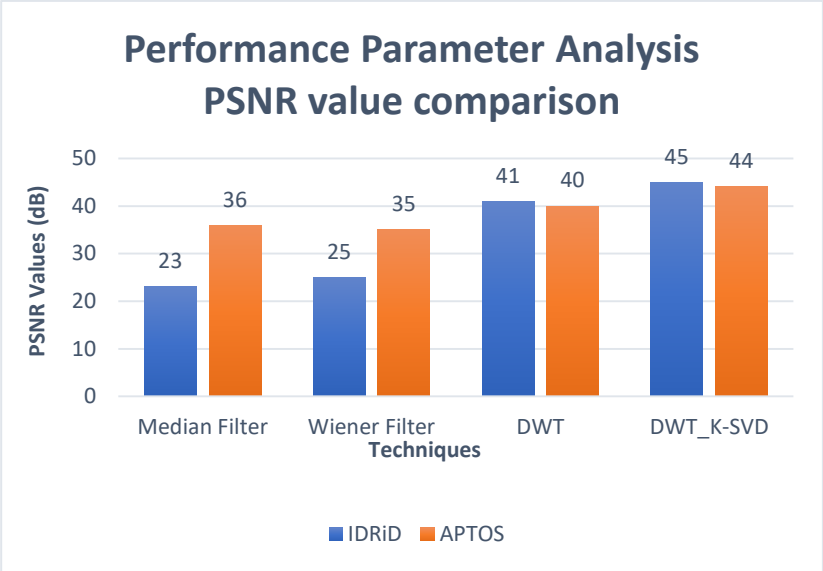


Figure 4.9: PSNR Performance metric analysis for APTOS and IDRiD datasets

Figure 4.10 represents the SSIM performance metric analysis of APTOS and IDRiD datasets obtained by the median, Wiener, DWT and DWT_K-SVD techniques. The SSIM values obtained by the DWT_K-SVD technique on APTOS and IDRiD are 0.92 and 0.90, respectively. The DWT_K-SVD technique outperforms existing techniques on APTOS and IDRiD datasets by obtaining maximum SSIM value compared to the other techniques.

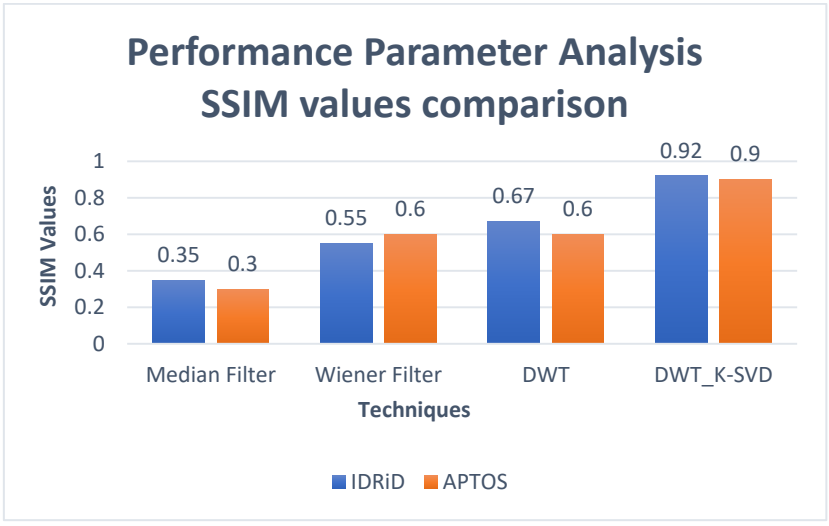


Figure 4.10: SSIM Performance metric analysis for APTOS and IDRiD datasets

The DWT method is relatively operative when compared to the median and wiener methods. The use of adaptive Gaussian thresholding in the hybrid method helps preserve image details while effectively reducing the image noise. To minimize the error and enhance the performance of the denoising process, DWT is combined with the K-SVD method. To de-

noise retinal images, a dictionary learning approach is effective. The DWT_K-SVD produces optimum results compared to the median, wiener and DWT on APTOS and IDRiD datasets. Hence, the results of DWT_K-SVD method are efficient in denoising the retinal fundus images.

4.3 Histogram-based Contrast Enhancement Techniques

After denoising, the histogram-based image enhancement techniques are applied to improve the image contrast. This helps in clearly visualizing the tiny blood vessels. Techniques such as HE, AHE and CLAHE are applied to DWT_K-SVD Denoised image. HE is a simple technique to enhance the image contrast with nominal computational work. A uniform allocation of the pixel intensities to grayscale levels. This method increases the global contrast of the images. He is not efficient for retinal fundus images due to its high-intensity noises (Zheng Y *et al.*, 2012; Odstrcilik J *et al.* (2014). In AHE, histograms are estimated for separate image segments to redistribute the luminance image values Jintasuttisak T and Intajag S (2014). AHE enhances the local contrast and edge. AHE performs better when compared to the HE method Shaik F *et al.* (2010). CLAHE (Jintasuttisak T and Intajag S 2014; Pisano E *et al.* 1998) is applied to enrich the low contrast in the images. The noise amplification is controlled by the clip limit parameter using CLAHE. The clipping at a prearranged area preceding to Cumulative Distribution Function (CDF) estimation controls the noise amplification.

The enhanced images are separated into lesser non-overlapping regions known as blocks. The Block Size (BS) and the Clip Limit (CL) are the vital parameters used in CLAHE to control and generate the image with enhanced quality. The image quality is enhanced by maximizing the BS and CL values, which enhances the image contrast. Figure 4.11 shows the CLAHE pixel redistribution where the HE is applied to each tile, and it's clipped.

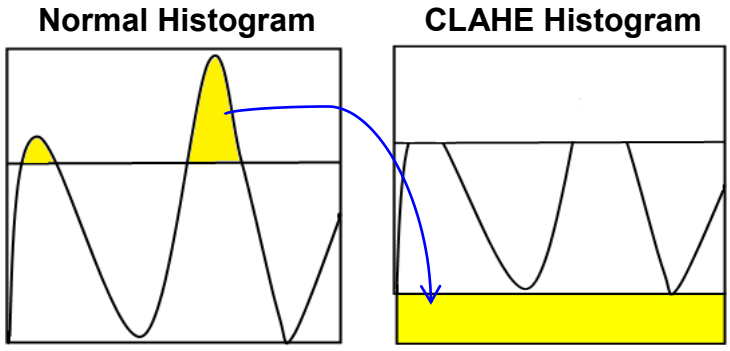


Figure 4.11: CLAHE pixel redistribution

The pixels are reallocated to specific gray levels because the image strengths are inadequate. The reallocated and the real histogram are dissimilar. The enhanced and real image retains the gray levels (Jintasuttisak T and Intajag S 2014; Xu Z *et al.* 2009). The CLAHE method performs the following steps for image enhancement:

- The input DWT_K-SVD pre-processed image is split into minor non-overlapping regions known as tiles. Here, MxN= Total number of tiles.
- The histogram of each tile is estimated as per the gray level values.
- The contrast restricted histogram of each region is estimated by the clip limit, which is represented as $Np_{avg} = \frac{(NpX\ x\ NpY)}{N_g}$ where Np_{avg} denotes the average number of pixels, N_g denotes the number of grey levels, NpX and NpY denotes the number of pixels in X and Y. Here, CL is represented as $N_{cl} = N_{ncl}\ x\ Np_{avg}$ where N_{cl} denotes the clip limit, N_{ncl} denotes the normalized [0,1] clip limit. When the pixel number is more significant, N_{cl} then the pixels are trimmed, and the clipping rule is applied.
- The left-over pixels are reallocated, and the Rayleigh transform is applied to improve the intensity value. This works better for retinal images at high alpha values Dos Santos *et al.* (2020).
- For rescaling, a linear contrast stretch is applied. The tile areas are joint, and the border artifacts are eliminated.

In this study, HE, AHE and CLAHE techniques are applied in the preprocessing stage after denoising. Figure 4.12 and Figure 4.13 show the original noisy image and contrast-enhanced image obtained by the CLAHE method on APTOS and IDRiD datasets.

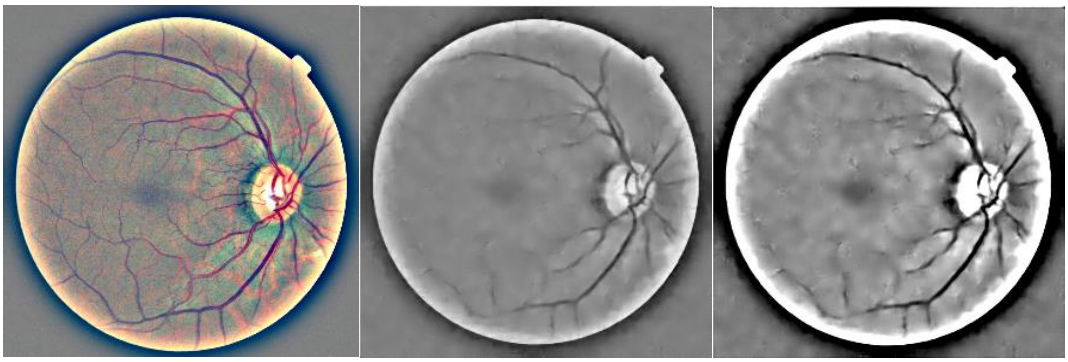


Figure 4.12: Original noisy image, Denoised image-DWT_K-SVD and contrast-enhanced image-CLAHE on the APTOS Dataset

Initially, the original noisy image from the APTOS dataset is represented and the denoised image produced using the DWT_K-SVD is represented. Then, the contrast-enhanced image using CLAHE is represented in Figure 4.12. After denoising, contrast-enhanced techniques are applied to the denoised images to clearly visualize the tiny blood vessels that appear around the vital region.

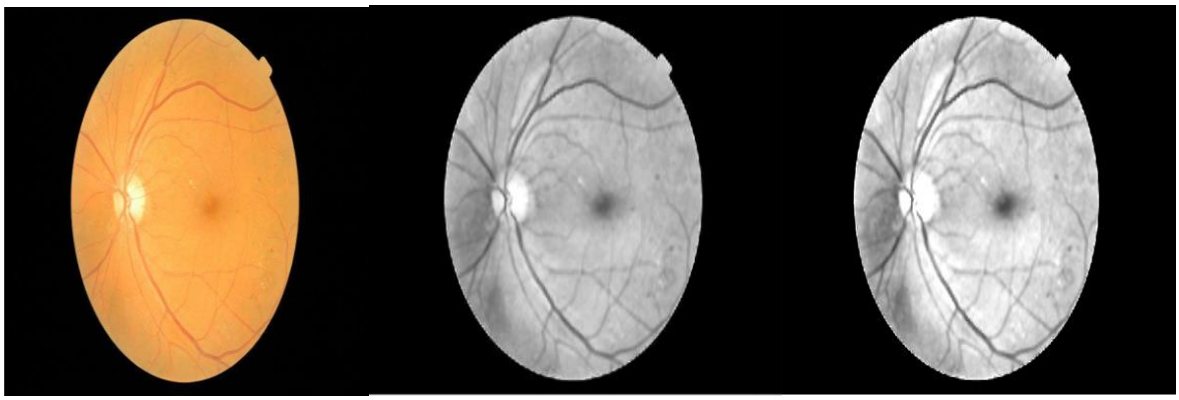


Figure 4.13: Original noisy image, Denoised image-DWT_K-SVD and contrast-enhanced image-CLAHE on the IDRiD Dataset

Similarly, the original noisy image obtained from the IDRiD dataset is shown in Figure 4.13. Subsequently, the DWT_K-SVD method is applied to denoise the images, resulting in the representation seen in Figure 4.13. Following denoising, contrast enhancement techniques are implemented to improve visibility, particularly highlighting the tiny vessels, as shown in Figure 4.13.

The PSNR, SNR and SSIM are the three-performance metrics used in this study to assess the proposed CLAHE method. Figure 4.14 represents the PSNR performance metrics analysis of APTOS and IDRiD datasets obtained by HE, AHE and CLAHE methods. The proposed CLAHE overtakes the other techniques on APTOS and IDRiD. The PSNR values of 45dB and 44dB achieved in the CLAHE method are higher than the other methods on both APTOS and IDRiD datasets.

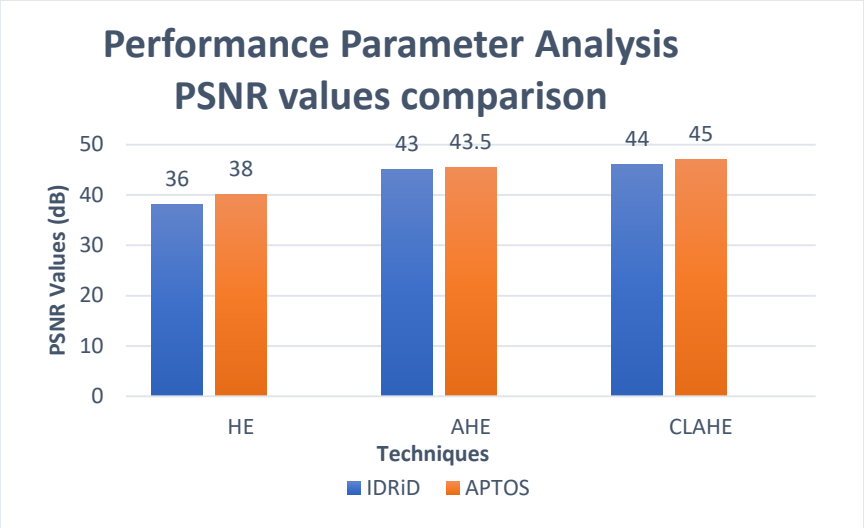


Figure 4.14: Contrast Enhancement - PSNR metric comparison on the APTOS and IDRiD datasets

Figure 4.15 represents the SNR performance metrics analysis of APTOS and IDRiD datasets obtained by HE, AHE and CLAHE methods. The SNR values obtained by the CLAHE method on APTOS and IDRiD are 71 dB and 72 dB, respectively. The proposed CLAHE overtakes the existing techniques on APTOS and IDRiD by obtaining maximum SNR value compared to the other techniques.

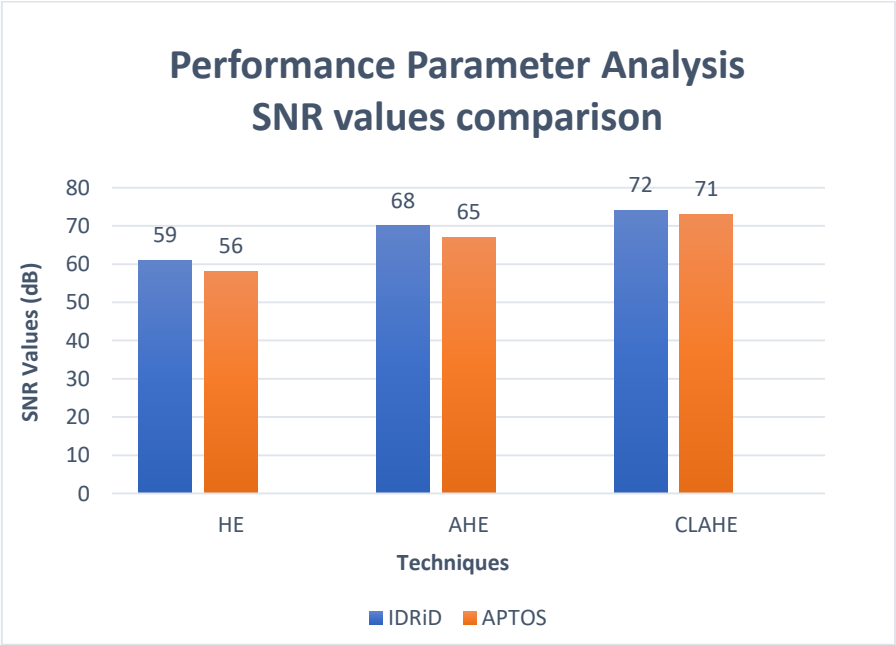


Figure 4.15: Contrast Enhancement - SNR metric comparison on the APTOS and IDRiD datasets

Figure 4.16 represents the SSIM performance metrics analysis of APTOS and IDRiD obtained by HE, AHE and CLAHE methods. The SSIM values obtained by the CLAHE on APTOS and IDRiD are 0.90 and 0.91, respectively. The CLAHE outperforms the other techniques on APTOS and IDRiD by obtaining maximum SSIM value compared to the other techniques.

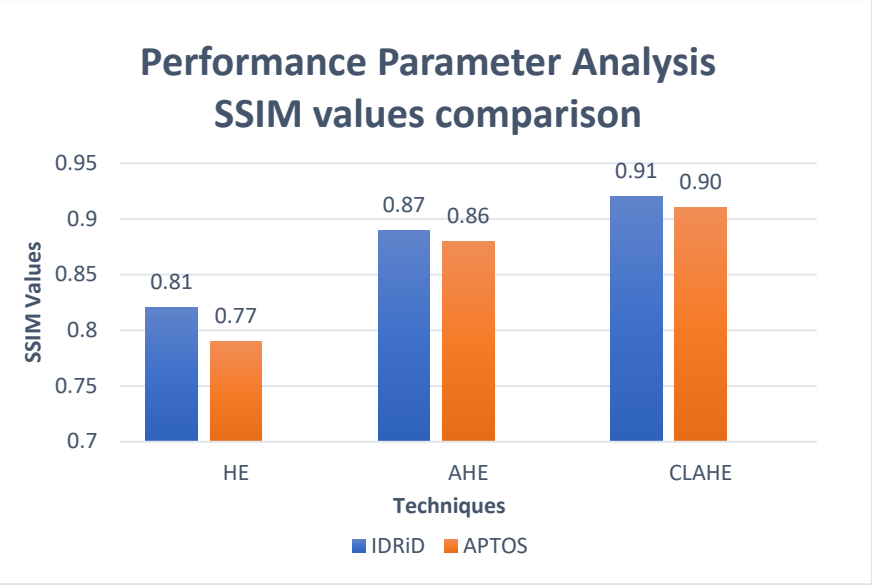


Figure 4.16 Contrast Enhancement - SSIM metric comparison on the APTOS and IDRiD datasets

4.4 Static Augmentation on the Preprocessed Dataset

Static image augmentation is performed as the second step in this two-stage image preprocessing framework. This section discusses the augmentation methods applied to the pre-processed retinal image and the training dataset after static augmentation on APTOS and IDRiD datasets.

The augmentation process is applied to the pre-processed data (training set). Synthetic images are created to handle class imbalance problems. Additional training images are generated by applying various transformations to existing images using data augmentation, which in turn balances the dataset classes. This helps in preventing overfitting and enhances the model's ability to generalize to new data. The augmented images are then fed into DL models for classification tasks. The various pipelines of operations in augmentation are given below:

i) **Random rotation:** Rotating the image clockwise randomly by selecting an angle in the range of degrees.

ii) **Horizontal flip:** Flipping the image from left to right randomly with the specified probability.

iii) **Vertical flip:** Flipping the image from top to bottom randomly with the specified probability.

iv) **Translation:** The image shifts in horizontal and vertical directions with the specified maximum complete fraction.

The model is trained using the augmented dataset. Figure 4.17 shows the augmentation flow. Figure 4.18 shows the pipeline operations involved in augmentation. Figure 4.19 shows the augmented dataset images.

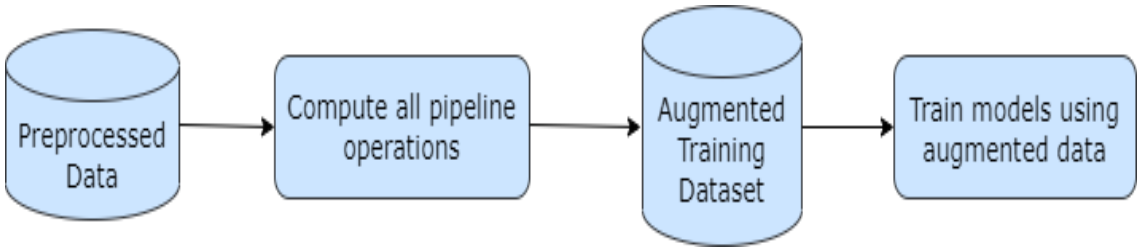


Figure 4.17: Augmentation flow

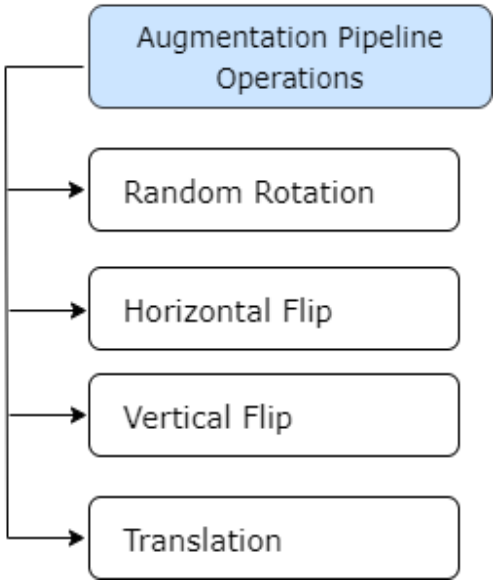
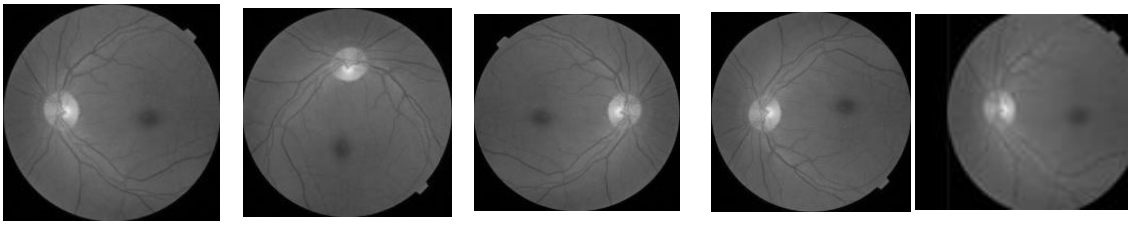


Figure 4.18: Augmentation pipeline operations performed on preprocessed data



a. Pre-processed Image **b. Rotation (Degree=90)** **c. Horizontal Flip** **d. Vertical Flip** **e. Translation**

Figure 4.19: Augmented dataset images

Table 4.1 shows the training dataset after static augmentation on APTOS and IDRiD datasets. The actual number of training images and the number of augmented images in APTOS and IDRiD datasets based on severity levels are given. The static data augmentation approach has been used in this context to balance the dataset images. The classes with the least number of samples are augmented with respect to the class that has the maximum number of image samples. The augmented images are appended to the training dataset. The dataset classes are then balanced to provide better generalization and unbiased results.

Table 4.1: APTOS and IDRiD training datasets after static augmentation

Class / Severity level	APTOS 2019 Dataset		IDRiD Dataset	
	Imbalanced dataset classes before augmentation	Balanced dataset classes after augmentation	Imbalanced dataset classes before augmentation	Balanced dataset classes after augmentation
0 – Normal	1435	1435	134	134
1 - Mild NPDR	322	1435	18	134
2 – Moderate NPDR	791	1435	131	134
3 - Severe NPDR	146	1435	53	134
4 - Proliferative DR	234	1435	74	134
Total no. of images	**2930**	**7175**	**410**	**670**

4.5 Summary

In this chapter, several image preprocessing techniques are applied to the retinal fundus image to eliminate the image noise. APTOS and IDRiD are used to evaluate the model for training. The categories of images collected and annotated samples, DR grades and

other information enclosed in the DR database. The image preprocessing techniques such as median, Wiener, DWT, DWT_K-SVD for image denoising are described in detail. The results of the previously mentioned filtering techniques are showcased and analyzed. A study on filtering techniques is performed, and the test results show that the values produced by the DWT_K-SVD beat the other filtering methods. Performance metrics like MSE, PSNR and SSIM are used in this study to evaluate the DWT_K-SVD method. Image enhancement techniques such as HE, AHE and CLAHE are applied to enhance the image contrast. The CLAHE method reduces the image noise compared to the HE and AHE methods. Performance metrics such as PSNR, SNR, and SSIM are used to assess the contrast enhancement techniques on both datasets. BS and CL are the two parameters that influence the CLAHE results. Hence, based on the comparative study, the DWT_K-SVD method de-noises the retinal image noise efficiently, and the CLAHE overtakes the other histogram-based methods to enhance the image contrast. The augmentation methods such as random rotation, horizontal flip, vertical flip, and translation are applied to the pre-processed retinal image and the training dataset after static augmentation for APTOS and IDRiD datasets are presented.