CHAPTER 8

CONCLUSION AND FUTURE ENHANCEMENT

8.1 Conclusion

This study focused on improving the Diabetic Retinopathy (DR) Stage-wise Classification performance, which helps ophthalmologists in the decision-making process. DR disease affects the vision of a diabetic patient, which can lead to blindness in its advanced stages. The seepage on blood vessels in the retina in diabetic patients is the cause of permanent blindness. If DR disease is identified at an initial stage and treated precisely, then eye blindness can be prevented. A digital photograph of a retinal image is used for screening patients with DR and Glaucoma diseases. The abnormal retinal images can be measured with the help of retinal vessel structure that specifies the state of the disease. Hence, an advanced system is required to classify the stage of eye disease. Computer Aided Diagnosis (CAD) aids the ophthalmologist in categorizing the normal and abnormal retinal images. OD, blood vessel thickness, hard exudates, vein diameter measurement, and hemorrhages are a few of the features extracted from the retinal images.

In this study, features like OD, blood vessel thickness, etc., are extracted from the retinal image automatically for normal and abnormal classification. Parameters like accuracy, precision, recall and F1-score are applied to assess the model's performance. The DR stages, such as mild non-proliferative DR, moderate non-proliferative DR, severe non-proliferative DR, and proliferative DR, are considered. Hence, there is a need for more accurate DR diagnosis that assists ophthalmologists in making decisions. The various issues in DR stage classification, such as the accuracy of the classification based on the number of training and testing images, the retinal images with imbalanced classes, CNN structures, the hyper-parameters of CNN structures and other additional vital features that aid in accuracy of classification are considered in this study.

Firstly, image preprocessing techniques were developed to denoise the retinal images and to improve the image contrast. These techniques effectively reduce Gaussian and salt & pepper noise, as it is the most found noise on retinal fundus images. It also served 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 diversity of samples, enabling the model to learn and generalize better. Furthermore, strategies were devised to address class imbalance issues within the Retinal image database. These strategies were instrumental in improving the accuracy of DR stage classification by ensuring that all classes received adequate representation. Preprocessing the retinal image prior to further analysis is a vital process to remove the unwanted components from the original image. In this study, a two-stage image preprocessing framework is developed that includes image denoising and augmentation. The noise removal, image enhancement and correction procedures provide positive effects on the image analysis results. Image preprocessing techniques like filter-based and wavelet-based denoising methods are applied to the dataset images. Another preprocessing step is contrast enhancement; here, histogram-based image enhancement techniques are used. Then, augmentation is performed on the pre-processed images to expand the dataset. The different preprocessing techniques are essential to the retinal image prior to giving them as input to the deep learning models.

Image denoising methods like median, Wiener, DWT, and DWT_K-SVD are used. For contrast enhancement, methods like HE, AHE and CLAHE are applied. The MSE, PSNR and SSIM are the three-performance metrics used in this study to assess the proposed DWT_K-SVD method.

Experimental results and performance analysis are performed. 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 CLAHE method in 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 method outperforms the other histogram-based methods to enhance the image contrast. The most commonly used augmentation techniques like random rotation, translation, horizontal and vertical flip were applied as pipeline of operations thus expanding and balancing the dataset images class-wise.

The first model of the study is the creation of an MSA-ResNetGB model to improve the performance of DR stage classification when compared to the other pretrained CNN models. The enhancement of Convolutional Neural Network (CNN) architectures played a pivotal role in recognizing subtle signs of DR, including soft EXs, hard EXs, MAs, and HEs.

The high-level features are extracted for classifying the different DR stages. The extracted features are classified as normal and abnormal images by providing input to the classifier to classify retinal image. The ResNetGB encoder module consists of several layers in the deep network. The encoder extracts feature through the consecutive residual blocks. A scaling mechanism is employed to create semantic learning for those features with various spatial resolutions. ResNetGB model uses deep networks known as vanishing gradients. The attention mechanism in the ResNetGB model consists of a sequence of convolution units used for multi-scale feature analysis. The decoder module is the final part of the network in the proposed model, which consists of fully connected layers to map the feature vectors to the outcome in the principal component analysis layer. Two different benchmark datasets, such as APTOS and IDRiD, are used to evaluate the proposed model. Accuracy, precision, recall and F1-score are the metrics used to measure the MSA-ResNetGB model.

Experimental results and performance analysis are performed for the MSA-ResNetGB model on APTOS and IDRiD datasets. A comparison is performed with the other CNN-based pretrained models. It has been found that the MSA-ResNetGB model achieved better results when trained on preprocessed data and their results are discussed in this section. The accuracy of the proposed MSA-ResNetGB model is 94.40% on APTOS dataset. The confusion matrix of the MSA-ResNetGB model on the APTOS dataset is presented. The performance metrics are analyzed for the MSA-ResNetGB model on the IDRiD. The accuracy of the proposed MSA-ResNetGB model is 94.18% on the IDRiD dataset. The confusion matrix of the MSA-ResNetGB on the IDRiD dataset is presented. The performance of APTOS dataset exceeds the performance of the IDRiD dataset. When both the datasets are compared, there is a performance loss in the IDRiD dataset due to a smaller number of image samples.

The second model of the study is creating a novel model, namely a SGAN-ECR, to overcome the limitations and enhance the classification performance of MSA-ResNetGB model. The key objective of the SGAN-ECR is to generate high-contrast and RF images for classifying DR stages based on severity levels. The SGAN model consists of a pixel-2-pixel GAN with a UNet++ generator and a patch-GAN discriminator. A pixel-2-pixel GAN generator is used to train the shape of dissimilarities of the retinal fundus images. For image segmentation with accurate local contrast patterns, a patch-GAN discriminator is used. In the

SGAN model, high-contrast synthetic RF images are generated, which are utilized for training the MSA-ResNet regression and classification structures along with the original images. The proposed SGAN-ECR model consists of SGAN-based image augmentation, the structure of SGAN and DR stage classification using ensemble classification regression model. Initially, benchmark retinal images alone were used for training the model, but in this SGAN model, along with the original images, high-contrast images generated from SGAN were also used for training. The images are trained by the MSA-ResNet structure, which consists of an MLP classifier. The trained model is applied to assess the samples for DR stage classification. The efficiency of classification is enhanced by estimating the cross-entropy loss factor for classification and the mean square error loss factor for the regression process.

Experimental results and performance analysis are performed for the SGAN-ECR model on APTOS and IDRiD datasets. A comparison is performed with the other CNN-based pretrained models. It has been found that the SGAN-ECR model achieved better results when trained on unprocessed data and their results are discussed in this section. The accuracy of the proposed SGAN-ECR model is 97.81% on APTOS dataset. The confusion matrix of the SGAN-ECR model on the APTOS dataset is presented. The accuracy of the proposed SGAN-ECR model is 96.12% on the IDRiD dataset. The confusion matrix of the SGAN-ECR on the IDRiD dataset is presented. The performance accuracy of the APTOS dataset exceeds the performance accuracy of the IDRiD dataset.

The third model of the study is creating an efficient model, namely, SGAN-OECR, to automatically assign the hyper-parameter values. The classification efficiency was optimized by thoroughly investigating factors such as Retinal Fundus (RF) image resolution and tuning CNN hyper-parameters, including learning rate, number of iterations, network depth, and batch size. An Enhanced Mine Blast Algorithm (EMBA) is proposed for optimizing the hyper-parameter values of the MSA-ResNetGB model. The convolution layer count, filter count, filter size, FC layer count and the hidden layers in FC layer are the hyper-parameters used to improve the architecture. The SGAN-ECR model with parameters optimized is named SGAN-OECR. An optimized SGAN model is also trained with the images used for the SGAN-ECR model. To identify the deformities of the test images of DR lesion a trained model is applied. The MLP classifier in final FC layer classifies DR lesions in MSA-ResNet. The classifier accuracy depends on the hyper-parameters used in MSA-ResNet. The proposed

SGAN-OECR model consists of hyper-parameters for the MSA-ResNet structure, using an enhanced mine blast algorithm to tune the hyper-parameters.

Experimental results and performance analysis are performed for the SGAN-OECR on APTOS and IDRiD datasets. A comparison is performed with the other CNN-based pretrained models. It has been found that the SGAN-OECR model achieved better results when trained on unprocessed data and their results are discussed in this section. The accuracy of the proposed SGAN-OECR model is 98.21% on the APTOS dataset. The confusion matrix of the SGAN-OECR model on the APTOS dataset is presented. The accuracy of the proposed SGAN-OECR model is 97.09% on the IDRiD dataset. The confusion matrix of the SGAN-OECR on the IDRiD dataset is presented. The performance of APTOS dataset exceeds the performance of the IDRiD dataset.

In the final analysis, the performances of all three models were assessed and compared using the considered metrics. Each model underwent training on both the original and preprocessed images as separate datasets, and the results were subsequently obtained. The findings revealed that the MSA-ResNetGB model exhibited strong performance when applied to preprocessed images. Conversely, the SGAN-ECR and SGAN-OECR models excelled when working with unprocessed images. This difference in performance is attributed to the fact that SGAN models were trained using high-contrast synthetic images generated by SGAN, as opposed to relying solely on the original dataset images.

Limitations:

- The Deep Learning models proposed in this research work perform well on the current system setting but require high computational resources, limiting their accessibility to those with access to high-performance computing resources.
- The model's performance heavily relies on the availability of a substantial amount of labeled data, which may be challenging to obtain in certain domains.
- Scaling the Deep Learning model to larger datasets may be challenging due to the high computational requirements and the need for a vast volume of labeled data.
- The process of labeling a large volume of data can be labor-intensive and may require significant human resources, hindering the acquisition of necessary labeled data for model training.

However, the primary objective of this research work has been successfully achieved by improving the performance of DR stage classification using an optimized convolutional neural network-based ensemble classification and regression framework with a focus on aiding Ophthalmologists in decision-making. The fulfillment of secondary objectives, including noise reduction, data augmentation, class imbalance mitigation, abnormal signs recognition, and classification efficiency optimization, collectively reinforced the model's capabilities and laid the foundation for advancing the diagnosis and treatment of Diabetic Retinopathy.

8.2 Recommendations for future study

The study can be protracted to comprise the following:

- The model can be generalized to classify other diabetic eye diseases such as Glaucoma, Macular Edema, Cataracts and so on.
- The proposed deep learning models can be implemented in cloud-reliant platforms.
- Hybrid optimization algorithms can be considered in tuning the CNN hyper-parameters.
- The different combinations of neural network architectures can be explored to benefit from the strengths of various models, like transformers, CNNs, and recurrent neural networks (RNNs).
- Advanced strategies can be proposed to create smaller and more efficient models for edge devices and low-resource environments.
- Domain-specific pre-trained models can be developed to accelerate the development of medical applications.