CHAPTER 3

RESEARCH METHODOLOGY

This chapter briefs the outline of the three models proposed in this study. The base network for these three proposed models is the Residual Network (ResNet-32). This chapter discusses the pre-trained ResNet-34 CNN structure, experimental setup, dataset description, dataset distribution based on the DR stages and the research framework.

3.1 Basics of Residual Network (ResNet) Architecture

Computer vision applications use a deep learning concept, which has become very popular in recent times, especially when working on huge volumes of data. CNN architecture consists of thousands of convolutional layers, which are simply scalable to a huge number of layers. The scalability leads to an issue known as a vanishing gradient. Residual blocks are introduced in the ResNet architecture to overcome the issue. Figure 3.1 shows the architecture of the ResNet-34 Layer used in this research, and Figure 3.2 shows the skip connection process. The skip connection technique connects the initiation of a layer to supplement layers by hopping a few layers in between and recycles the initiations of the former layer. The hopping process speeds up the training initially by reducing the network with few layers.

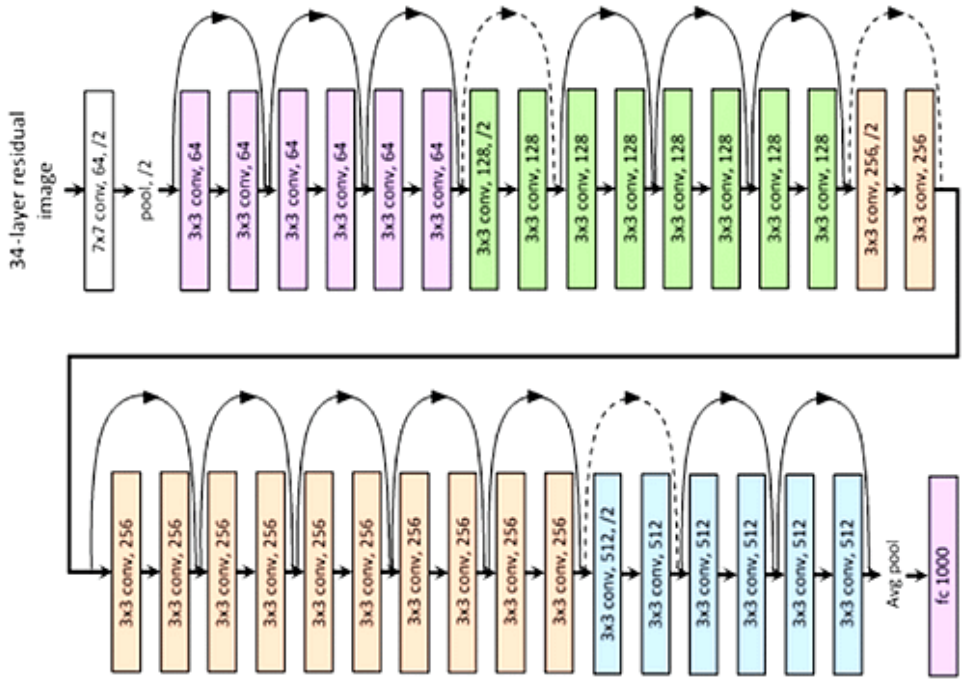

Figure 3.1: Architecture of ResNet-34 Layered

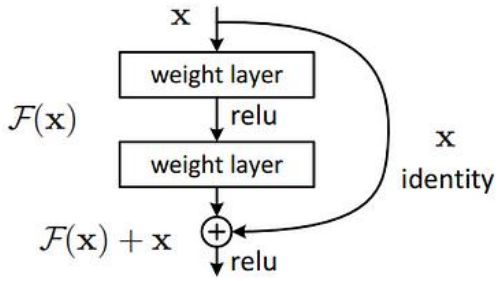

Figure 3.2: Skip Connection in ResNet-34

The network is reoriented by extending all the layers and the residual parts of the network, which is known as the residual part. This part can explore additional feature space in the input image. During nonlinearity, ResNet models hop two or three layers and batch normalization is performed in between. An advanced ResNet architecture trains 'skip weights' to control the number of layers.

3.2 Experimental Setup and Dataset Description

The system configuration details, and the dataset description are given in this section. The details of fundus image acquisition for validation and analysis are discussed in the subsequent sections. The APTOS and IDRiD image database from KAGGLE were used to evaluate the proposed model that is appropriate for the training.

The experiments are performed using Intel Core i7 7th generation processor, 2.7 GHz, 16 GB RAM, 1TB SSD, and Windows 10 operating system. The proposed model is evaluated using Matlab 2020b tool. The GPU used is NVIDIA GeForce RTX 2080. The following section defines the datasets which includes categories of images, accumulated samples quantity, annotated, DR grades and facts that were enclosed in the DR database.

3.2.1 APTOS 2019 Blindness Detection dataset

In this research, RF images from Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 dataset were used. The APTOS database contains a total of 3662 fundus image collection acquired from assorted rural Indian citizens. For classification purposes, APTOS dataset samples are preferred. The fundus images are classified into five different stages based on severity scale where 0 indicates No DR, 1 indicates mild DR, 2 indicates moderate DR, 3 indicates severe DR, and 4 indicates PDR. The main RF image sample is classified as a healthy image that is comprised of no DR. The subsequent images are classified based on the DR conditions with different levels of distribution as in Figure 3.3. Every consequent classification consists of more imperfect retinas. Figure 3.3 illustrates RF samples in APTOS.

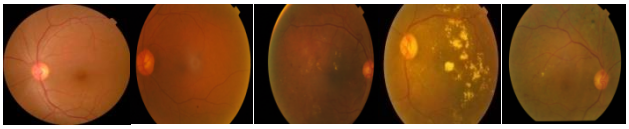

Figure 3.3: APTOS 2019 blindness detection dataset samples from each class (0-4)

3.2.2 Indian Diabetic Retinopathy Image Dataset (IDRiD)

The next set of RF images used in this study is from the IEEE ISBI -2018 IDRiD (Prasanna P et al. 2018). A variety of 516 images with DR and Diabetic Macular Edema (DME) consisting of 413 training and 103 testing images were used for disease grouping. Figure 3.4 shows RF samples from every class in IDRiD. The DR grade is classified and labeled into five classes based on the severity levels on a scale of 0-4, like APTOS.

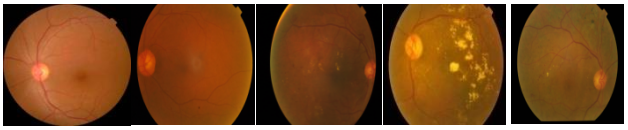

Figure 3.4: Indian Diabetic Retinopathy Image Dataset (IDRiD) samples from each class (0-4)

3.3 DR Stages/Severity Levels

The dataset description and the dataset distribution based on the severity levels are discussed subsequently. Table 3.1 illustrates the DR stages/severity levels.

Table 3.1: Ophthalmoscopic Features based on DR Stages/Severity Levels

S.No.	DR Stage / Severity level	DR Ophthalmoscopic Features
1	Mild NPDR	Minimum one microaneurysm observed
2	Moderate NPDR	1. Microaneurysm & hemorrhage in 2 quadrants 2. Mild intraretinal microvascular abnormalities found 3. Cotton wool spots & exudates are found
3	Severe NPDR	Any 1 following symptoms, 1. Venous beading in 2 quadrants 2. Intraretinal abnormalities in 1 quadrant 3. Microaneurysms & extensive hemorrhages in 4 quadrants
4	PDR	Early Neovascularization found at the optic disc or elsewhere

The DR ophthalmoscopy features are described below:

- Microaneurysm - leak fluids, lipids, proteins formation and red spots near macular area.
- Hemorrhage - Occurs due to capillaries leakages.
- Exudates - lipoproteins leakage.
- Cotton-wool spots - whitish lesions in the retinal area.
- Intraretinal Microvascular abnormalities - comprise irregular lines (red) connecting venules with the arterioles.
- Venous abnormalities - dilation, beading, and looping in the capillary area.
- Neovascularization - abnormal blood vessels development at the OD.

3.4 Dataset Image Distribution over Classes

The Aptos 2019 dataset comprises 3662 image samples distributed over five distinct classes, with the 0-Normal class having the maximum samples and the 3-Severe NPDR having the minimum samples. In the IDRiD dataset, 512 image samples are distributed over five distinct classes, with the 0-Normal class including the most samples and the 1-Mild NPDR class containing the fewest. Table 3.2 illustrates the total number of images in APTOS and IDRiD datasets. In both APTOS 2019 and IDRiD datasets, the entire data has been split into two parts: 80% training and 20% testing. The total image count in the APTOS training dataset is 2930, and in the testing, dataset is 732. The total image count in the IDRiD training dataset is 410, and in the testing, dataset is 103.

Table 3.2: Total Number of Images in APTOS and IDRiD Datasets

DR Class / Stage	APTOS 2019 Dataset	IDRiD Dataset
0 – Normal	1794	168
1 - Mild NPDR	403	22
2 - Moderate NPDR	989	164
3 - Severe NPDR	183	66
4 - Proliferative DR	293	93
Total no. of images	**3662**	**513**

3.5 Research Framework

The complete process of this research is depicted in Figure 3.5. Initially, the process begins with data acquisition. In this research work, two different datasets, the APTOS 2019 Blindness Detection and the IDRiD, are used. Then, Pre-processing is performed in two stages: 1. De-noising and Contrast enhancement 2. Augmentation. During image de-noising, various techniques, such as median, wiener, DWT, and DWD_K-SVD (proposed technique), are applied. The objective of performing pre-processing is to eliminate Gaussian and salt & pepper noises and highlight DR features for better DR diagnosis. The Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM) are the three-performance metrics used in this study to assess the DWT_K-SVD method.

During the contrast enhancement process, HE, AHE and CLAHE approaches are applied. The PSNR, SNR and SSIM are the three-performance metrics used in this study to assess HE, AHE and CLAHE methods. Then, augmentation is performed using a pipeline of operations such as random rotation, horizontal flip, vertical flip, and translation. These are the most used augmentation techniques in medical imaging. Three different works are proposed for feature extraction and classification: the MSA-ResNetGB model, SGAN-ECR model and SGAN-OECR model.

The first model, namely the MSA-ResNetGB model, enhances the performance of DR stage classification. The high-level features are extracted for classifying the different DR stages based on the severity levels. The ResNetGB encoder, multi-scale feature extraction analysis, MSA strategy, decoder and classification units are the various processes involved in this model. Accuracy, precision, recall and F1- score are the performance metrics used to evaluate the MSA-ResNetGB model.

The second model, namely SGAN-ECR, was proposed to overcome the limitations present in the MSA-ResNetGB model and to enhance the classification performance by generating and training synthetic images along with the original images. The ensemble classification regression model provides a better generalization ability for classifying DR stages based on severity. The various processes involved in the proposed SGAN-ECR model are SGAN-based image augmentation, the structure of SGAN and DR stage classification using ensemble classification regression model. The metrics of MSA-ResNetGB is used to evaluate the SGAN-ECR model.

The third model, namely SGAN-OECR, was proposed to optimize the hyper-parameters of the MSA-ResNet classification and regression model. The various processes involved in the proposed SGAN-OECR model are configuring the MSA-ResNet structure's hyper-parameters by using an enhanced mine blast algorithm to tune the hyper-parameters. The metrics of MSA-ResNetGB are used to evaluate the SGAN-OECR model.

In the end, the performance of all three models is assessed and compared using the selected metrics. Each of these models was trained separately on both the original images and pre-processed images, and the outcomes were then presented for analysis.

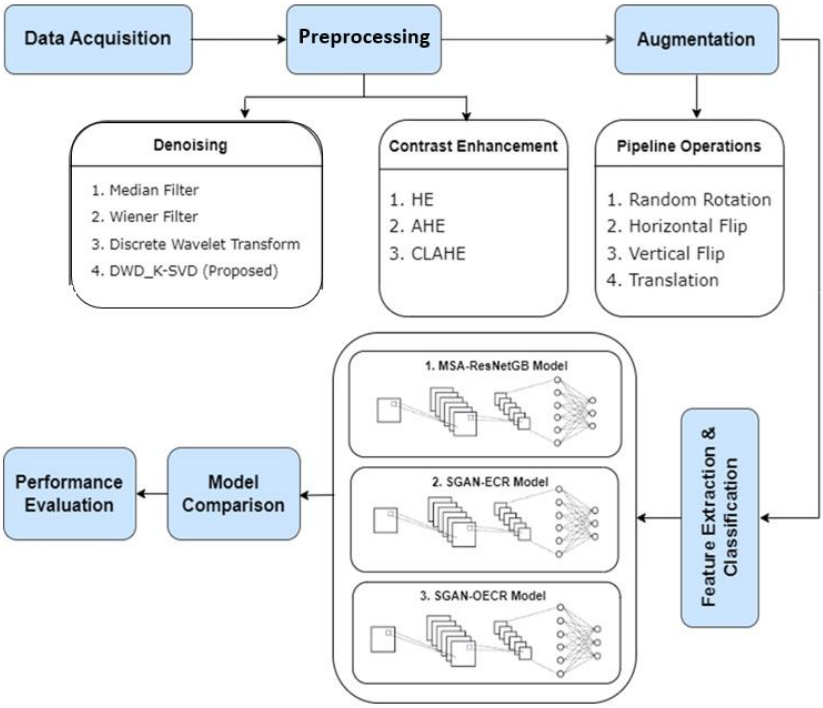

Figure 3.5: Research Framework

3.6 Summary

In this chapter, the base network, which is the ResNet-32, is discussed in detail. The experimental setup and the details of fundus image acquisition for validation and analysis are discussed. The APTOS and IDRiD image database from KAGGLE were used to evaluate the proposed model that is appropriate for the training. The details of the dataset are described, such as categories of images, number of image samples, DR grades and ground truths that were enclosed in the DR database. The dataset distribution based on the DR stages is discussed. The outline of the research procedure, which consists of data acquisition, pre-processing, augmentation, and feature extraction is explained in brief with the proposed deep learning models such as MSA-ResNetGB, SGAN-ECR and SGAN-OECR.