

Detection of Diabetic Retinopathy Using Custom CNN to Segment the Lesions

Saleh Albahli^{1,2,*} and Ghulam Nabi Ahmad Hassan Yar³

¹Department of Information Technology, College of Computer, Qassim University, Buraydah, Saudi Arabia

²Department of Computer Science, Kent State University, Kent, OH, USA

³Department of Electrical and Computer Engineering, Air University, Islamabad, Pakistan

*Corresponding Author: Saleh Albahli. Email: salbahli@qu.edu.sa

Received: 17 October 2021; Accepted: 25 November 2021

Abstract: Diabetic retinopathy is an eye deficiency that affects the retina as a result of the patient having Diabetes Mellitus caused by high sugar levels. This condition causes the blood vessels that nourish the retina to swell and become distorted and eventually become blocked. In recent times, images have played a vital role in using convolutional neural networks to automatically detect medical conditions, retinopathy takes this to another level because there is need not for just a system that could determine is a patient has retinopathy, but also a system that could tell the severity of the procession and if it would eventually lead to macular edema. In this paper, we designed three deep learning models that would detect the severity of diabetic retinopathy from images of the retina and also determine if it would lead to macular edema. Since our dataset was a small one, we employed three techniques for generating images from the ones we have, the techniques are Brightness, color and, contrast (BCC) enhancing, Color jitters (CJ), and Contrast Limited Adaptive Histogram Equalization (CLAHE). After the dataset was ready, we used it to train the ResNet50, VGG16, and VGG19 models both for determining the severity of the retinopathy and also the chances of macular edema. After validation, the models yielded very reasonable results.

Keywords: Convolutional neural networks; deep learning; diabetic retinopathy; diabetes mellitus; ResNet50; VGG16; VGG19

1 Introduction

Retinopathy is a disease of the retina. It occurs due to several reasons in a diabetic person. These reasons include hemorrhage, extrudes, and microaneurysms.

Macular edema is fluid build-up in the macula (area in the center of the retina). This causes the macula to swell and hence causes blurry and distorted vision. This fluid buildup is caused by retinopathy, when hemorrhage occurs it also causes the fats and other fluids to leak along with blood.

This is how the segmentation and disease grading task of the Indian Diabetic Retinopathy Image Dataset (IDRiD) dataset [1] given are linked. We will first find the severity of retinopathy and macular edema caused



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

by hemorrhage, extrudes, and microaneurysms. Then we will segment areas affected by hemorrhage, extrudes, and microaneurysms. Fig. 1 shows the lesions in the eye and their respective masks [1].

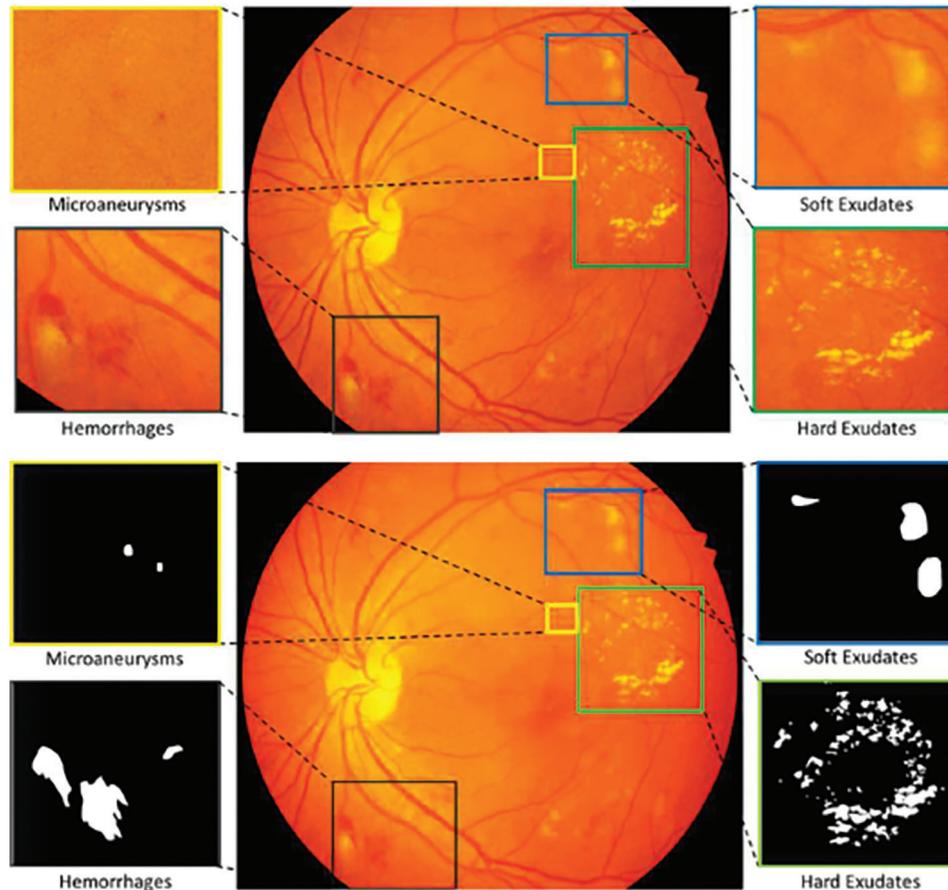


Figure 1: Different retinal lesions and their masks

Diabetic retinopathy (DR) has four stages of progression, which goes from having no DR to proliferative DR. The first stage is type 0, where there are no abnormalities observed meaning there is no DR. The second stage is type 1 called the mild non-proliferative retinopathy and it is characterized by observing microaneurysms, it is considered the earliest stage. The next stage, type 2 is moderate non-proliferative retinopathy. At this stage, the blood vessels that nourish the retina become distorted and swell thereby losing their ability to carry blood, it is characterized by observing microaneurysms, dots, hemorrhages, and spots. The fourth stage is type 3 which is the severe non-proliferative retinopathy that occurs due to blood not reaching the retina due to blockage, hence making the retina grow fresh blood vessels.

The final stage is type 4 which is the proliferative diabetic retinopathy which is an advanced stage that is characterized by either neovascularization of the disc or vitreous hemorrhage. Tab. 1 shows a summary of the stages.

The main contributions of this paper are as follows:

1. Regeneration of more datasets from the existing dataset using three different methods: Brightness, color and, contrast (BCC) enhancing, Color jitters (CJ), and Contrast Limited Adaptive Histogram Equalization (CLAHE).

2. Performing disease grading classification using three different deep learning models: ResNet50, VGG16, and VGG19.
3. Segmenting the result to determine which grade it is, from grade 0 to 1.
4. Comparing our work with similar works and suggesting ways to improve the field of retinopathy detecting in general.

Table 1: Severity levels of retinopathy

DR stage	Observation	Characteristics	Severity level
Type 0	Normal retina	No observation	No DR
Type 1	Small bulges in the tiny blood vessels of the retina.	Micro-aneurysms	Mild non-proliferative DR
Type 2	Blood vessels become distorted and swell.	Micro-aneurysms, dots, hemorrhages and spots	Moderate nonproliferative DR
Type 3	Total blockage of blood vessels. Creating new blood vessels.	Macular ischemia Retinal detachment	Severe nonproliferative DR
Type 4	Growth of abnormal and fragile new blood vessels.	Neovascularization of the disc Vitreous hemorrhage	Proliferative DR

2 Related Works and Motivation

Qummar et al. [2] proposed a system that utilized the publicly available Kaggle dataset of retina images five ensemble deep Convolutional Neural Network (CNN) models (ResNet50, InceptionV3, Xception, Dense121, Dense169) which could encode the rich features and improve the classification task for different stages of DR. After the experiment, results show that the proposed model detects all the stages of DR unlike the current methods and does a better job of classification when compared to other state-of-the-art models trained on the same Kaggle dataset. Tymchenko et al. [3] proposed a system that could automatically detect the stages of diabetic retinopathy by applying deep learning on photography of the human fundus. Additionally, they applied the multistage approach to transfer learning, which made use of a similar dataset but with different labeling. The proposed method could be used as a screening method for early detection of diabetic retinopathy after it got a sensitivity and specificity of 99% and is ranked 54 of 2943 among competing methods (quadratic weighted kappa score of 0.925466) on APTOS 2019 Blindness Detection Dataset (13000 images).

Rakhlin [4] made use of deep Convolutional Neural Networks (CNNs) to design a system that could diagnose eye fundus images. They used the publicly available Kaggle dataset to train the model while for testing, they used a portion of the dataset used for training and Messidor-2 reference standard. Neither the withheld Kaggle dataset nor the Messidor-2 was used for training. For Messidor-2 we achieved sensitivity 99%, specificity 71%, and area under the curve (AUC) 0.97. These results are close to recent state-of-the-art models trained on much larger data sets and surpass the average results of diabetic retinopathy screening when performed by trained optometrists. They assured that with the continuous development of

their models, they expect to increase the accuracy of the model and expand it to cataract and glaucoma diagnostics.

Nguyen et al. [5] presented an automated classification system that analyzed fundus images with varied illumination and fields of view to generate a severity grade for diabetic retinopathy (DR) by using Convolutional Neural Networks (CNN) such as VGG-16 and VGG-19. The proposed system got a sensitivity of 80%, an accuracy of 82%, specificity of 82%, and AUC of 0.904 for classifying images into 5 categories ranging from 0 to 4, where 0 is no DR and 4 is proliferative DR.

Li [6] proposed an optical coherence tomography (OCT) image-based diagnostic technology for automated early DR diagnosis, including at both grades 0 and 1. The work was aimed at helping ophthalmologists with the evaluation and treatment, thereby reducing the rate of vision loss by enabling timely and accurate diagnosis. They designed and evaluated a novel deep learning network—OCTD Net, for early-stage DR detection. While one of the networks extracted features from the original OCT image, the other extracted retinal layer information. The accuracy, sensitivity, and specificity were 92%, 90%, and 95%, respectively.

Hemanth et al. [7] employed the use of image processing with histogram equalization, and contrast limited adaptive histogram equalization techniques to develop a convolutional neural network that would perform the classification task. The method was validated using 400 retinal fundus images within the MESSIDOR database, and average values for different performance evaluation parameters were obtained as accuracy 97%, sensitivity (recall) 94%, specificity 98%, precision 94%, FScore 94%, and GMean 95%.

Gangwar et al. [8] address the problem of automatic diabetic retinopathy detection and proposed a novel deep learning hybrid to solve the problem. They used transfer learning on pre-trained Inception-ResNet-v2 and added a custom block of CNN layers on top of Inception-ResNet-v2 for building the hybrid model. The performance of the model was evaluated on the Messidor-1 diabetic retinopathy dataset and APTOS 2019 blindness detection (Kaggle dataset). They got an accuracy of 72.33% and 82.18% on the Messidor-1 and APTOS datasets, respectively.

Chakrabarty [9] proposed a method whose primary goal was to automatically classify patients having diabetic retinopathy and not having the same, given any High-Resolution Fundus Image of the Retina. For that, an initial image processing was done on the images which include mainly, conversion of colored (RGB) images into perfect greyscale and resizing it. Then, a Deep Learning Approach was applied in which the processed image was fed into a Convolutional Neural Network (CNN) to predict whether the patient is diabetic or not. This methodology was applied to a dataset of 30 High-Resolution Fundus Images of the retina. The results, so obtained are a 100% predictive accuracy and a Sensitivity of 100% also. Such an Automated System can easily classify images of the retina among Diabetic and Healthy patients, reducing the number of reviews of doctors.

Wan et al. [10] attempted to find an automatic way of classifying a given set of fundus images. They used convolutional neural networks (CNNs) for DR detection, which included 3 major difficult challenges: classification, segmentation, and detection. Coupled with transfer learning and hyper-parameter tuning, they adopted AlexNet, VggNet, GoogleNet, ResNet, and analyze how well these models do with the DR image classification. They used the publicly available Kaggle platform for training these models. The best classification accuracy is 95.68% and the results have demonstrated the better accuracy of CNNs and transfer learning on DR image classification.

Qureshi et al. [11] proposed an automatic recognition of the DR stage based on a new multi-layer architecture of active deep learning (ADL). To develop the ADL system, they used a CNN model to automatically extract features compare to handcrafted-based features. However, the training of the CNN procedure required an immense size of labeled data which made it almost difficult in the classification phase.

As a result, a label-efficient CNN architecture is presented known as ADL-CNN by using one of the active learning methods known as an expected gradient length (EGL). This ADL-CNN model can be seen as a two-stage process. At first, the proposed ADL-CNN system selects both the most informative patches and images by using some ground truth labels of training samples to learn the simple to complex retinal features. Next, it provided useful masks for prognostication to assist clinical specialists with the important eye sample annotation and segment regions-of-interest within the retinograph image to grade five severity levels of diabetic retinopathy. To test and evaluate the performance of the ADL-CNN model, the EyePACS benchmark is utilized and compared with state-of-the-art methods. The statistical metrics are used such as sensitivity (SE), specificity (SP), F-measure and classification accuracy (ACC) to measure the effectiveness of the ADL-CNN system. On 54,000 retinograph images, the ADL-CNN model achieved an average SE of 92.20%, SP of 95.10%, F-measure of 93%, and ACC of 98%. Hence, the new ADL-CNN architecture is outperformed for detecting DR-related lesions and recognizing the five levels of severity of DR on a wide range of fundus images.

Chetoui et al. [12] introduced the use of different texture features for DR, mainly Local Ternary Pattern (LTP) and Local Energy-based Shape Histogram (LESH). They showed that they outperform LBP extracted features. Support Vector Machines (SVM) were used for the classification of the extracted histogram. A histogram binning scheme for features representation was proposed. The experimental results showed that LESH is the best performing technique with an obtained accuracy of 90.4% using SVM with a Radial Basis Function kernel (SVM-RBF). Similarly, the analysis of the ROC curve showed that LESH with SVM-RBF gives the best AUC (Area Under Curve) performance with 0.931.

Zeng et al. [13] proposed a model that accepts binocular fundus images as inputs and learns their correlation to help to make a prediction. In the case with a training set of only 28,104 images and a test set of 7,024 images, an area under the receiver operating curve of 0.951 is obtained by the proposed binocular model, which is 0.011 higher than that obtained by the existing monocular model. To further verify the effectiveness of the binocular design, a binocular model for five-class DR detection was also trained and evaluated on a 10% validation set. The result shows that it achieved a kappa score of 0.829 which is higher than that of the existing non-ensemble model.

Dutta et al. [14] proposed a method where fundus images containing diabetic retinopathy were taken into consideration. The idea behind this paper was to propose an automated knowledge model to identify the key antecedents of DR. The proposed Model was trained with three types of techniques, backpropagation NN, Deep Neural Network (DNN), and Convolutional Neural Network (CNN) after testing models with central processing unit (CPU) trained Neural network gives the lowest accuracy because of one hidden layer whereas the deep learning models are outperforming NN. The Deep Learning models are capable of quantifying the features as blood vessels, fluid drip, exudates, hemorrhages, and microaneurysms into different classes. The model will calculate the weights which give the severity level of the patient's eye. The foremost challenge of this study was the accurate verdict of each feature class threshold. To identify the target class thresholds weighted Fuzzy C-means algorithm was used.

2.1 Research Gap

A few problems were encountered by researchers who took to solving the retinopathy detection problem, a lot of these problems are inevitable and don't have a solution right now, this is because the field of deep learning is still new to a lot of people and data gathering always brings problems. Data are scarce and most of the data available are not usable due to one reason or the other. Fig. 2 gives an insight into the research gaps of retinopathy using CNN to segment the lesions. Also, Tab. 2 compares and shows the previous approaches in the area.

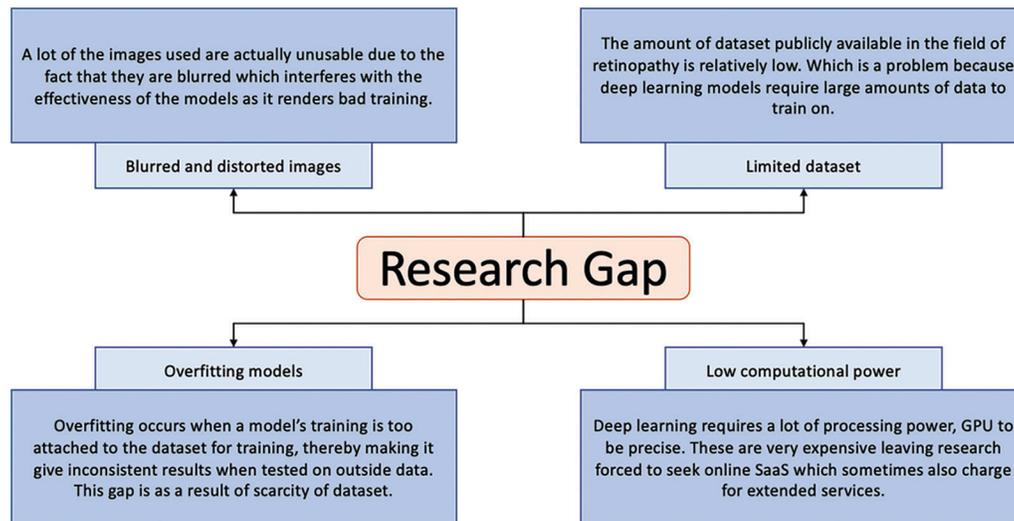


Figure 2: Study of the research gap of Retinopathy using CNN to segment the lesions

Table 2: Summary of the previous approaches, their methodologies, gaps, and improvement of retinopathy using CNN to segment the lesions

Reference	Methodology	Dataset	Findings	Gaps identified
Qummar et al. [2]	ResNet50, InceptionV3, Xception, Dense121, Dense169	Kaggle	Recall of 97%, Precision of 84%, F1-score of 90%	The model's performance can be enhanced with the right measures.
Tymchenko et al. [3]	Multistage approach to transfer learning	APTOS 2019 Blindness Detection Dataset	Sensitivity of 99%, Specificity of 99%, Quadratic weighted kappa score of 0.925466	The model performs great and can also be improved by finetuning the model.
Rakhlin [4]	Convolutional Neural Networks (CNNs).	Kaggle and Messidor-2	Sensitivity 99%, Specificity 71%, AUC 0.97	The accuracy can be improved and the complexity can be reduced
Nguyen et al. [5]	VGG-16 and VGG-19	–	Sensitivity of 80%, Accuracy of 82%, Specificity of 82%, AUC of 0.904	The accuracy can be improved by making the model deeper.
Li [6]	OCTD Net	–	Accuracy of 92%, Sensitivity of 90%, Specificity of 95%	Fine-tuning the parameters can improve the model.

(Continued)

Table 2 (continued)				
Reference	Methodology	Dataset	Findings	Gaps identified
Hemanth et al. [7]	Convolutional neural Network	MESSIDOR database	Accuracy of 97%, Sensitivity (recall) 94%, Specificity of 98%, Precision of 94%, FScore of 94%, GMean of 95%.	The prediction accuracy can further be enhanced
Gangwar et al. [8]	Transfer learning on pre-trained Inception-ResNet-v2 with added custom block of CNN layers.	Messidor-1 diabetic retinopathy dataset and APTOS 2019 blindness detection (Kaggle dataset).	Accuracy of 72.33% on, Messidor-1, Accuracy of 82.18% on APTOS dataset	The accuracy of the model is low and can be enhanced with modifications to the model.
Chakrabarty [9]	Convolutional Neural Network	Dataset of 30 High Resolution Fundus Images of the retina	Predictive accuracy of 100%, Sensitivity of 100%	No model can get 100%, hence the model has overfitted.
Wan et al. [10]	AlexNet, VggNet, GoogleNet and ResNet	Kaggle	Accuracy of 95.68%	The prediction accuracy can be improved.
Qureshi et al. [11]	ADL-CNN	Dataset with 54,000 Retinography images	Average SE of 92.20%, SP of 95.10%, F-measure of 93%, Accuracy of 98%	The model does well but can be improved by fine-tuning the model.
Chetoui et al. [12]	Local Energy based Shape Histogram (LESH) using SVM with a Radial Basis Function kernel (SVM-RBF)	–	AUC (Area Under Curve) of 0.931	The model can be enhanced to improve performance.
Zeng et al. [13]	CNN	Dataset with 28,104 images for training and 7,024 images for testing	The area under the receiver operating curve of 0.951, Kappa score of 0.829	Making the model deeper can increase the performance of the model.
Dutta et al. [14]	Backpropagation BNN, Deep Neural Network (DNN) and Convolutional Neural Network (CNN)	Kaggle	BNN accuracy: 42%, DNN accuracy: 86.3%, CNN: 78.3%	The accuracies obtained are relatively low. Modifying the model will improve the results.

From the previous review of the current literature utilized conventional methods and DL architectures, we can conclude their main limitations are as follows:

1. Limited dataset
2. Blurred and distorted images
3. Overfitting models
4. Low computational power

2.2 Problem Formulation and Suggested Approach

Retinopathy is one of the diseases which has a scarcity of specialists, there is a need now more than ever for a system that could automatically detect and identify patients who suffer from it. Our work began at the data acquisition stage where we got our images from the Indian Diabetic Retinopathy Image Dataset (IDRID) [1]. The data preprocessing stage came second, preprocessing was very crucial because even though the dataset was standard, we needed to tailor it to fit the problem we were trying to solve. Then finally, the data is entered into the classification models where it is trained to automatically detect retinopathy. Fig. 3 shows the pipeline of how the experiment of CNN is used to segment the lesions.

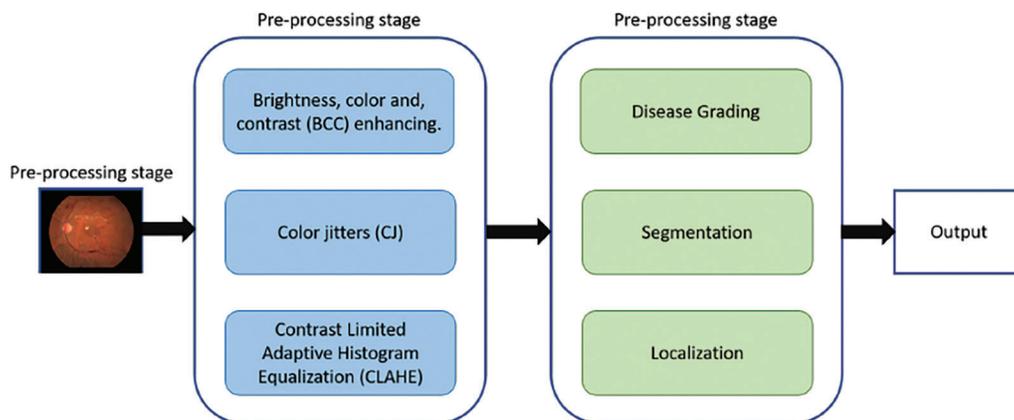


Figure 3: Pipeline for the CNN task to segment the lesions

3 Methodology

3.1 Dataset

Indian Diabetic Retinopathy Image Dataset (IDRID) [1] consists of three different tasks:

1. Segmentation: This includes segmentation of four causes of retinopathy that includes, Microaneurysms (MA), Haemorrhages (HE), Hard Exudates (EX), and Soft Exudates (SE). It also includes images for the segmentation of the optic disc (OD).
2. Disease Grading: This includes detection that how much the eye is affected by retinopathy and how much are the chances of macular edema due to retinopathy.
3. Localization: This includes localization of optic disc and Fovea center

In this paper, only segmentation and disease grading are targeted because they are linked with each other. Lesions given in the segmentation task are the cause of retinopathy and it also determines that what are the chances of macular edema.

In the segmentation task, there is a total of 81 images that are then divided into training (54 images) and test set (27 images). Masks for all 81 images are given and for each lesion, a different mask is given. Although some images do not have a specific lesion or that image the mask of that specific lesion is not given.

In the disease grading task, there is a total of 516 images that are further divided into training (413 images) and testing set (103 images). For retinopathy, 5 grades are ranging from 0 to 4, and for macular edema, 3 grades are ranging from 0 to 3.

3.2 Pre-Processing

Before performing any task, images were pre-processed and three more datasets were generated. The first dataset was generated by enhancing the brightness, color and, contrast (BCC) of the original images. The second dataset was obtained by applying color jitters (CJ) to the original image.

In color jitters brightness, contrast and, a saturation of images are changed randomly. The third dataset was obtained by applying Contrast Limited Adaptive Histogram Equalization (CLAHE) on original images.

3.3 Disease Grading Classification

At first, the disease grading is targeted using different classification models. Classification models used in this study are ResNet50, VGG-16, and VGG19. Three datasets generated during pre-processing along with original images were used to train the classification models. These models were pre-trained on the ImageNet dataset having 1000 classes and were finetuned by adding more layers for our disease grading task. The advantage of using a pre-trained model is to mitigate the effect of fewer datasets. [Tab. 3](#) shows the additional fine-tuning layers added to the original architecture of the models.

Table 3: Fine-tuning layers for X-ray and CT-Scan classifier

X-ray classification			CT classification		
Type	Output	Kernel	Type	Output	Kernel
Average Pooling	2048	2×2	Average Pooling	2048	2×2
Flatten	8192	–	Flatten	8192	–
Dense	1024	–	Dense	1024	–
Dropout(0.5)	1024	–	Dropout(0.5)	1024	–
Dense	1024	–	Dense	1024	–
Dropout (0.5)	1024	–	Dropout (0.5)	1024	–
Dense	3	–	Dense	2	–

ResNet-50 is a model given by Microsoft [15] and has the quality feed output of some layer directly to the input of some other layer by bypassing layers in between. This property of ResNet-50 is called identity mapping. This helps with solving the problem of vanishing gradient. [Fig. 4](#) shows the working of identity mapping.

VGG-16 and VGG-19 are the models given by “Visual Graphic Group” at Oxford University [16]. VGG-16 and VGG-19 contain 16 and 19 layers respectively arranged in a pyramid shape. [Fig. 5](#) shows the pyramid architecture of VGG-16 and [Fig. 6](#) shows the architecture for VGG-19.

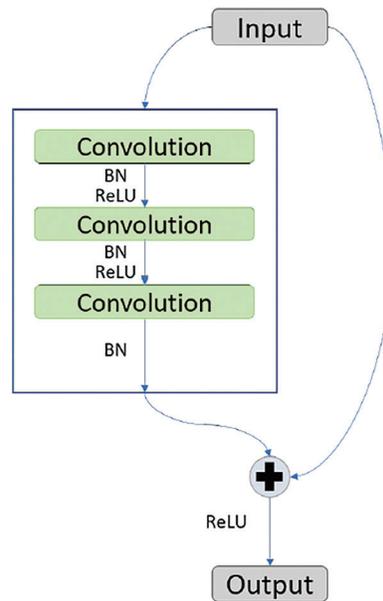


Figure 4: Identity mapping block of ResNet50

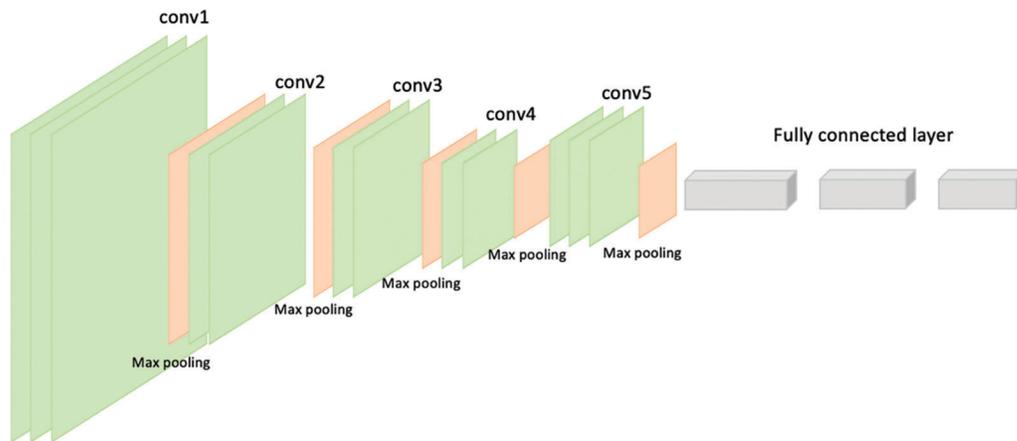


Figure 5: Architecture of VGG-16

3.4 Segmentation

After finding the severity of the disease, different causes of retinopathy were segmented. For segmentation UNet was used. UNet consists of two parts, contraction and expansion.

The contraction part extracts features from the image using Downsampling and max-pooling layers. While the expansion part localizes the segmentation area. [Tab. 4](#) shows the UNet architecture used.

4 Results

At first, images were pre-processed using different techniques. [Fig. 7](#) shows the results of random images after applying different pre-processing techniques to the original image. It can be seen from the images that features inside the image has been highlighted, for example, bright yellow spots in the processed images are the exudates. These spots were not as bright in the original image.

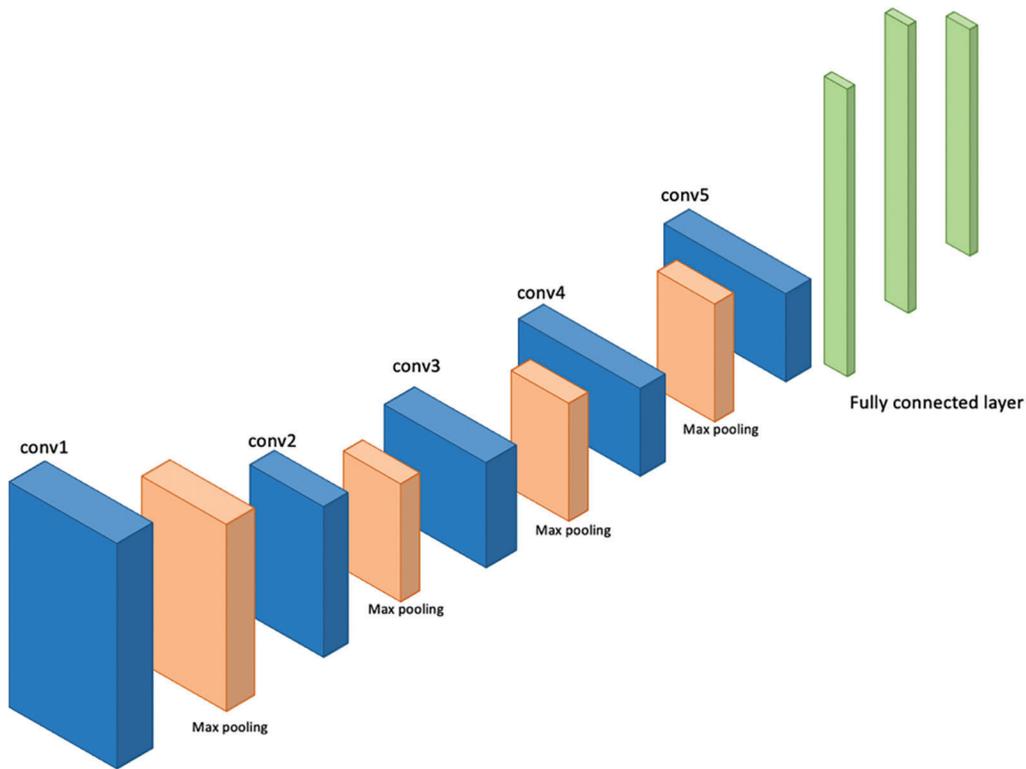


Figure 6: Architecture of VGG-19

Table 4: Architecture of UNet

Layer	Type	Kernel size	Filters
1	Input Layer	N/A	N/A
Contraction Path			
2	Convolution	3×3	16
3	Dropout (0.1)	N/A	N/A
4	Convolution	3×3	16
5	MaxPooling	2×2	1
6	Convolution	3×3	32
7	Dropout (0.1)	N/A	N/A
8	Convolution	3×3	32
9	MaxPooling	2×2	1
10	Convolution	3×3	64
11	Dropout (0.2)	N/A	N/A
12	Convolution	3×3	64
13	MaxPooling	2×2	1
14	Convolution	3×3	128

(Continued)

Table 4 (continued)			
Layer	Type	Kernel size	Filters
15	Dropout (0.2)	N/A	N/A
16	Convolution	3×3	128
17	MaxPooling	2×2	1
18	Convolution	3×3	256
19	Dropout (0.3)	N/A	N/A
20	Convolution	3×3	256
Expansive Path			
21	Conv Transpose	2×2	128
22	Concatenate(21,16)	N/A	N/A
23	Convolution	3×3	128
24	Dropout (0.2)	N/A	N/A
25	Convolution	3×3	128
26	Conv Transpose	2×2	64
27	Concatenate(26,12)	N/A	N/A
28	Convolution	3×3	64
29	Dropout (0.2)	N/A	N/A
30	Convolution	3×3	64
31	Conv Transpose	2×2	32
32	Concatenate(31, 8)	N/A	N/A
33	Convolution	3×3	32
34	Dropout (0.1)	N/A	N/A
35	Convolution	3×3	32
36	Conv Transpose	2×2	16
37	Concatenate(36, 4)	N/A	N/A
38	Convolution	3×3	16
39	Dropout (0.1)	N/A	N/A
40	Convolution	3×3	16
41	Conv (Sigmoid)	1×1	1

After generating the new datasets through pre-processing, classification models were trained on those datasets separately. Tab. 5 shows the training, testing, and validation accuracy along with losses for finding the severity of retinopathy. In the case of finding the severity of retinopathy VGG-16 gave the best testing accuracy for the dataset obtained by adjusting brightness, color, and contrast. While for validation accuracy, ResNet-50 gave the best accuracy for the dataset enhanced by using CLAHE.

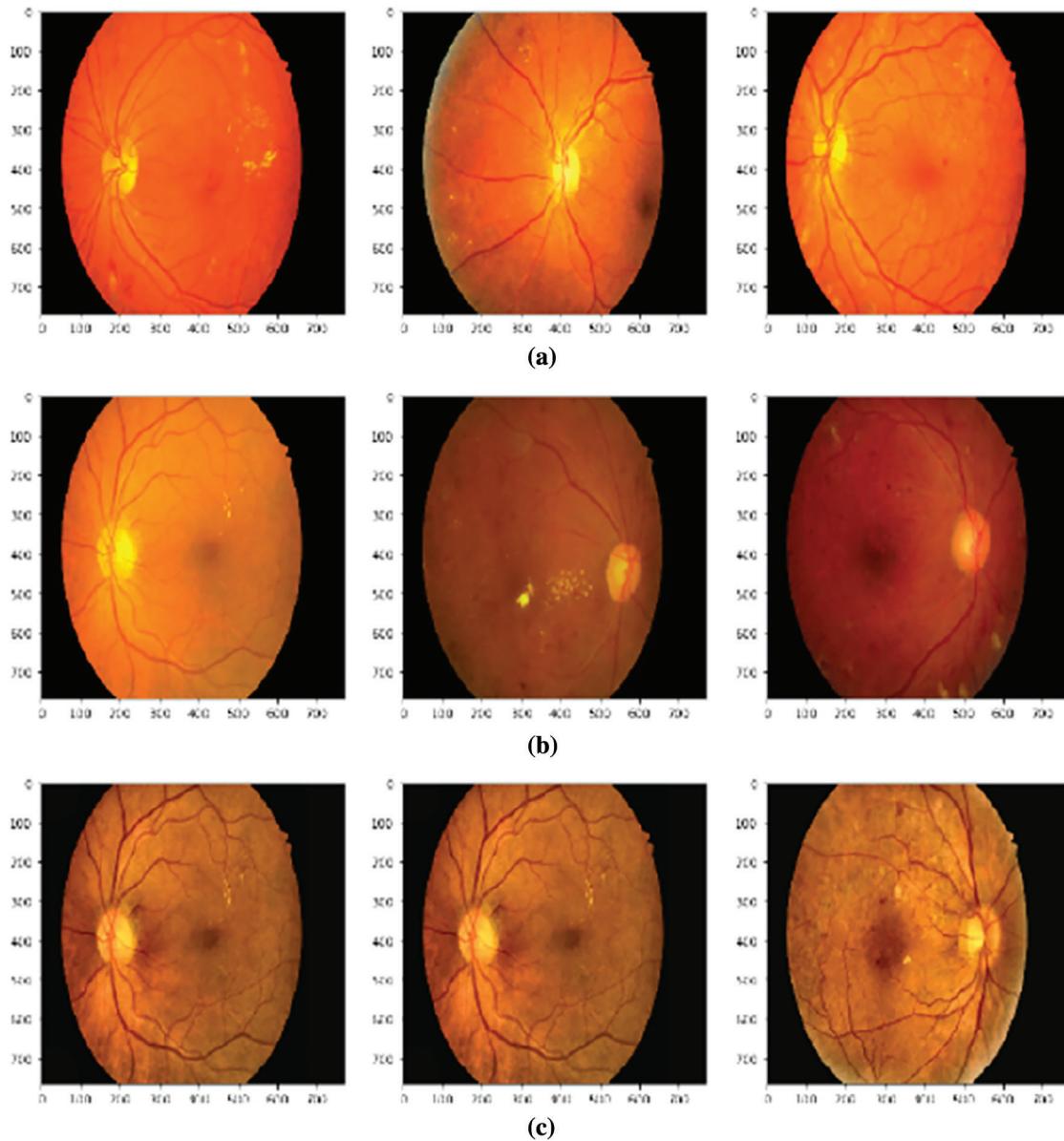


Figure 7: Effect of different pre-processing techniques (a) Images after adjusting brightness, color and contrast, (b) Images after applying color jitters (c) Images after applying CLAHE

In case of finding the chances of having macular edema, [Tab. 6](#) shows that ResNet-50 and VGG-19 gave the best testing accuracy for original images and for datasets obtained by enhancing brightness, color, and contrast respectively. Best validation accuracy was obtained *via* ResNet-50 on the original images dataset. Results for finding the chances of having macular edema can be seen in [Tab. 7](#).

After the severity is found, segmentation of different retinopathic lesions was done using UNet. UNet was trained for all four types of datasets and each kind of lesion segmentation.

[Tab. 8](#) shows the accuracies and losses obtained for each dataset and each kind of lesion. Best accuracy results are written in bold.

Table 5: Training, validation, and test accuracy achieved by different models for finding severity of retinopathy

Model	Dataset	Training accuracy	Training loss	Testing accuracy	Testing loss	Val accuracy	Val loss
ResNet50	Original	62.88%	0.9192	54.37%	1.1422	54.37%	1.0853
VGG19		62.97%	0.9317	50.49%	1.2358	52.43%	1.2329
VGG16		63.27%	0.8778	46.60%	1.2968	49.51%	1.3221
ResNet50	BCC	58.24%	0.9831	57.28%	1.1164	60.19%	1.2024
VGG19		61.51%	0.9603	58.25%	1.1326	54.37%	1.2006
VGG16		60.00%	0.9849	53.40%	1.2882	59.22%	1.3099
ResNet50	CJ	55.77%	1.0213	54.37%	1.2391	55.34%	1.2355
VGG19		59.80%	0.9659	47.57%	1.3498	48.54%	1.3105
VGG16		60.30%	0.9711	51.46%	1.3667	47.57%	1.3143
ResNet50	CLAHE	61.27%	0.9401	53.40%	1.0898	59.22%	1.1169
VGG19		61.82%	0.9201	54.37%	1.2877	49.51%	1.1585
VGG16		62.98%	0.9597	50.49%	1.1537	44.66%	1.1909

Table 6: Training, validation, and test accuracy achieved by different models for finding the chances of macular edema

Model	Dataset	Training accuracy	Training loss	Testing accuracy	Testing loss	Val accuracy	Val loss
ResNet50	Original	82.32%	0.4518	78.64%	0.6521	82.52%	0.5908
VGG19		82.92%	0.4817	68.93%	0.7793	68.93%	0.7853
VGG16		82.87%	0.4592	74.76%	0.7318	75.73%	0.668
ResNet50	BCC	81.31%	0.4914	77.67%	0.7163	76.70%	0.6914
VGG19		82.22%	0.4916	78.64%	0.771	77.67%	0.7359
VGG16		80.96%	0.5176	73.79%	0.8255	71.84%	0.8409
ResNet50	CJ	78.69%	0.5488	73.79%	0.6686	72.82%	0.7296
VGG19		80.81%	0.4917	72.82%	0.9105	73.79%	0.8573
VGG16		78.07%	0.5738	72.82%	0.8649	68.93%	0.8408
ResNet50	CLAHE	79.78%	0.5217	76.70%	0.826	76.70%	0.8452
VGG19		79.45%	0.5237	72.82%	0.7351	75.73%	0.7019
VGG16		79.55%	0.5473	73.79%	0.7677	76.70%	0.7762

Table 7: Related works of having macular edema

Network architecture performance	Network architecture performance
Rakhlin [4]	Specificity 71%
Nguyen et al. [5]	Accuracy of 82%.
Gangwar et al. [8]	Accuracy of 72.33% on Messidor-1, Accuracy of 82.18% on APTOS dataset

Table 8: Training, and test accuracy and loss achieved by different UNet for different datasets for different lesions and optical disc segmentation

Dataset	Class	Training accuracy	Training loss	Test accuracy	Test loss
Original	EX	78.12%	0.1749	63.61%	0.26
	HE	72.44%	0.2528	57.65%	0.3353
	MA	94.58%	0.0457	89.43%	0.0845
	OD	86.64%	0.1668	80.93%	0.2158
	SE	93.75%	0.089	89.15%	0.1409
BCC	EX	80.24%	0.1611	65.22%	0.2402
	HE	73.55%	0.2325	58.94%	0.3115
	MA	94.23%	0.0453	91.09%	0.0775
	OD	87.05%	0.1616	81.43%	0.2085
	SE	94.07%	0.0873	88.75%	0.1474
CJ	EX	78.59%	0.1713	64.25%	0.2567
	HE	72.06%	0.2567	56.96%	0.3399
	MA	94.08%	0.0464	88.52%	0.0948
	OD	89.12%	0.146	84.83%	0.1866
	SE	94.54%	0.0848	88.80%	0.1422
CLAHE	EX	79.08%	0.1707	64.51%	0.2527
	HE	74.12%	0.2297	59.35%	0.3004
	MA	94.86%	0.0436	89.98%	0.0818
	OD	87.24%	0.1601	80.72%	0.2174
	SE	93.90%	0.0885	89.69%	0.1337

5 Comparison with Other Works

A lot of work has been done towards using artificial intelligence to detect retinopathy and in this section, we present an analysis of recent works that have been done towards this. [Tab. 7](#) shows a summary of the results obtained from the works done by Rakhlin [4], Nguyen et al. [5], and Gangwar et al. [8] which are all models developed to detect diabetic retinopathy from images. The models do well but not very well, this is mainly due to not properly fine-tuning the parameters and hyper-parameters of the models. Our model achieves an accuracy of 83% which is fairly high even though it will be improved as more and more dataset is collected.

6 Conclusion and Future Work

Retinopathy is the main cause of eye defects in diabetic patients. To find the severity of retinopathy and the chances of the eye being damaged due to macular edema are important. Not only that finding which area of the eye is affected by what kind of lesion can help us know the main cause of retinopathy. This paper trained different pre-trained models to find the severity of retinopathy and chances of macular edema. Once the severity is found, UNet was used to segment out the areas affected by different kinds of lesions.

We found that ResNet-50 is the best model for finding the severity of retinopathy as it gave the maximum accuracy for images enhanced by adjusting brightness, color, and contrast.

For finding the chances of the eye being affected by macular edema ResNet-50 again showed maximum testing and validation accuracy but this time for original images. In the case of segmentation, the accuracies were not very far away from each other. For EX and MA, the BCC dataset showed the maximum testing accuracy. For OD, the CJ dataset showed maximum testing accuracy. In the end, the dataset enhanced through CLAHE showed the maximum testing accuracy for HE and SE.

In the future, other pre-trained models can be tested to find if other models can enhance the results further. Moreover finetuning of layers can be further enhanced for classification purposes. Also instead of using an Artificial Neural Network for classification, other models like support vector machines (SVM), random forest (RF), and decision tree (DT) can be used. In terms of segmentation, many other models can be tested i.e., UNet++.

Acknowledgement: We would like to thank the Deanship of Scientific Research, Qassim University for covering the publication of this project.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh *et al.*, “Indian diabetic retinopathy image dataset (IDRID): A database for diabetic retinopathy screening research,” *Data*, vol. 3, no. 3, pp. 25, 2018.
- [2] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband *et al.*, “A deep learning ensemble approach for diabetic retinopathy detection,” *IEEE Access*, vol. 7, pp. 150530–150539, 2019.
- [3] B. Tymchenko, P. Marchenko and D. Spodarets, “Deep learning approach to diabetic retinopathy detection,” arXiv preprint arXiv:2003.02261, 2020.
- [4] A. Rakhlin, “Diabetic retinopathy detection through integration of deep learning classification framework,” *bioRxiv*. p. 225508, 2018.
- [5] Q. H. Nguyen, R. Muthuraman, L. Singh, G. Sen, A. C. Tran *et al.*, “Diabetic retinopathy detection using deep learning,” in *Proc. of the 4th Int. Conf. on Machine Learning and Soft Computing*, Haiphong City, Viet Nam, pp. 103–107, 2020.
- [6] X. Li, L. Shen, M. Shen, F. Tan and C. S. Qiu, “Deep learning based early stage diabetic retinopathy detection using optical coherence tomography,” *Neurocomputing*, vol. 369, no. 5, pp. 134–144, 2019.
- [7] D. J. Hemanth, O. Deperlioglu and U. Kose, “An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network,” *Neural Computing and Applications*, vol. 32, no. 3, pp. 707–721, 2020.
- [8] A. K. Gangwar and V. Ravi, “Diabetic retinopathy detection using transfer learning and deep learning,” in *Evolution in Computational Intelligence*, NIT Surathkal, India, Springer, pp. 679–689, 2021.
- [9] N. Chakrabarty, “A deep learning method for the detection of diabetic retinopathy,” in *5th IEEE Uttar Pradesh Section Int. Conf. on Electrical, Electronics and Computer Engineering (UPCON)*, Gorakhpur, India, IEEE, pp. 1–5, 2018.
- [10] S. Wan, Y. Liang and Y. Zhang, “Deep convolutional neural networks for diabetic retinopathy detection by image classification,” *Computers & Electrical Engineering*, vol. 72, no. 10, pp. 274–282, 2018.
- [11] I. Qureshi, J. Ma and Q. Abbas, “Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning,” *Multimedia Tools and Applications*, vol. 80, no. 8, pp. 691–721, 2021.

- [12] M. Chetoui and M. A. Akhloufi, "Explainable end-to-end deep learning for diabetic retinopathy detection across multiple datasets," *Journal of Medical Imaging*, vol. 7, no. 4, pp. 044503, 2020.
- [13] X. Zeng, H. Chen, Y. Luo and W. Ye, "Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network," *IEEE Access*, vol. 7, pp. 744–753, 2019.
- [14] S. Dutta, B. Manideep, S. M. Basha, R. D. Caytiles and N. Iyengar, "Classification of diabetic retinopathy images by using deep learning models," *International Journal of Grid and Distributed Computing*, vol. 11, no. 1, pp. 89–106, 2018.
- [15] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," in *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 770–778, 2019.
- [16] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556, 2017.