Automated Diabetic Retinopathy DetectionUsing Convolutional Neural Networks For Feature Extraction And Classification (ADRFEC)

P. AnilKumar

Assistant Professor, Department of CSE, Sir C.R.Reddy college of Engineering, India.

Abstract:
Diabetic Retinopathy (DR) is a significant complication of Diabetes Mellitus, leading to various retinal abnormalities that can impair vision and, in severe cases, result in blindness. Approximately 80% of patients with long-standing diabetes for 10–15 years develop DR. The manual process of diagnosing and detecting DR for timely treatment is both time-consuming and unreliable, mainly due to resource constraints and the need for expert opinion. To address this challenge, computerized diagnostic systems utilizing Deep Learning (DL) Convolutional Neural Network (CNN) architectures have been proposed to learn DR patterns from fundus images and assess disease severity. The proposed model performs an exhaustive analysis of these architectures upon fundus images, and derives the best performing DL architecture for DR feature extraction and fundus image classification. Amongst all the models, ResNet50 has achieved the highest training accu- racy whereas VGG-16 has achieved the lowest training accuracy. Again, VGG-16 has achieved lowest validation accuracy whereas ResNet101 has achieved highest validation accuracy.

Keywords: Diabetic Retinopathy, Fundus image, Convolutional Neural Network, Deep Learning, Image classification

1. Introduction

Diabetic Retinopathy (DR) is an ocular condition stemming from Diabetes Mellitus (DM), characterized by inflammation and rupture of retinal blood vessels, leading to the formation of various irregular retinal lesions. It is commonly observed in individuals with prolonged diabetes, typically spanning 10–15 years. Statistics indicate that approximately 80% of diabetic patients with extended diabetes durations exhibit diverse stages of DR [1]. In India alone, an estimated 73 million individuals are afflicted with diabetes [2]. Timely consultation of diabetic patients with experts for DR assessment has become imperative. Early diagnosis can mitigate DR progression and reduce the risk of severe blindness [3]. Without proper diagnosis, DR advances through various stages characterized by the formation of lesions such as Microaneurysms (MAs), Exudates (EXs), Hemorrhages (HEs), Cotton Wool Spots (CWSs), abnormal structure of the Optic Disc (OD), abnormal Foveal Avascular Zone (FAZ), Neovascularizations, Intra Retinal Microvascular Abnormalities (IRMAs), among others [4]. The presence of these lesions can lead to vision impairment, and as the disease progresses to severe stages, it can culminate in complete blindness [5].

The experts of the domain, called as Ophthalmologist examines the human retina using a high-resolution digitized fundus camera for capturing fundus image. The fundus perceives various DR retinal lesions for image annotation and follow-up necessary treatments. The process of physical diagnosis is painstaking, and the dearth of unavailability of proper means for treatment makes early detection of DR, a challenging task. Thus, such a medical disorder necessitates advanced means for proper diagnosis and treatment. To overcome the challenges of labor-intensive DR detection, researchers have proposed intelligent expert systems, using Deep Learning (DL) for analysis and in-depth study of DR features from fundus images. Intelligent systems [6] are effective with respect to time, feature extraction, error recognition, and early diagnosis and treatment compared to traditional methods. These smart systems take fundus images as input, which are enhanced and analyzed for extraction of significant features, for classification of DR as moderate DR, moderate NODR, mild DR, severe DR, Proliferative DR [7]. Thus, intelligent systems using DL are proposed as an early and potentially scalable alternative for DR detection. Conventional ML models and data analysis approaches are shallow in nature, and have shown poor performance in learning and training complex non- linear features from larger datasets and hence, are unable to exhibit better analysis and interpretation [8]. Besides, DL based CNN models [9] have overshadowed ML models in performance, through inbuilt preprocessing, convolutional operations, better learning and generalization with deeper networks [10], less overfitting , data imbalance mitigation , optimization etc. Therefore, various state-of-the-art ML-based models [11] called DL models are proposed for deep feature extraction and image [12] classification tasks. The inherent caliber of such models due to their hierarchical structures, enhances the learning of the model, to achieve state-of-the-art performance. In diagnosis of DR, different DL architectures proposed earlier, have exhibited different feature extraction and classification performances. However, in a single experimental set up, the effectiveness of assessing a comprehensive DL model is not exhaustive, and is unreliable. Therefore, this paper proposes a DL model entitled Diabetic Retinopathy Feature Extraction and Classification (DRFEC), an intelligent expert system using DL architectures, for an exhaustive comprehensive evaluation for detection of DR. The proposed approaches is an intensive meta-analysis of 26 pretrained DL networks for identification of the optimal architecture suitable for a larger dataset such as Kaggle EyePACS DR detection with significant skewness, and thereby identify architectural patterns and corresponding loopholes in performance, to ease the process of traditional methods of detec- tion. The model identifies suitable architecture-optimized goals to mitigate overfitting and poor generalization. The main contributions of the manuscript have been enlisted below:

The proposed model performs a comprehensive comparative assessment and evaluation to determine the behaviour of 5 DL models on the DR dataset.To conduct meta-analysis of traditional as well as state-of-the-art DL architectures on the highly skewed Kaggle DR detection dataset, with minimal resources and identify the best amongst them for future application and analysis.

The proposed model identifies high bias and high variance during DR image classification to identify the best DL classifier. This manuscript is sectionalized into various sections. Section 2 is an analysis of various related works and models proposed earlier, for identification of loopholes in early DR detection. Section 3 gives a detailed illustration of the proposed methodology, and establishes the significance of DL models for better feature extraction and classification. Section 4 illustrates the significance of the 5 DL models, based on the analysis and comparison of their performances, for determination of an optimal DL architecture(s) for DR detection. Section 5 concludes on a note identifying the suitable DL architecture(s) for DR detection.

2. Literature Review:

Diabetic Retinopathy (DR) is a chronic health condition that necessitates early detection and treatment [13]. Manual examination and detection of DR are unreliable and prone to errors, highlighting the importance of using intelligent systems for faster prediction. Therefore, researchers have explored advanced feature extraction and image classification techniques, especially using Machine Learning (ML) and Deep Learning (DL) methods, for early DR detection. DL techniques have shown superiority over ML-based approaches due to their ability to process large datasets efficiently, handle overfitting, generalize well, and provide accurate predictions. This literature review presents various works in the field of DR detection using DL techniques applied to fundus images [14].

Wang et al. [15] proposed a boosted Convolutional Neural Network (CNN) architecture using EfficientNet B3 for feature extraction from breast cancer images. Gurcan et al. [16] developed an automated DR classification system based on deep CNN and ML methods, achieving competitive classification accuracy. Sarki et al. [17] conducted a systematic study on preprocessing operations for Diabetic Eye Disease (DED) detection. Mayyaa et al. [18] conducted a methodical review on automated microaneurysm detection for DR, identifying various strengths and weaknesses. Hattiya et al. [19] evaluated different CNN architectures for DR detection using retina images.

Several studies have focused on transfer learning for DR detection. Kamal et al. [20] proposed a transfer learning model for DR detection, while Bodapati et al. [21] used transfer learning and feature extraction techniques for DR detection. Shah et al. [22] developed a DL model for distinguishing refera- ble DR using macula-centered fundus images. Pour et al. [23] utilized EfficientNet for feature extraction and classification in DR detection.

Furthermore, studies have investigated ensemble methods and model optimization techniques for DR detection. Ji et al. [24] proposed an augmented DNN model using ensemble learning for DR detec- tion, while Michele et al. [25] employed an ensemble CNN approach for palmprint recognition. Sau and Bansal [26] proposed a Fitness based Newly Updated Grasshopper Optimization Algorithm (FNU-GOA) for optimizing DL models in DR detection

Sarki et al. [27] conducted a methodical study on preprocessing operations for Diabetic Eye Dis- ease (DED) detection. They identified various preprocessing techniques such as Contrast-Limited Adap- tive Histogram Equalization (CLAHE), morphological operations, and image segmentation for blood vessel segmentation. These techniques were applied to datasets including DRISHTI-GS, Messidor, Reti- nal Dataset, and Messidor-2. The proposed automated classification framework utilized image enhancement, image augmentation, segmentation, and classification, aiming to address challenges associated with small datasets containing poor category DR samples.

Mayyaa et al. [28] conducted a methodical review focusing on the diagnostic use of automated microaneurysm detection for DR. They examined various methodologies and identified their strengths and weaknesses, highlighting the challenges that need to be addressed in designing effective algorithms for early diagnosis of DR.

Hattiya et al. [29] appraised the AlexNet DL mechanism as an ideal CNN architecture for DR de-tection. They compared different CNN architectures including MobileNet, DenseNet201, InceptionV3, ResNet50, NASNetMobile, and MNASNet using a dataset comprising 23,513 retina images. Their evaluation aimed to identify the most suitable CNN architecture for effective DR detection.

Chetoui et al. [30] introduced an EfficientNet-based feature extraction and classification model utilizing EfficientNet B7 and Global Average Pooling for DR detection. Their model leveraged datasets from

Kaggle EyePACs and APTOS 2019, employing Gradient-weighted Class Activation Mapping (Grad-CAM) to extract features such as EXs, HEs, and MAs.

Tymchenko et al. [31] proposed a DCNN encoder-based feature extraction method for DR detec-tion, utilizing pre-trained models including EfficientNet-B5, EfficientNet-B4, SE-ResNeXt50, and an ensemble of 20 models.

Chaturvedi et al. [32] suggested a modified DenseNet121 network for DR detection on the AP- TOS 2019 dataset, utilizing 3662 fundus images.

Samanta et al. [33] anticipated a fine-tuned DenseNet121 model trained on 3050 images for DR detec-tion. They tested the model's performance against various architectures including VGG16, InceptionV1, InceptionV3, InceptionV2, Xception, AlexNet, ResNet-50, and DenseNet.

Sarki et al. [34] proposed an inclusive assessment of 13 pretrained CNNs for DR detection using the MESSIDOR and Kaggle datasets.

Ji et al. [35] introduced an augmented DNN model employing Inception V3, DenseNet121, and ResNet50 for transfer learning to enhance computational proficiency and classification in DR detection. They analyzed the performance of various subnetworks for each DNN on large OCT image datasets comprising 83,484 images, achieving accuracy and stability with InceptionV3 and ResNet50. However, they noted that DenseNet121 showed no significant improvement, possibly due to its dense connection architecture.

Michele et al. [36] proposed a fine-tuned pretrained feature extraction and classification model utilizing MobileNetV2, dropout, and a linear SVM classifier for palmprint recognition. They utilized a PolyU palmprint dataset comprising 6000 images.

Hui et al. [37] anticipated a modified extreme inception-based U-Net segmentation module for extracting effective features using multitask learning and distance representation from remote sensing images. Their model demonstrated improved performance with multitasking on the datasets.

Jiang et al. [38] introduced an interpretable ensemble ResNets-based feature mining and classifi- cation approach for DR detection. Their DL model consisted of Inception V3, ResNet152, and Incep-tionResNetV2, combined with the Adaboost algorithm. They trained the model on 28,244 images and achieved superior performance using the integrated DL model compared to individual models.

Orlando et al. [39] proposed an ensemble LeNet-CNN approach for detecting MAs, HEs, and redlesions for DR detection. Their model utilized handcrafted features, CNN features, and a combination of both, achieving better results with the combination of CNN and handcrafted features.

Suriyal et al. [40] proposed a real-time DR model utilizing MobileNet on internet-deprived port- able devices for DR detection. They trained the model on 16,798 images.

Huang et al. [41] introduced a compact novel network architecture called CondenseNet, which combines dense connectivity with novel learned group convolutions for image classification tasks using CIFAR-10, ImageNet, and CIFAR-100 datasets. This novel CNN architecture achieved cost-efficient performance compared to other architectures such as MobileNets, DenseNet-190, and ShuffleNets.

Pogorelov et al. [42] compared global features extracted from gastrointestinal tract images using transfer learning models like ResNet50 and Inception V3. They proposed a modified CNN and evaluated the predictions of both ML and DL classifiers. Their results showed better performance using ResNet50 for feature extraction compared to Inception V3.

Gulshan et al. [43] proposed a DL InceptionV3 model for the detection of DR and diabetic macu-lar edema in fundus images. Their methodology utilized the EyePACS-1 dataset and MESSIDOR-2 da-

taset. They employed an ensemble of ten networks, and their linear average was used for prediction on 128,175 images.
Nneji et al. [44] proposed a two-channel preprocessing weighted fusion deep learning networkfor the detection of DR using fundus images. They utilized CLAHE fundus images and contrast- enhanced canny edge detection (CECED) fundus images from Kaggle and MESSIDOR datasets. Their model employed VGG-16 and a modified Inception-V3 for feature extraction and merged the channel output using a weighted fusion approach.
Dong et al. [45] introduced a DL model using InceptionV3 and VGG-16 for the detection of DR, which was compared with mainstream models like GoogLeNet, AlexNet, and ResNet50 using 2693 wide-field optical coherence tomography augmented images. They claimed that wide-field optical co- herence tomography images are better than regular optical coherence tomography images due to their better field-of-view (FOV). However, the authors acknowledged limitations in their model due to a very limited dataset, which could affect its reliability.
Padmanayana and Anoop [46] proposed a CNN model for the detection of DR, comparing the perfor- mance of various optimizers such as Adagrad, RMSProp with momentum, and Adam. They used the APTOS 2019 Kaggle dataset for training and 1000 images collected from a private institute for testing. Preprocessing included weighted CLAHE, Gaussian blur, and Ben Grahams fraction maxpooling.
Sivapriya et al. [47] proposed a Recurrent Neural Network (RNN) to identify hard exudates (EXs) for the detection of DR. They used a limited dataset of 400 images from the MESSIDOR dataset and performed various preprocessing operations without significant changes in the RNN architecture. However, their approach reflected conventional and unreliable behavior compared to earlier works.
Bora et al. [48] introduced an automated risk analysis system for the detection of DR develop- ment in patients with diabetes but without DR. They reported a risk stratification tool that learns various associated risk factors leading to the development of DR from the stage of no DR. Their analysis utilized a large retrospective longitudinal dataset, including over half a million eyes/single images.
Deepa et al. [49] proposed a multistage ensemble DCNN using InceptionV3 and Xception for deep feature extraction and SVM-based ensemble classification for the detection of DR. Their architec- ture involved concatenating multi-stage patch-based and image-based probability vectors. They utilized voting and stacking for probability vector concatenation and employed an ensemble of SVM classifiers for further prediction on the ensembled classification output. However, their proposed architecture was computationally expensive and time-consuming due to its complexity.
Saeed et al. [50] developed a two-stage Principal Component Analysis (PCA) based transfer learning algorithm that initializes the pretrained model with extracted (64 × 64) Regions of Interest (ROIs) of Microaneurysms (MAs), Exudates (EXs), and normal features from fundus images of Kaggle's EyePACs and MESSIDOR datasets for DR detection. They introduced fixed-dimension based adaptive maxpooling to predict the label of the ROI and compared the performance of DL models including VGG-19, ResNet152, and Dual Path Network 107 (DPN107), where the fully connected layers are replaced with PCA for unsupervised feature discrimination. They used Decision Tree (DT), Random Forest (RF), and Gradient Boosting (GB) for classification and found that ResNet152 outperformed the other models.
Tsai et al. [51] proposed a DL model using Inception V3, ResNet101, and DenseNet121 on both the global Kaggle's EyePACs dataset and a local dataset from Taipei City Hospital (TCH) for DR detection. They observed a higher overestimation rate on the local dataset compared to the global dataset due to di-

ferences in regional and ethnic factors. DenseNet121 performed better than InceptionV3 andResNet101. However, their model was biased and failed to learn region-specific features.

Atwany et al. [52] performed a literature review analyzing the significance of supervised, semi-supervised, and vision transformer methodologies and learning paradigms for DL models in DR detection. They used Kaggle's EyePACs dataset, DDR dataset, and fundus images from Beijing Tongren Eye Centre for the study. They concluded that supervised learning is not suitable for noisy data and identified Semi-Supervised Learning (SSL) as effective but unexplainable, and less prone to inductive bias. However, SSL was found to be non-robust for small-scale datasets. They also reviewed amalgamation techniques for dataset synthesis such as Generative Adversarial Network (GAN) and Variational Autoencoders (VAE), and identified less complex DL attention models like Vision Transformers (ViT).

Das et al. [53] proposed a two-way CNN classifier incorporating a Squeeze-and-Excitation memory module and CNN architecture for DR detection using DIARETDB1 and a local dataset. They emphasized the importance of data augmentation for improved model performance.

Lim et al. [54] conducted a literature review on gradient-based interpretability methods in DL models for DR detection, including saliency maps, integrated gradients, layer-wise relevance propagation, occlusion testing, sensitivity analysis, class activation maps, gradient-weighted class activation maps, and layer- wise relevance propagation. They highlighted the drawbacks of these methods in accurately detecting lesions for a given class and the lack of reliable ground truth.

AbdelMaksoud et al. [55] proposed an integrated CNN model called E-DenseNetBC-121, which combines EyeNet and DenseNet architectures for DR detection using datasets such as EyePACS, IDRiD, MESSIDOR, and APTOS 2019.

Li et al. [56] developed a DL algorithm-based software for grading 1674 images for DR detection but failed to detect Diabetic Macular Edema (DME), which is responsible for early DR detection. They used fundus images from the publicly available EyePACs dataset and Shanghai General Hospital.

Deepa et al. [57] proposed a comprehensive two-phase feature extraction algorithm for DR detection, utilizing Xception architecture along with textural and transform-based techniques to detect Microaneurysms (MAs). They employed a Siamese Network-based CNN for hierarchical clustering, followed by patch-wise local and global feature extraction using a fine-tuned Xception model. The extracted features from 2290 images were classified using a Radial Basis Function (RBF) kernel-based SVM and compared with other classifiers such as Random Forest (RF), Adaboost, and Multilayer Perceptron (MLP).

Sau and Bansal [58] introduced a Fitness-based Newly Updated Grasshopper Optimization Algo-rithm (FNU-GOA) to optimize a DL model and threshold values in the active contour method for seg- menting blood vessels, MAs, EXs, and HEs for DR detection. They compared FNU-GOA with other optimization algorithms such as Particle Swarm Optimization (PSO), Grey Wolf Optimization Algo-rithm (GWO), Whale Optimization Algorithm (WOA), and Grasshopper Optimization Algorithm (GOA), as well as ML classifiers like Neural Network (NN), RNN, Long Short Term Memory (LSTM), and Deep NN. However, the meta-heuristic optimized algorithm achieved poor specificity and failed to classify negative samples correctly.

Shaik and Cherukuri [59] proposed a multi-stage end-to-end DNN pipeline called Hinge Atten- tion Network (HA-Net), which utilized gated attention VGG-16 discriminator and reconstruction auto-encoder to produce attention maps. They aimed to learn latent representations in images from Kaggle's

APTOS 2019 dataset and ISBI-2018 IDRiD2 dataset for DR detection using various optimizers. However, the model achieved poor accuracy on a limited-graded dataset, employing various attention descriptors with minimal samples. It struggled with the curse of dimensionality, failed to generate a decision boundary, learn inter-spatial and inter-channel correlation in latent features, and reduce latent spa- tial representations learned from baseline models such as VGG-16, VGG-19, ResNet50, ResNet50V2, Xception, MobileNet, Inception V3, and InceptionResNetV2, as well as from overlapping data.

The proposed DRFEC (Diabetic Retinopathy Feature Extraction and Classification) framework aims to address the limitations of conventional models in diabetic retinopathy (DR) detection by leveraging state-of-the-art deep learning (DL) techniques. Conventional models often lack depth and fail to capture the complex patterns in DR data, while DL models have demonstrated superior performance due to their ability to learn intricate features from large datasets.

One of the main drawbacks of conventional models is their reliance on small and limiteddatasets, which hinders their ability to generalize well to real-world scenarios. Additionally, the imbalance and lack of features in datasets like Kaggle's EyePACS dataset further restrict their effectiveness. Moreover, processing larger datasets through data augmentation can lead to data explosion, posing challenges in terms of resource constraints.

Existing models often exhibit similar performances and struggle to improve significantly due to their reliance on a similar number of images and failure to address regional and ethnic biases. Furthermore, these models tend to exhibit high variance, indicating poor learning processes and limited generalizability to real-time datasets.

The proposed DRFEC framework seeks to overcome these challenges by training a DL model on an imbalanced dataset of DR images. The framework will gradually optimize the baseline model through hyperparameter tuning, model training, and evaluation. By leveraging DL techniques and systematically analyzing DR data from its inception, the framework aims to identify and differentiate parameters responsible for various performance outcomes and ultimately improve DR detection accuracy.

1.3 The proposed DRFEC:

DRFEC system aims to automate the detection of diabetic retinopathy (DR) using machine learning (ML) techniques, particularly deep learning (DL). Manual examination of DR is slow and not suitable for early detection, necessitating the development of an efficient automated system.

The DRFEC system is designed to perform preprocessing, deep feature extraction, and image classification using a comprehensive DL model. This model includes a wide range of architectures, such as VGG-19, VGG-16, DenseNet121, ResNet50 and ResNet101.

These DL models are pre-trained on the ImageNet dataset, which contains a vast number of la- beled images across various categories. By leveraging these pre-trained models, the DRFEC system can effectively extract features from retinal images without the need for extensive manual feature engineering.

The DRFEC system is trained and evaluated on the Kaggle DR Detection dataset obtained from EyePACS. This dataset contains a large number of retinal images labeled with DR severity levels, making it suitable for training and evaluating the performance of the DL models. By leveraging a diverse set of DL architectures and pre-trained models, the DRFEC system aims to achieve robust and accurate detection of DR, enabling early diagnosis and intervention to prevent vision loss in patients with diabetes.

It seems like you're describing the technical setup and workflow of the DRFEC (Diabetic Reti- nopathy Feature Extraction and Classification) system. Figure 1 likely illustrates the steps involved in deploying DL (deep learning) models for DR image classification. Here's a breakdown of the compo- nents and steps mentioned:

Technical Specifications:

Operating System: 64-bit operating systemProcessor: ×64-based processor

OS Version: Windows 10 Pro

DL Package: Keras from TensorFlowProgramming Language: Python 3.8

TensorFlow Version: 2.4RAM: 64GB

CPU: Intel(R) Xeon(R) W-2155 CPU @ 3.30GHz 3.31 GHz

1.3.1 The proposed system Workflow Steps:

1.3.1.1. Data Pre-processing: This step likely involves cleaning, resizing, and normalizing the input im-ages to prepare them for feeding into the DL models.

1.3.1.2 Feature Extraction: DL models are used to extract deep features from the preprocessed images. This process leverages representation learning to capture intricate patterns and structures in the images.

1.3.1.3. Model Training: The extracted features are used to train the DL models. This involves optimiz-ing the model parameters to minimize the classification error.

1.3.1.4. Model Evaluation: The trained models are evaluated using a validation dataset to assess theirperformance in classifying DR images accurately.

1.3.1.5 Model Selection: Based on the evaluation results, the most suitable DL model is selected for DR image classification.

1.3.1.6 Deployment: The selected DL model is deployed for real-world applications, where it can auto- matically classify DR images with high accuracy.

the DRFEC system combines state-of-the-art DL models with efficient preprocessing techniques to au- tomate the detection and classification of diabetic retinopathy from retinal images. The proposed DRFEC for DR detection at an early stage shown in below figure 1.

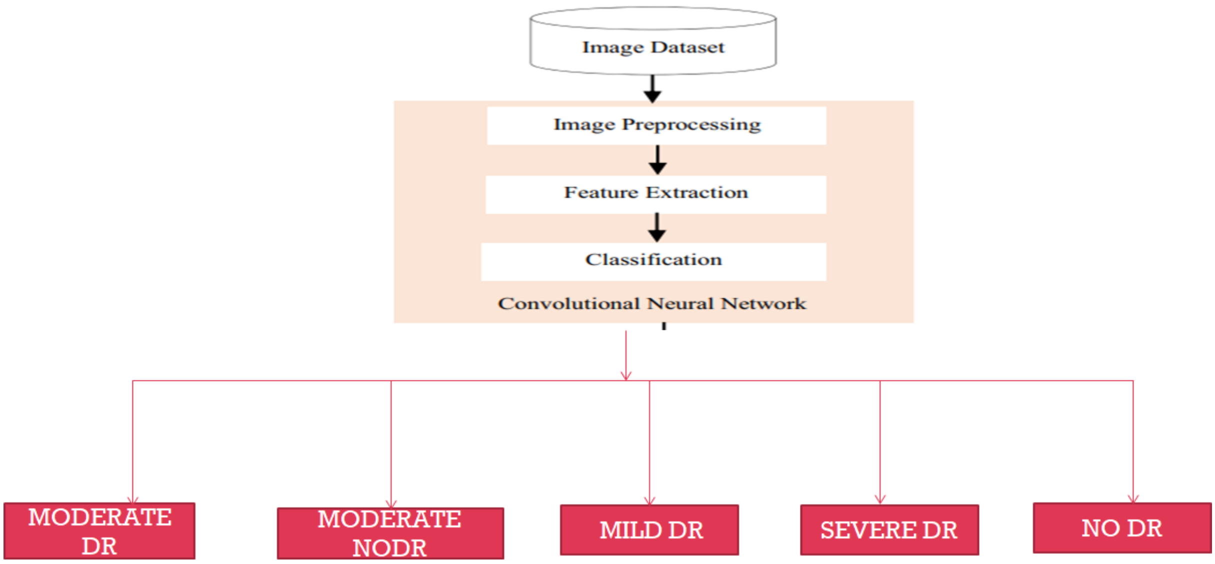


Figure1: Layout of the proposed DRFEC for DR detection at an early stage

In the training dataset, there are a total of 236 images labeled as Proliferate_DR, 154 images labeled as Severe, 799 images labeled as Moderate, 1444 images labeled as No_DR, and 296 images labeled as Mild.

Table 1 :Number of labelled images

Stage of DR	No. of fundus images
No DR	1444
Mild DR	296
Moderate DR	799
Severe DR	154
PDR	236

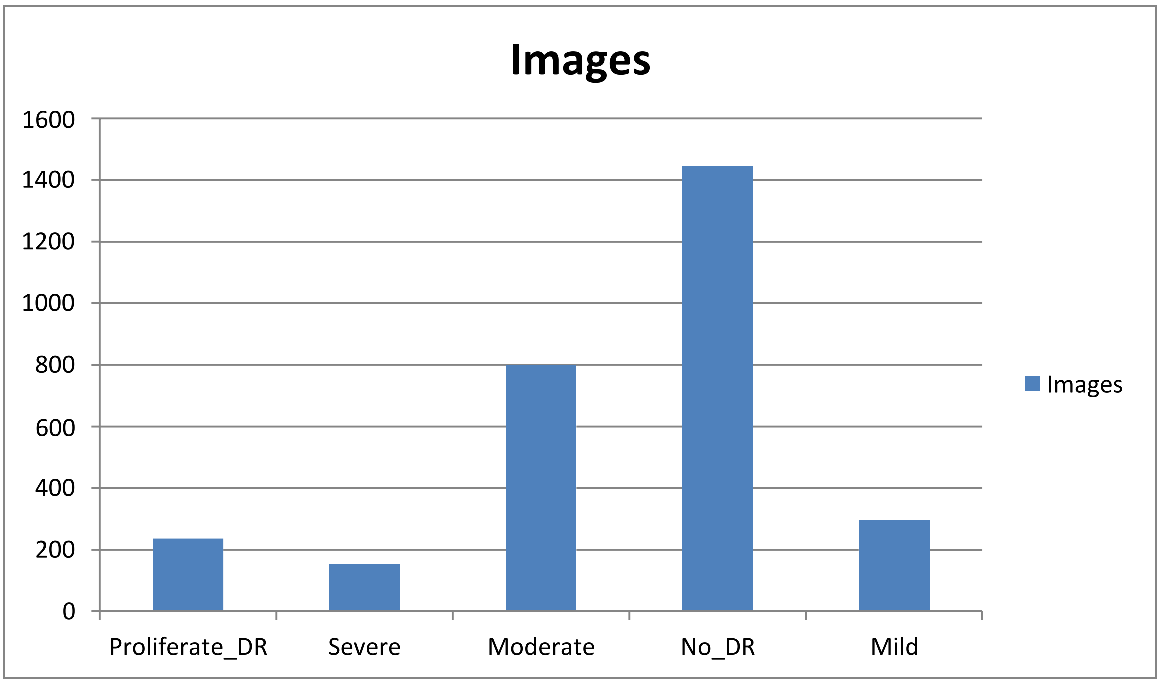


Figure2: Long tail distribution of the skewed DR Dataset

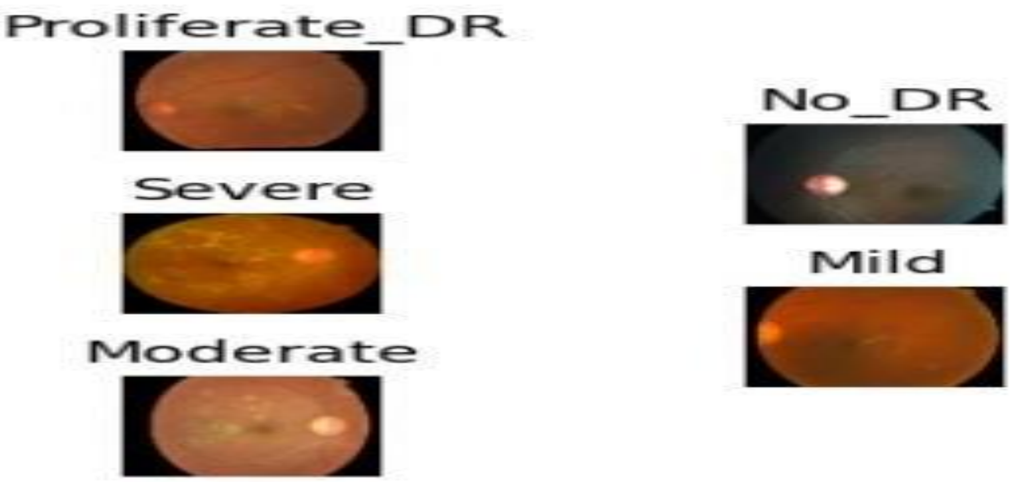


Figure3: Different types of Sample input

3.4 VGG-16:

VGG-16 is a convolutional neural network architecture proposed by Karen Simonyan and Andrew Zisserman from the Visual Geometry Group (VGG) at the University of Oxford. It was introduced in their paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" in 2014.

The architecture is characterized by its simplicity and depth. It consists of 16 layers, hence the name "VGG-16," with 13 convolutional layers followed by 3 fully connected layers. The convolutional layers use small 3x3 filters with a stride of 1 and same padding, and max-pooling layers with 2x2 filters and a stride of 2. This consistent architecture of stacking convolutional layers enables VGG-16 to learn intricate features at different scales.

Although VGG-16 has been surpassed in terms of performance by more modern architectures like ResNet and Inception, it remains influential and is often used as a baseline in many computer vision tasks and benchmarks. It is also a common choice for transfer learning due to its availability and simplicity. The Diagram of VGG 16 is shown in below Figure 4.

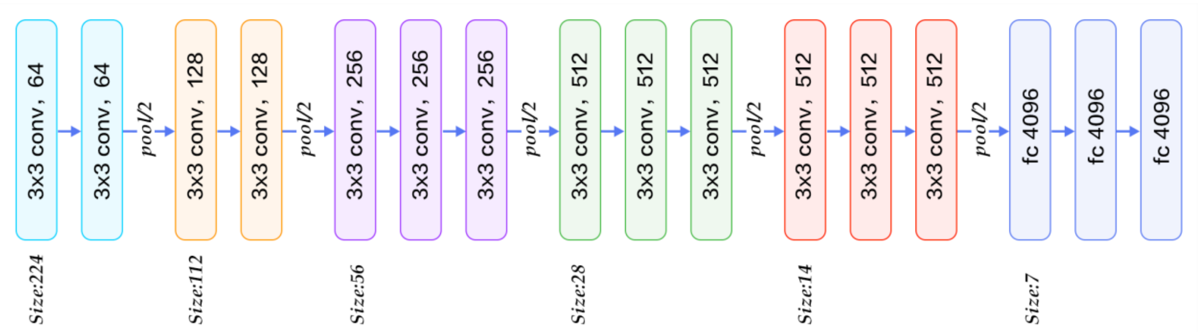


Figure 4: Structure of VGG 16

3.5 VGG-19:

VGG19 is an extension of the VGG16 architecture, proposed by the same authors, Karen Simonyan and Andrew Zisserman from the Visual Geometry Group (VGG) at the University of Oxford. It was introduced in their paper "Very Deep Convolutional Networks for Large-Scale Image Recognition" in 2014.

As the name suggests, VGG19 consists of 19 layers, making it deeper than VGG16. It has a similar architecture to VGG16, with 16 convolutional layers and 3 fully connected layers. The main difference between VGG19 and VGG16 is the increased depth achieved by adding more convolutional layers.

Like VGG16, VGG19 utilizes small 3x3 filters with a stride of 1 and same padding for convolutional layers, and 2x2 max-pooling layers with a stride of 2. This design choice aims to learn hierarchical features of increasing complexity as information passes through the network.

While VGG19 offers increased representational power compared to VGG16, it also comes with higher computational complexity and memory requirements. As with VGG16, VGG19 has been widely used as a benchmark architecture in various computer vision tasks and is often employed for transfer learning due to its availability and effectiveness. The Diagram of VGG 16 is shown in below Figure 5.

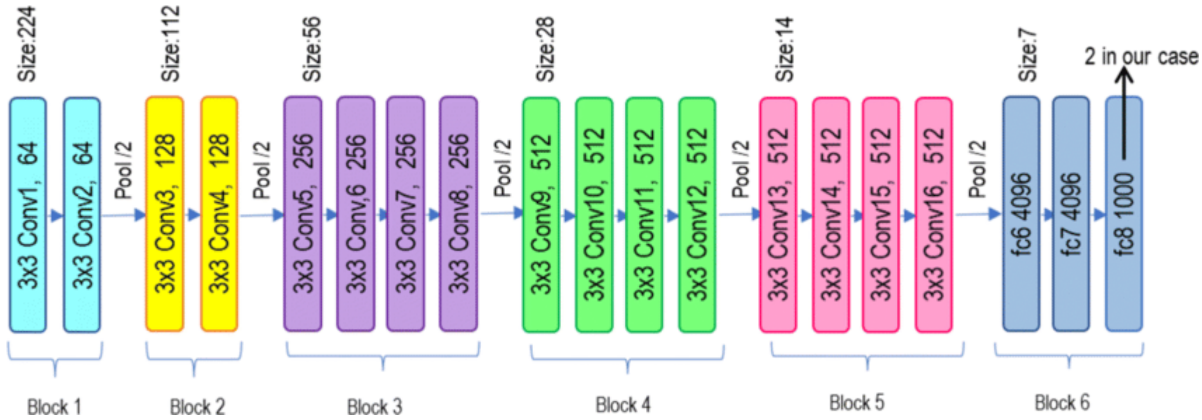


Figure 5 : Structure of VGG19

3.6 ResNet101:

ResNet101 is a convolutional neural network architecture that belongs to the family of Residual Networks (ResNets). It was proposed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun from Microsoft Research in their paper "Deep Residual Learning for Image Recognition" in 2016.

The main innovation introduced by ResNet is the residual learning framework, which addresses the problem of vanishing gradients in very deep neural networks. In traditional deep networks, adding more layers can sometimes degrade performance due to difficulties in training caused by vanishing gradients. ResNet tackles this issue by introducing skip connections, also known as identity shortcuts or skip connections, that skip one or more layers. These skip connections allow the gradients to flow more easily during training, enabling the training of very deep networks with hundreds of layers.

ResNet101 specifically refers to a ResNet architecture with 101 layers. It consists of a series of convolutional layers, batch normalization, ReLU activations, and residual blocks. These residual blocks contain skip connections, which help propagate gradients effectively during training.

ResNet101 has been widely adopted in various computer vision tasks, such as image classification, object detection, and image segmentation, due to its excellent performance and ability to train very deep networks effectively. It is often used as a benchmark architecture and as a basis for transfer learning in many deep learning projects. The Diagram of ResNet101is shown in below Figure 6.

3.7 ResNet50:

ResNet50 is a convolutional neural network architecture that is part of the ResNet (Residual Network) family. It was introduced by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun from Microsoft Research in their paper "Deep Residual Learning for Image Recognition" in 2015.

ResNet50, as the name suggests, consists of 50 layers. It builds upon the basic idea of residual learning introduced in the original ResNet paper. The key innovation of ResNet is the use of skip connections or shortcuts, which allow the network to bypass one or more layers, facilitating the flow of gradients during training. This helps in training very deep neural networks effectively by mitigating the vanishing gradient problem.

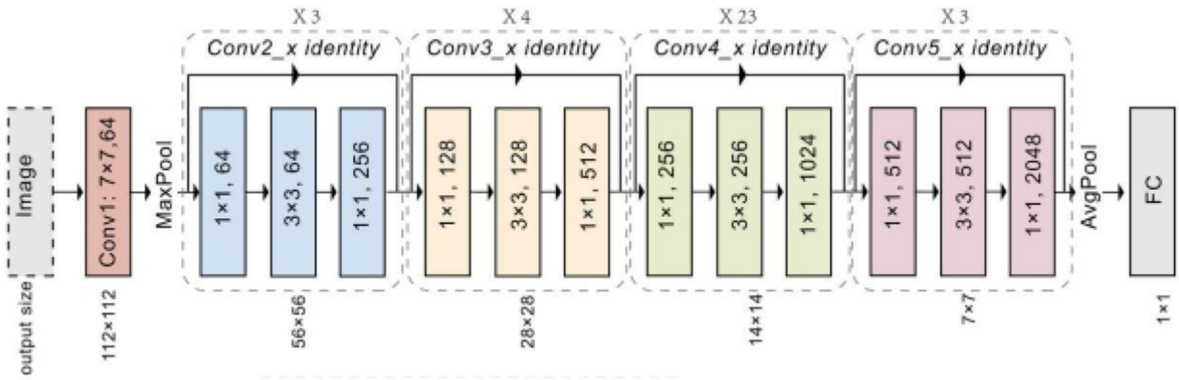
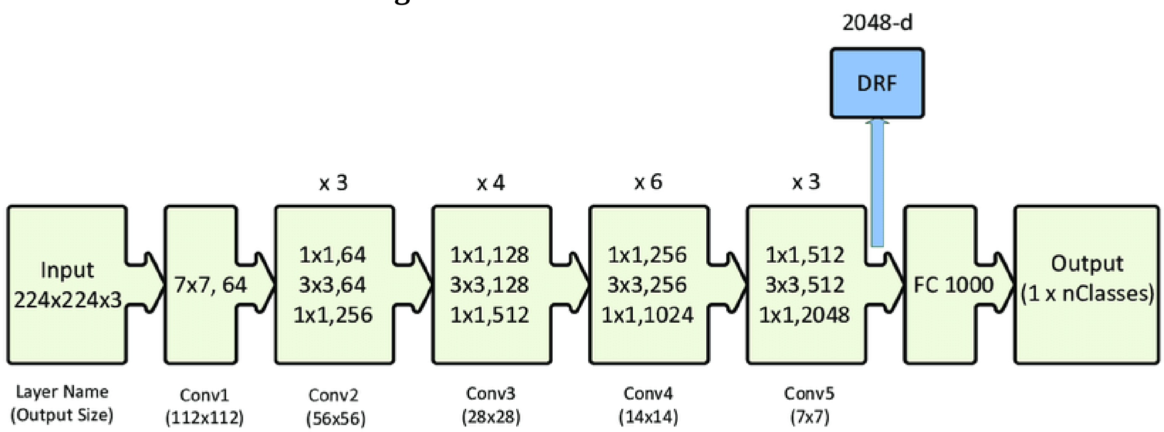


Figure 6: Structure of ResNet101

The architecture of ResNet50 includes convolutional layers, batch normalization, ReLU activations, and residual blocks. These residual blocks contain skip connections, which enable thenetwork to learn more efficiently. ResNet50 has a specific configuration of residual blocks that balances model complexity with computational efficiency.

ResNet50 has been widely adopted in various computer vision tasks, including image classification, object detection, and image segmentation. It has achieved state-of-the-art performance on benchmark datasets and is often used as baseline architecture for comparison in research and practical applications. Additionally, ResNet50 is commonly used for transfer learning, where pre-trained models are fine-tuned on specific tasks due to its availability and effectiveness. The Diagram of ResNet50is shown in below Figure 7.

Figure 7: Structure of ResNet50



3.8 DenseNet121:

DenseNet121 is a convolutional neural network architecture introduced by Gao Huang, Zhuang Liu, Laurens van der Maaten, and Kilian Q. Weinberger in their paper "Densely Connected Convolutional Networks" in 2017.

The "Dense" in DenseNet refers to the dense connectivity pattern within the network architecture. Unlike traditional convolutional neural networks where each layer is connected only to the subsequent layers, DenseNet establishes direct connections between all layers, resulting in densely connected blocks. These dense connections facilitate feature reuse and gradient flow throughout the network, leading to improved learning efficiency and better performance, particularly in scenarios with limited training data.

DenseNet121 specifically refers to a DenseNet architecture with 121 layers. It comprises multiple dense blocks, each containing a series of convolutional layers with batch normalization and ReLU activation, as well as skip connections that concatenate the feature maps from all preceding layers. The dense blocks are separated by transition layers, which include convolutional layers followed by pooling operations to reduce the spatial dimensions of the feature maps.

DenseNet architectures, including DenseNet121, have demonstrated impressive performance on various computer vision tasks, such as image classification, object detection, and image segmentation. They have become popular choices due to their efficient use of parameters, strong feature propagation, and high accuracy. Additionally, DenseNet models are often used for transfer learning, where pre- trained models are fine-tuned on specific tasks, leveraging the learned features from large-scale datasets like ImageNet. The Diagram of DenseNet121 is shown in below Figure 8.

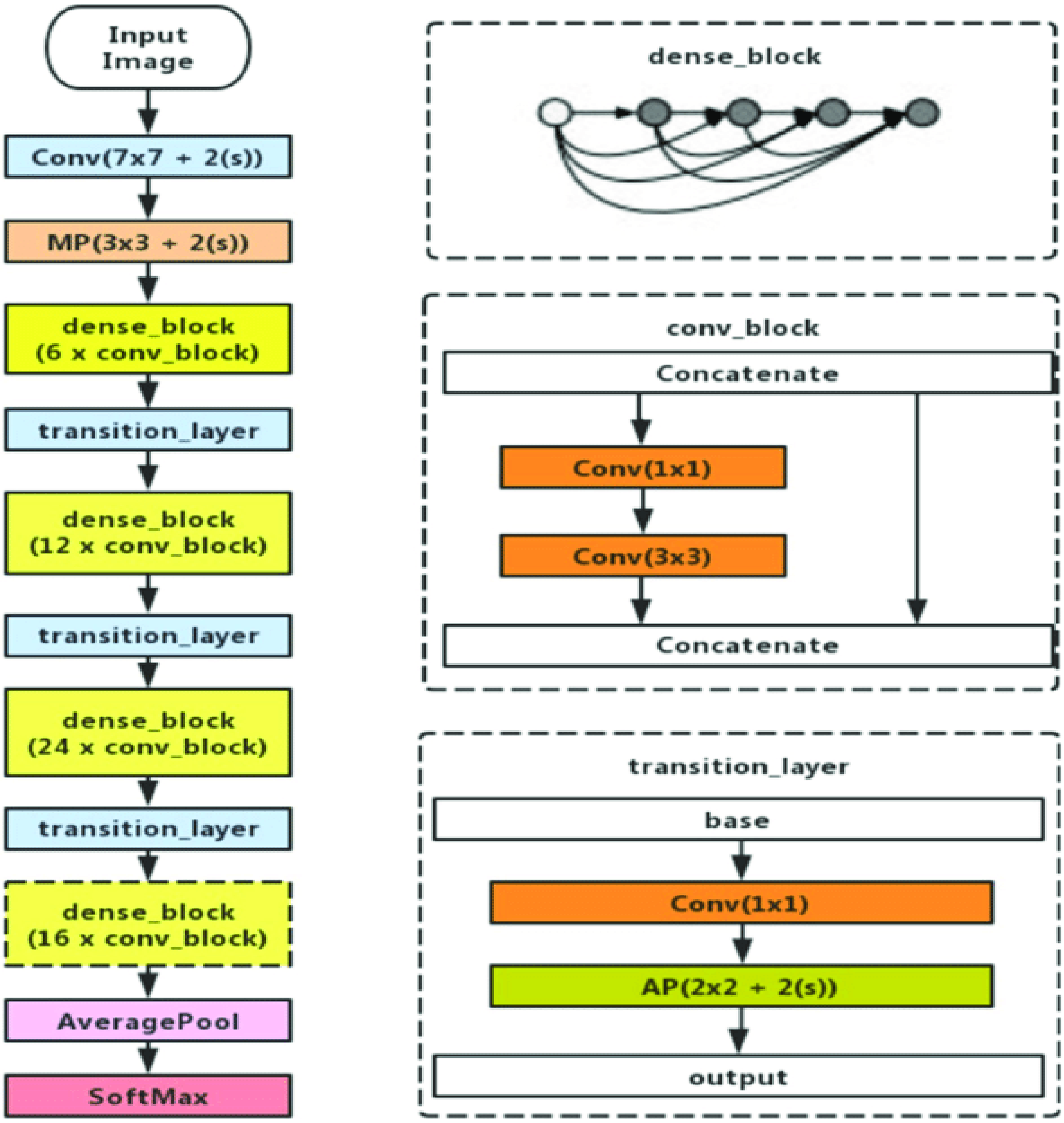


Figure 8: Structure of DenseNet121

4. Result Analysis:

The proposed model is a composition of DL CNN models such as VGG-16, VGG-19, DenseNet121, ResNet50 and ResNet101 which are trained on a training dataset of 27,446 fundus images and tested on a test dataset of 7680 fundus images, where both the training and test dataset contains images belonging to 5 classes.

-1	0	1
-1	0	1
-1	0	1

=

i-1, j+1	i, j+1	i+1, j+1
i-1,j	i, j	i+1, j
i-1, j-1	i, j-1	i+1, j-1

(i-1, j+1) x (-1) + (i, j+1) x0 + (i+1, j+1) x1+(i-1, j) x (-1) +(i, j) x0 + (i+1, j) x1 + (i-1, j-1) x (-1) + (i, j-1) x0+ (i+1, j-1) x1

3	0	1	2	7	4
1	5	8	9	3	1
2	7	2	5	1	3
0	1	3	1	7	8
4	2	1	6	2	8
2	4	5	2	3	9

1	0	-1
1	0	-1
1	0	-1

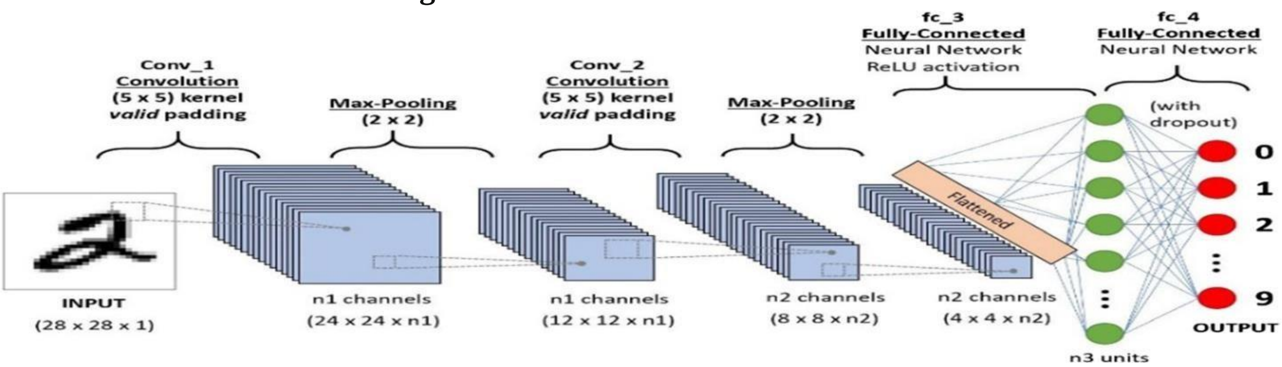
-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

=

Fig. 9 Convolution of image array with Prewitt kernel for detection of vertical edges

These DL models are applied to extract enhanced features necessary for better learning and image classification tasks. Figure 10 illustrates the functioning of a convolutional neural network (CNN), which takes a grayscale input image (28x28x1), performs convolution and pooling operationsfor feature extraction, and utilizes a fully connected neural network-based classification approach using ReLU activation, dropout, and softmax activation of DR. The DL models are individually trained and tested upon the dataset of interest to evaluate their performance.

Figure 10: Convolutional Neural Network



The proposed model illustrates and demonstrates the behaviour of these models in a chronological fashion, and has achieved different training and validation values. The DL models are trained for a target size of (224 × 224), using inbuilt preprocessing of the CNNs, a batch size of 32, andusing the fully connected classification layers for classification accompanied with softmax activation function. Besides, the DL models uses the state-of-the-art Adam optimizer, a learning rate of 0.0001 and categorical cross entropyloss function, for training for 50 epochs.

4.1. VGG-16:

The VGG-16 model has a total trainable parameter of 138,357,544 and 0 non-trainable parameters. On regularization, through minimization of two layers of the neural network, the model has a total trainable parameter of 134,281,029 and 0 non-trainable parameters, for a target size of (224 × 224), using the intermediate layers for output. It has achieved a training accuracy of 90.91% and validation accuracy of 62.07%, in the 50th epoch. It has also achieved a training loss of 0.26 and a validation loss of 2.32. It can be inferred that the training accuracy and validation loss are higher than validation accuracy and training loss, respectively. This implies that the model is complex and is overfitting.To regularize the model and prevent it from overfitting, the layers of the network are reduced and the number of neurons are minimized or dropped to reduce model parameters. Besides, other forms of regularization such as preprocessing, data augmentation and batch normalization can also be applied to regularize the model. Figure 11 depicts training and validation loss of VGG-16. Figure 12 depicts training and validation accuracy of VGG-16.

Figure 11: Training and validation loss of VGG-16

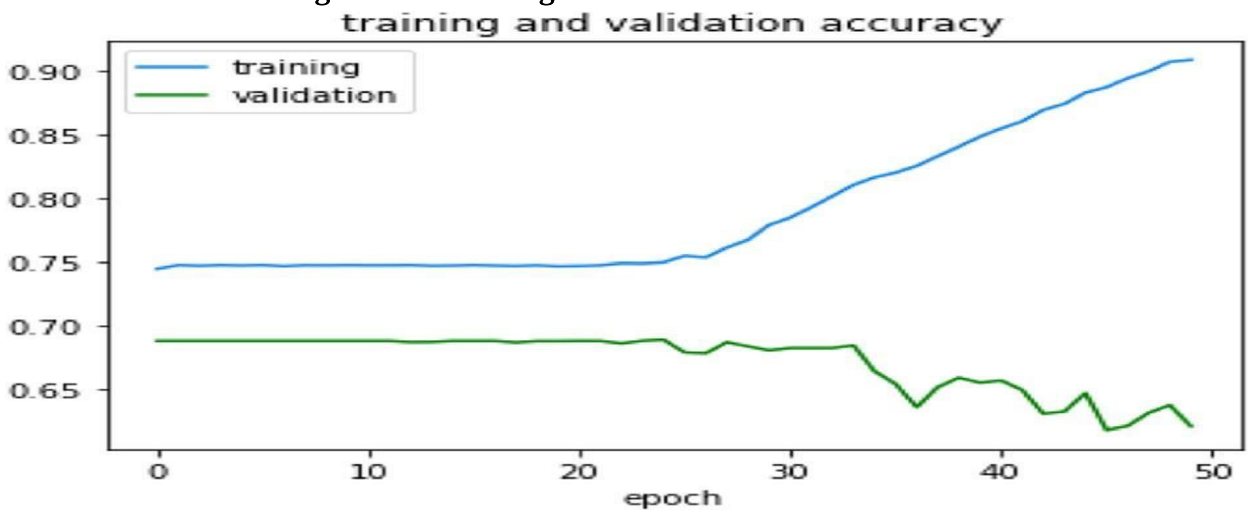
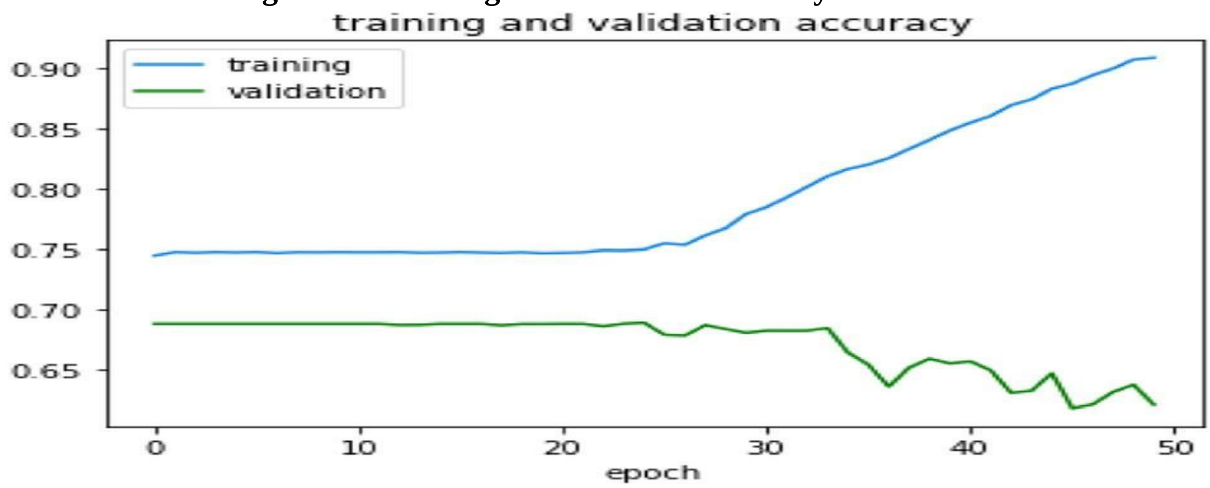


Figure 12 :Training and validation accuracy of VGG-16



4.2 VGG-19:

The VGG-19 model has a total trainable parameter of 143,667,240 and 0 non-trainable parameter. On regularization, through minimization of two layers of the neural network, the model has a total trainable parameter of 139,590,725 and 0 non-trainable parameter, for a target size of (224 × 224). It has achieved a training accuracy of 97.98% and validation accuracy of 73.37%, in the 50th epoch. It has also achieved a training loss of 0.062 and a validation loss of 2.07. It can be inferred that the training accuracy and validation loss are higher than validation accuracy and training loss, respectively. This implies that the model is complex and is overfitting. It has many parameters that are capable of memorizing the training data, and hence are only capable of extracting information and not create it. Figure 13 depicts training and validation loss of VGG-19. Figure 14 depicts training and validation accuracy ofVGG-19.

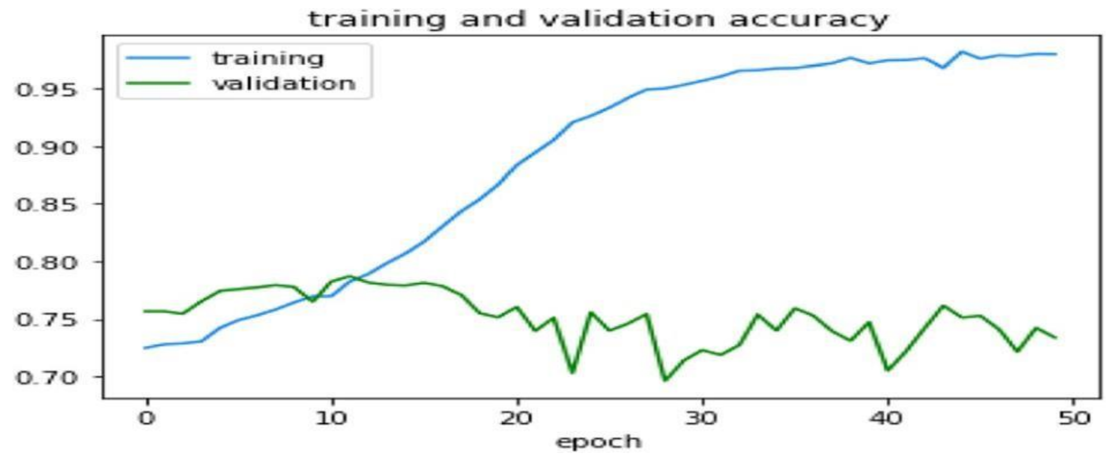


Figure 13: Training and validation loss of VGG-19

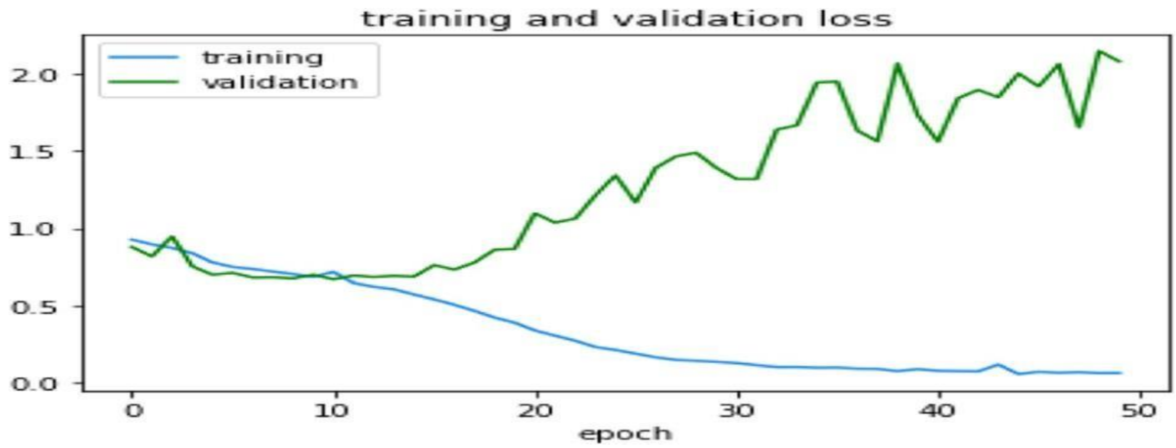


Figure 14 Training and validation accuracy of VGG-19

4.3 ResNet101:

The ResNet101 model has a total of 44,707,176 of which 44,601,832 parameters are trainable and 105,344 parameters are non-trainable, for a target size of (224 × 224). The model is regularized by reducing the number of layers of the neural network by 2. This has led the model to have a total of 42,668,421 of which 42,563,077 parameters are trainable and 105,344 parameters are non-trainable, for a target size of (224 × 224), upon the intermediate layers foroutput. It has achieved a training accuracy of 99.32% and a validation accuracy of 77.32%. It has also achieved a training loss of 0.0197 and a validation loss of 1.8714. It is observed that the training accuracy and validation loss of the model are higher than validation accuracy and training loss, respectively which has led to the overfitting of the model. On reducing and minimizing the number of layers in the neural network by 2, for the purpose of regularization to reduce the overfitting caused, has however brought minimal changes in the gap of training and validation loss. Figure 15 depicts training and validation loss of ResNet101. Figure 16 depicts training and validation accuracy of ResNet101.

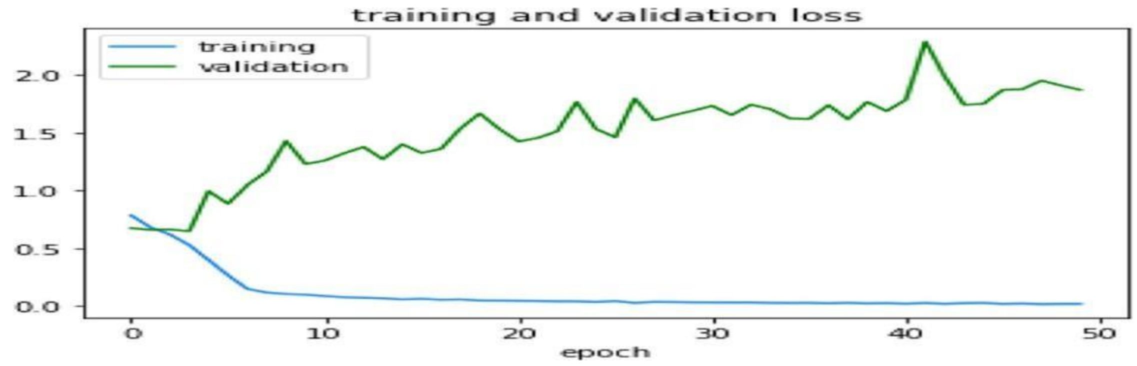


Figure 15: Training and Validation loss of ResNet101

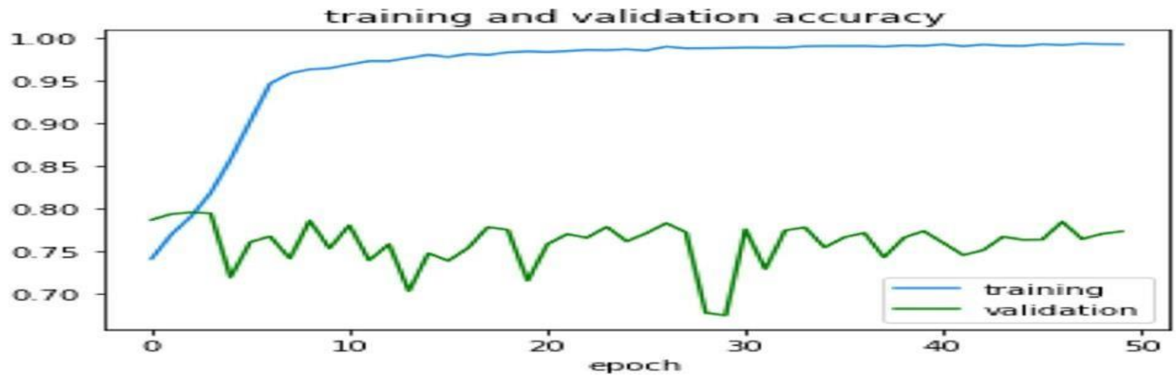


Figure 16: Training and Validation accuracy of ResNet101

4.4 ResNet50:

The ResNet50 model has a total of 25,636,712 of which 25,583,592 parameters are trainable and 53,120 parameters are non-trainable, for a target size of (224 × 224). The model is regularized by reducing the number of layers of the neural network by 2. This has led the model to have a total of 23,597,957 of which 23,544,837 parameters are trainable and 53,120 parameters are non-trainable, for a target size of (224 × 224), using the intermediate layers for output. It has achieved a training accuracy of 99.37% and a validation accuracy of 71.64%.

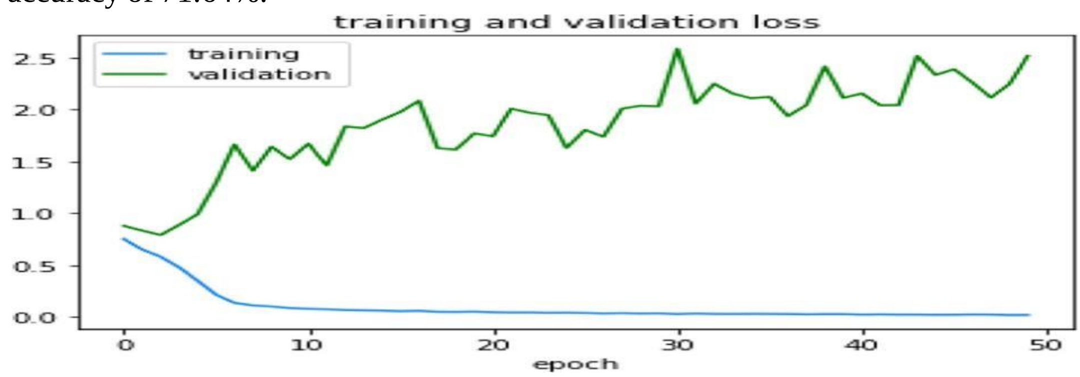


Figure 17: Training and Validation loss of ResNet50

It has also achieved a training loss of 0.0178 and a validation loss of 2.51. It is observed that the training accuracy and validation loss of the model are higher than validation accuracy and training loss, respectively which has led to the overfitting of the model. On reducing and minimizing the number of layers in the neural network by 2, for the purpose of regularization to reduce the overfitting caused, has however brought minimal changes in the gap of training and validation loss, except for minimization in the number of parameters. Figure 17 depicts training and validation loss of ResNet50. Figure 18 depicts training and validation accuracy ofResNet50.

Figure 18: Training and Validation accuracy of ResNet50

4.5 DenseNet121:

The DenseNet121 model has a total of 8,062,504 parameters of which 7,978,856 parametersare trainable and 83,648 parameters are non-trainable, for a target size of (224 × 224). The model is regularized by reducing the number of layers of the neural network by 2. This has ledthe model to have a total of 7,042,629 parameters of which 6,958,981 parameters are trainableand 83,648 parameters are non-trainable, for a target size of (224 × 224), using the interme- diate layers for output. It has achieved a training accuracy of 98.80% and a validation accuracy of 73.58%. It has also achieved a training loss of 0.0344 and a validation loss of 1.52. It is observed that the training accuracy and validation loss of the model are higher than validationaccuracy and training loss, respectively which has led to the overfitting of the model. On reducing and minimizing the number of layers in the neural network by 2, for the purpose ofregularization to reduce the overfitting caused, has however brought minimal changes in the gap of training and validation loss, except for minimization in the number of parameters. Figure 19 depicts training and validation loss of DenseNet121. Figure 20 depicts training and validation accuracy of DenseNet121.

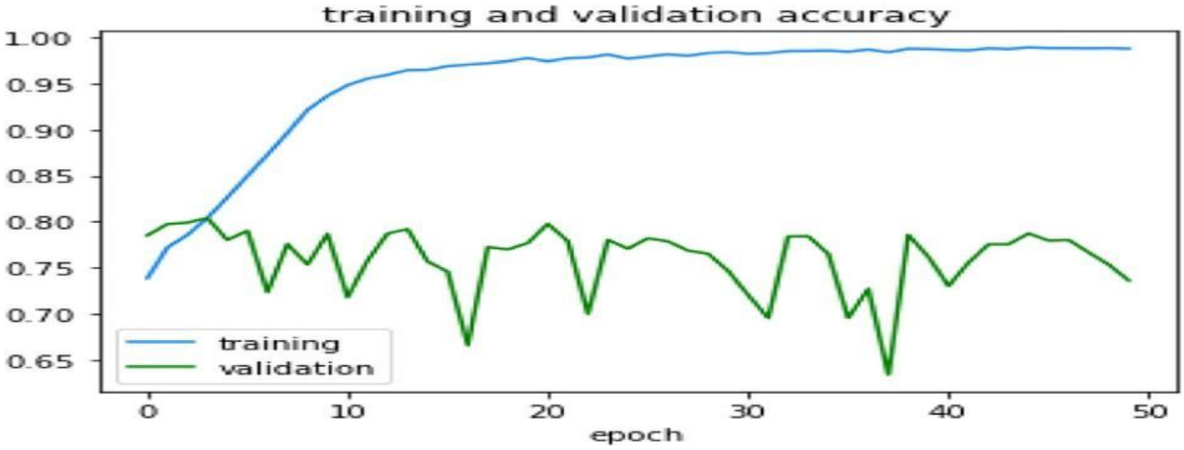


Figure 19: Training and validation accuracy of DenseNet121

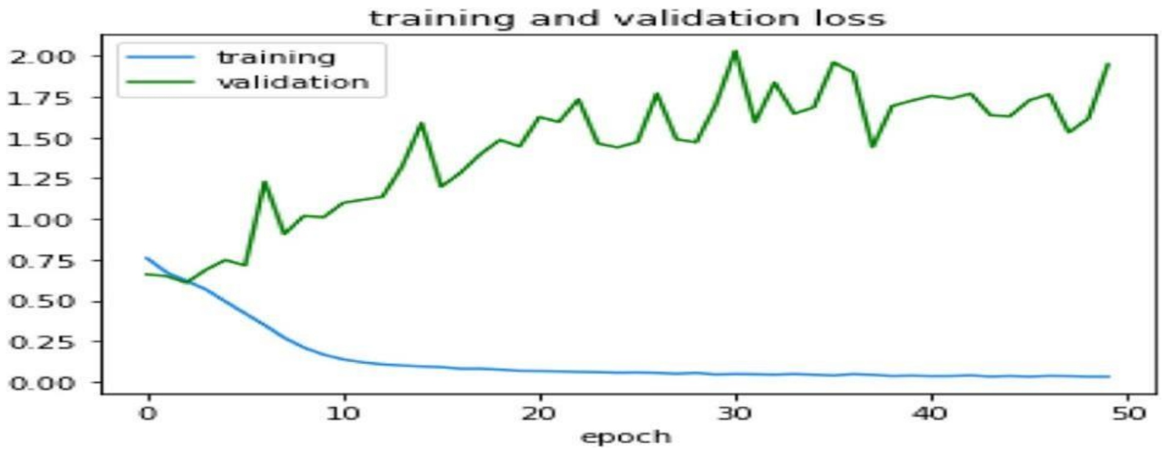


Figure 20: Training and validation loss of DenseNet121

The total summary of DRFEC methods are given below table-6.1 :

Table-2: Accuracy and lose of DRFEC methods

S.No	Method	Accuracy		Loss	
		Training	Validation	Training	Validation
1	**VGG16**	90.91%	62.07%	0.26%	2.32%
2	**VGG19**	97.98%	73.37%	0.062%	2.07%
3	**ResNet101**	99.32%	77.32%	0.019%	1.87%
4	**ResNet50**	99.37%	71.64%	0.017%	2.51%
5	**DenseNet121**	98.80%	73.58%	0.034%	1.52%

5. Conclusion:

This paper proposes an automated system for early detection of DR called Diabetic Retinopathy Feature Extraction and Classification (DRFEC) which employs DL CNN models such as VGG-16, VGG-19, DenseNet121,ResNet50 and ResNet101 for DR feature extraction and image classification. The proposed model performs an exhaustive analysis of these architectures upon fundus images, and derives the best performing DL architecture for DR feature extraction and fundus image classification. Amongst all the models, ResNet50 has achieved the highest training accu- racy whereas VGG-16 has achieved the lowest training accuracy. Again, VGG-16 has achieved lowest validation accuracy whereas ResNet101 has achieved highest validation accuracy. In the designed approach, the imbalanced dataset has caused overfitting of the models, but in addition to that the complexity of the DL models have also significantly contributed to overfitting, poor generalization, poor training time, poor gradient flow, and optimization and framework constraints. Moreover, VGG16 has shown highest overfitting whereas ResNet50has shown the lowest overfitting.

In future, the best DL architecture determined through the proposed work can be used to incorporate state-of-the-art techniques such as attention mechanism for extraction of relevant features, for detection of subtle lesions, for early detection of DR. The proposed model will further be extended through architectural engineering, and using a balanced fundus image benchmark dataset for comparison with benchmark models and to overcome the limited resource constraint, to achieve more convincing results. The model also aims to mitigate overfitting and poor generalization in future works, through fine-tuning of DL models for computation of significant evaluation metrics such as AUROC. The proposed baseline model shall be used for comparison with newly proposed future model for better research direction.

References

1. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, Ferris FL, Klein R (2004) Retinopathy in diabetes. Diabetes Care 27(1):84–87. https://doi.org/10.2337/diacare.27.2007.S84
2. Sosale B, Aravind SR, Murthy H, Narayana S, Sharma U, Gowda SGV, Naveenam M (2020) Simple, Mobile-based Artificial intelligence algorithm in the detection of diabetic Retinopathy (SMART) study.BMJ Open Diabetes Res Care 8(1):1–6. https://doi.org/10.1136/bmjdrc-2019-000892
3. Diabetic Retinopathy, Updated (2022), Available at: https://www.nei.nih.gov/learn-about-eye-health/eye- conditions-and-diseases/diabetic-retinopathy, Accessed on : 16-5-2022
4. Walter T, Klein JC, Massin P, Erginay A (2002) A contribution of image processing to the diagnosis ofdiabetic retinopathy-detection of exudates incolor fundus images of the human retina. IEEE Trans MedImaging 21(10):1236–1243. https://doi.org/10.1109/TMI.2002.806290
5. Shukla, UV, Tripathy, K (2022) Diabetic Retinopathy, in: StatPearls [internet], Treasure Island (FL): StatPearls Pub-lishing
6. Bhatti UA, Huang M, Wang H, Zhang Y, Mehmood A, Di W (2018) Recommendation system for immunization coverage and monitoring. Human Vaccines Immunotherapeutics 14(1):165–171. https:// doi.org/10.1080/21645515.2017.1379639
7. Goh JKH, Cheung CY, Sim SS, Tan PC, Tan GSW, Wong TY (2016) Retinal imaging techniques for diabetic retinopathy screening. J Diabetes Sci Technol 10(2):282–294. https://doi.org/10.1177/1932296816629491
8. Janiesch C, Zschech P, Heinrich K (2021) Machine learning and deep learning. Electron Mark 31:685–695. https://doi.org/10.1007/s12525-021-00475-2
9. Bhatti UA, Yu Z, Chanussot J, Zeeshan Z, Yuan L, Luo W, Nawaz SA, Bhatti MA, Ain Q, Mehmood A(2021) Local similarity-based spatial–spectral fusion hyperspectral image classification with deep CNN and Gabor filtering. IEEE Trans Geosci Remote Sens 60(5514215):1–15. https://doi.org/10.1109/TGRS.2021.3090410
10. Zhang, C, Bengio, S, Hardt, M, Recht, B, Vinyals, O (2017) Understanding Deep Learning Requires Rethinking Generalization, ICLR, pp.1–15, 2017 https://doi.org/10.48550/arXiv.1611.03530
11. Kishor A, Chakraborty C, Jeberson W (2020) A Novel Fog Computing Approach for Minimization of Latency in Healthcare using Machine Learning
, Special Issue on Current Trends in Intelligent Multimedia Processing Systems. Int J Interact Multimed Artif Intell 6(7):7–17. https://doi.org/10.9781/ijimai.2020.12.004
12. Bhatti UA, Yu Z, Chanussot J, Zeeshan Z, Yuan L, Luo W, Nawaz SA, Bhatti MA, Ain Q, Mehmood A(2021) Local similarity-based spatial–spectral fusion hyperspectral image classification with deep CNN and Gabor filtering. IEEE Trans Geosci Remote Sens 60(5514215):1–15. https://doi.org/10.1109/TGRS.2021.3090410
13. Kishor, A, Chakraborty, C (2021) Artifcial Intelligence and Internet of Things Based Healthcare 4.0 Monitoring System, Wireless Personal Commu- nications, pp.1–17 https://doi.org/10.1007/s11277-021- 08708-5
14. Hagos MT (2020) Point-of-care diabetic retinopathy diagnosis: a standalone mobile application

approach, arXiv:2002.04066v1 [eess.IV]. Vol. abs/2002.04066, pp 1–84. Available at: https://www.semanticscholar.o r g / pa p e r / P o i n t - of - C a r e - D i a b e t i c - Re t i n op a th y - D i a g n o s i s % 3 A - A H a g o s/ 3b9a121e53cea1b30cd29eea13854eeb4faadff

15. Wang J, Liu Q, Xie H, Yang Z, Zhou H (2021) Boosted EfficientNet: Detection of Lymph Node Metastases in Breast Cancer Using Convolutional Neural Networks. Cancers 13(4):1–14. https://doi.org/10.3390/ cancers13040661
16. Gurcan OF, Beyca OF, Dogan O (2021) A Comprehensive Study of Machine Learning Methods on Diabetic Retinopathy Classification. Int J Com- put Intell Syst 14(2):1132–1141. https://doi.org/10.2991/ ijcis.d.210316.001
17. Sarki R, Ahmed K, Wang H, Zhang Y, Ma J, Wang K (2021) Image preprocessing in classification andidentification of diabetic eye diseases. Data Sci Eng 6:455–471. https://doi.org/10.1007/s41019-021-00167-z
18. Mayyaa, V, Kamath, SS, Kulkarni, U (2021) Automated microaneurysms detection for early diagnosis ofdiabetic retinopathy: A Comprehensive review, Computer Methods and Programs in Biomedicine Update,1, pp-1-15 https://doi.org/10.1016/j.cmpbup.2021.100013
19. Hattiya T, Dittakan K, Musikasuwan S (2021) Diabetic Retinopathy Detection Using Convolutional Neural Network: A Comparative Study on Different Architectures. Mahasarakham Int J Eng Technol 7(1):50–60.https://doi.org/10.14456/mijet.2021.8
20. Kamal KC, Yin Z, Wu M, Wu Z (2021) Evaluation of deep learning-based approaches for COVID-19 classification based on chest X-ray images, sig- nal. Image Vid Process 15:959–966. https://doi.org/10.1007/s11760-020-01820-2
21. Bodapati JD, Veeranjaneyulu N, Shareef SN, Hakak S, Bilal M, Maddikunta PKR, Jo O (2020) Blendedmulti-modal deep convnet features for diabetic retinopathy severity prediction. Electronics 9(6):914. https:// doi.org/10.3390/electronics9060914
22. Shah P, Mishra DK, Shanmugam MP, Doshi B, Jayaraj H, Ramanjulu R (2020) Validation of Deep Convolutional Neural Network based algorithm for detection of diabetic retinopathy — Artificial intelligence versus clinician for screening. Indian J Ophthalmol 68(2):398–405. https://doi.org/10.1007/s11760-020- 01820-2
23. Pour AM, Seyedarabi H, Jahromi SHA, Javadzadeh A (2020) Automatic detection and monitoring of diabetic Retinopathy using efficient convolu- tional neural networks and contrast limited adaptive histogram equalization. IEEE Access 8:136668–136673. https://doi.org/10.1109/ACCESS.2020.3005044
24. Ji Q, Huang J, He W, Sun Y (2019) Optimized deep convolutional neural networks for identification ofmacular diseases from optical coherence to- mography images. Algorithms 12(3):1–12. https://doi.org/10.3390/a12030051
25. Michele A, Colin V, Santika DD (2019) MobileNet convolutional neural networks and support vector Machines for Palmprint Recognition. Procedia Comput Sci 157:110–117. https://doi.org/10.1016/j.procs.2019.08.147
26. Sau, PC, Bansal, A (2022) A novel diabetic retinopathy grading using modified deep neural network withsegmentation of blood vessels and retinal ab- normalities, Multimedia Tools and Applications, pp.1–29 https://doi.org/10.1007/s11042-022-13056-y

27. Sarki R, Ahmed K, Wang H, Zhang Y, Ma J, Wang K (2021) Image preprocessing in classification andidentification of diabetic eye diseases. Data Sci 28. Eng 6:455–471. https://doi.org/10.1007/s41019-021-00167-z
29. Mayyaa, V, Kamath, SS, Kulkarni, U (2021) Automated microaneurysms detection for early diagnosis ofdiabetic retinopathy: A Comprehensive review, Computer Methods and Programs in Biomedicine Update,1, pp-1-15 https://doi.org/10.1016/j.cmpbup.2021.100013
30. Hattiya T, Dittakan K, Musikasuwan S (2021) Diabetic Retinopathy Detection Using Convolutional Neural Network: A Comparative Study on Different Architectures. Mahasarakham Int J Eng Technol 7(1):50–60.https://doi.org/10.14456/mijet.2021.8
31. Chetoui, M, Akhloufi, MA (2020) Explainable Diabetic Retinopathy using EfficientNET, In 2020 42ndAnnual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), pp.1966–1969, https://doi.org/10.1109/EMBC44109.2020.9175664
32. Tymchenko B, Marchenko P, Spodarets D (2020) Deep learning approach to diabetic retinopathy detection, arXiv:2003.02261v1 [cs.LG], pp 1–9. https://doi.org/10.48550/arXiv.2003.02261
33. Chaturvedi SS, Gupta K, Ninawe V, Prasad PS (2020) Automated diabetic retinopathy grading using deep convolutional neural network, arXiv:2004.06334v1 [eess.IV], pp 1–12. https://doi.org/10.48550/arXiv. 2004.06334
34. Samanta A, Saha A, Satapathy SC, Fernandes SL, Zhang Y (2020) Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset. Pattern Recogn Lett 135:293–298.https://doi.org/10.1016/j.patrec.2020.04.026
35. Sarki, R, Michalska, S, Ahmed, K, Wang, H, Zhang, Y (2019) Convolutional neural networks for milddiabetic retinopathy detection: an experimental study, bioRxiv, pp.1–18 https://doi.org/10.1101/763136
36. Ji Q, Huang J, He W, Sun Y (2019) Optimized deep convolutional neural networks for identification ofmacular diseases from optical coherence to- mography images. Algorithms 12(3):1–12. https://doi.org/10.3390/a12030051
37. Michele A, Colin V, Santika DD (2019) MobileNet convolutional neural networks and support vector Machines for Palmprint Recognition. Procedia Comput Sci 157:110–117. https://doi.org/10.1016/j.procs.2019.08.147
38. Hui J, Du M, Ye X, Qin Q, Sui J (2019) Effective building extraction from high-resolution remote sensingimages with multitask driven deep neural net- work. IEEE Geosci Remote Sens Lett 16(5):786–790. https://doi.org/10.1109/LGRS.2018.2880986
39. Jiang, H, Yang, K, Gao, M, Zhang, D, Ma, H, Qian, W (2019) An interpretable ensemble deep learningmodel for diabetic retinopathy disease classifi- cation, In 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 2045–2048, https://doi.org/10.1109/ EMBC.2019.8857160
40. Orlando JI, Prokofyeva E, del Fresno M, Blaschko MB (2018) An ensemble deep learning based approach for red lesion detection in fundus images. Comput Methods Prog Biomed 153:115–127. https://doi.org/10.1016/j.cmpb.2017.10.017
41. Suriyal, S, Druzgalski, C, Gautam, K (2018) Mobile assisted diabetic retinopathy detection using deep neural network, 2018 Global Medical Engi- neering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE), pp. 1–4, https://doi.org/10.1109/GMEPE-PAHCE.2018.8400760

42. Huang G, Liu S, Maaten L, Weinberger KQ (2018) CondenseNet: an efficient densenet using learned group convolutionss. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp 2752–2761. https://doi.org/10.1109/CVPR.2018.00291
43. Pogorelov, K, Riegler, M, Halvorsen, P , Griwodz, C, Lange, T, Randel, KR, Eskeland, SL, Dang-Nguyen,D , Ostroukhova, O, Lux, M, Spampinato, C (2017) A Comparison of Deep Learning with Global Features for Gastrointestinal Disease Detection, MediaEval'17, pp. 1–3 https://doi.org/10.1007/s11042-017-4989-y
44. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nel- son PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic Retinopathy in retinal fundus photographs.JAMA 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216
45. Nneji GU, Cai J, Deng J, Monday HK, Hossin MA, Nahar S (2022) Identification of diabetic Retinopathyusing weighted fusion deep learning based on Dual-Channel fundus scans. Diagnostics 12(540):1–19. https://doi.org/10.3390/diagnostics12020540
46. Dong, B, Wang, X, Qiang, X, Du, F, Gao, L, Wu, Q, Cao, G, Dai, C (2022) A Multi-Branch ConvolutionalNeural Network for Screening and Staging of Diabetic Retinopathy Based on Wide-Field Optical Coherence Tomography Angiography, IRBM, pp.1–7 https://doi.org/10.1016/j.irbm.2022.04.004
47. Padmanayana, Anoop BK (2022) Binary Classification of DR-Diabetic Retinopathy using CNN with Fundus Colour Images, Materials Today: Pro- ceedings, pp.1–5 https://doi.org/10.1016/j.matpr.2022.01.466
48. Sivapriya, G, Praveen, V, Gowri, P, Saranya, S, Sweetha, S, Shekar, K (2022) Segmentation of Hard exudates for the detection of Diabetic Retinopa- thy with RNN based sematic features using fundus images,Materials Today: Proceedings, pp.1–9 https://doi.org/10.1016/j.matpr.2022.05.189
49. Bora A, Balasubramanian S, Babenko B, Virmani S, Venugopalan S, Mitani A, Marinho GO, Cuadros J, Ruamviboonsuk P, Corrado GS, Peng L, Webster DR, Varadarajan AV, Hammel N, Liu Y, Bavishi P (2021) Predicting the risk of developing diabetic retinopathy using deep learning. Lan- cet Digit Health 2020(3):e10–e19. https://doi.org/10.1016/S2589-7500(20)30250-8
50. Deepa V, Kumar CS, Cherian T (2021) Ensemble of multi-stage deep convolutional neural networks forautomated grading of diabetic retinopathy using image patches, Journal of King Saud University – Computerand Information Sciences, pp 1–11 https://doi.org/10.1016/j.jksuci.2021.05.009
51. Saeed F, Hussain M, Aboalsamh HA (2021) Automatic diabetic Retinopathy diagnosis using adaptive fine-tuned convolutional neural network. IEEE Ac- cess 9:41344–41359. https://doi.org/10.1109/ACCESS.2021.3065273
52. Tsai CY, Chen CT, Chen GA, Yeh CF, Kuo CT, Hsiao YC, Hu HY, Tsai IL, Wang CH, Chen JR, HuangSC, Lu TC, Woung LC (2022) Necessity of local modification for deep learning algorithms to predict diabetic Retinopathy. Int J Environ Res Public Health 19(3):1–12. https://doi.org/10.3390/ijerph19031204
53. Atwany MZ, Sahyoun AH, Yaqub M (2022) Deep learning techniques for diabetic Retinopathy classifica-tion: a survey. IEEE Access 10:28642–28655. https://doi.org/10.1109/ACCESS.2022.3157632
54. Das S, Kharbanda K, Suchetha M, Raman R, Edwin DD (2021) Deep learning architecture based on segmented fundus image features for classifica- tion of diabetic retinopathy. Biomed Signal Process

Control68:1–10. https://doi.org/10.1016/j.bspc.2021.102600
55. Lim WX, Chen ZY, Ahmed A (2022) The adoption of deep learning interpretability techniques on diabeticretinopathy analysis: a review. Med Biol Eng Comput 60:633–642. https://doi.org/10.1007/s11517-021- 02487-8
56. AbdelMaksoud, E, Barakat, S, Elmogy, M (2022) A computer-aided diagnosis system for detecting variousdiabetic retinopathy grades based on a hybrid deep learning technique, Med Biol Eng Comput, pp.1–24https://doi.org/10.1007/s11517-022-02564-6
57. Li N, Ma M, Lai M, Gu L, Kang M, Wang Z, Jiao S, Dang K, Deng J, Ding X, Zhen Q, Zhang A, Shen T,Zheng Z, Wang Y, Peng Y (2022) A strati- fied analysis of a deep learning algorithm in the diagnosis of diabetic retinopathy in a real-world study. J Diab 14(2):111–120. https://doi.org/10.1111/1753-0407.13241
58. Deepa, V, Kumar, SC, Cherian, T (2022) Automated grading of diabetic retinopathy using CNN with hierarchical clustering of image patches by siamese network, Physical and Engineering Sciences in Medicine, pp.1–13 https://doi.org/10.1007/s13246-022-01129-z
59. Sau, PC, Bansal, A (2022) A novel diabetic retinopathy grading using modified deep neural network withsegmentation of blood vessels and retinal ab- normalities, Multimedia Tools and Applications, pp.1–29 https://doi.org/10.1007/s11042-022-13056-y
60. Shaik, NS, Cherukuri, TK (2022) Hinge attention network: A joint model for diabetic retinopathy severity grading, Applied Intelligence, pp. 1–17 https://doi.org/10.1007/s10489-021-03043-5