Journal of Theoretical and Applied Information Technology
15th August 2023. Vol.101. No 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPLEMENTING RESNET-50 TRANSFER LEARNING MODEL FOR DIAGNOSING OCT IMAGES FOR DETECTING AND CLASSIFYING DME ABNORMALITIES

[1] A. GEETHA DEVI, [2] SURYA PRASADA RAO BORRA,[3]D. HARI KRISHNA, [4]M. RAJ KUMAR NAIK, [5]K. VIDYA SAGAR, [6]LAKSHMI RAMANI BURRA

[1,2] Associate Professor, Department of ECE, PVP Siddhartha Institute of Technology, Vijayawada, India
[3]Associate Professor, Department of ECE, B V Raju Institute of Technology, Narasapur, Telangana
[4]Assistant Professor, Department of ECE, Chaitanya Bharathi Institute of Technology, Gandipet, Telangana
[5]Department of EIE, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad.
[6]Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram AP, India

Email: geetha.agd@gmail.com, suryaborra1679@gmail.com, harikrishna.dodde@bvrit.ac.in, raj.meghavath@gmail.com, vidyasagar_k@vnrvjiet.ac.in, ramanimythili@gmail.com

ABSTRACT

One of the major diseases that affect human eyes due to complications of diabetes is Diabetic Retinopathy (DR). A particular type of DR affecting the retina's central portion, called the macula, creates a vision problem. It is called Diabetic Macular Edema (DME). The blood vessels in the eye get damaged, and leaking fluid on the macula causes tissue thickening and swelling. The image may be occluded for various external and environmental reasons, which can degrade the image quality and provide a wrong diagnosis. The early diagnosis of DME is essential to avoid vision loss, and OCT images are used for prescreening because OCT is one of the non-invasive imaging modalities that can provide high-resolution retina images speedily. Several earlier research works have focused on analyzing various DR images using image processing methods, providing less prediction accuracy. This work aims to create an automatic transfer learning model for classifying DME using OCT images. ResNet-50 is created by training with 80% of the training OCT images, evaluated and validated by random, and 20% of the testing images. The reason for using the **ResNet-50** model is that it is a pre-trained model using ImageNet data. The proposed ResNet is experimented with a benchmark dataset with Python for testing and validating its output. The output is compared with the other earlier methods to evaluate the performance. The comparison shows that the proposed ResNet-50 model outperforms others by obtaining 97.56% accuracy.

Keywords: *Diabetic Macula Edema, OCT Images, Transfer Learning Models, ResNet-50, Diabetic Retinopathy, Medical Image Processing.*

1. INTRODUCTION

One of the complications of diabetes that affects the eyes is Diabetic Retinopathy (DR). It occurs when high blood sugar levels damage the blood vessels in the retina, the part of the eye that senses light and sends signals to the brain. Over time, these damaged blood vessels can leak blood and other fluids, causing swelling and clouding of the retina. It can lead to blurred vision, floaters, and even complete vision loss. DR is a severe condition that can develop without symptoms, so people with diabetes need regular eye exams to detect the symptoms of early diseases. Tight control of blood sugar, blood pressure, and cholesterol levels can help to reduce the risk of developing DR. Diabetic macular edema (DME) is a specific type of DR that affects the macula, the central part of the retina responsible for sharp, detailed vision. Age factors of DR affect the macula immediately in the retina.DME can be a severe complication of diabetes that can lead to permanent vision loss if left untreated. In DME, the blood vessels in the retina become damaged and leak fluid into the macula, causing swelling

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and thickening of the tissue. The symptoms of DME can include blurred or distorted vision, difficulty seeing colors or fine details, and dark or empty areas in central vision. Various imaging modalities, like Fundus images, OCT, and OCTA, are used for diagnosing retinal diseases. This paper uses OCT images for diagnosing DME diseases.

OCT (optical coherence tomography) is a non-invasive imaging technique that uses light waves to create detailed cross-sectional images of the retina, the layer of cells at the back of the eye that detects light and sends visual signals to the brain. These images can help eye doctors diagnose and monitor various eye conditions, including diabetic retinopathy, macular degeneration, and glaucoma. OCT images are typically displayed in grayscale and show the different layers of the retina and any abnormalities or damage that may be present. The images are obtained using a specialized machine that scans the eye with a beam of light and measures the time it takes for the light to bounce back, creating a 3D image of the retina. OCT imaging is a quick and painless procedure that can be performed in a doctor's office. The images obtained from OCT can help doctors monitor the progression of eye diseases, determine the effectiveness of treatments, and make more accurate diagnoses. Factors used for identifying the symptoms of DME while analyzing the OCT images are increased retinal thickness, cystoid spaces, disruption of the inner retinal layers, and loss of foveal depression. By analyzing these and other features on OCT images, ophthalmologists can manually diagnose DME and determine the appropriate course of treatment. To improve the accuracy this paper contributes,

- A fine-tunned ResNet-50 model is presented for classifying DME diseases in OCT images.
- The computational implementation process of the ResNet-50 model is illustrated from scratch theoretically with mathematical expressions.
- The entire process and the effects of parameter optimization for predicting the required output towards improving the accuracy are explained.
- The experimental results obtained and compared with other deep learning models are explained to understand the bitterness of the ResNet-50 model.

2. LITERATURE REVIEW

This section reviews various research works on processing OCT images and DME classification. Most researchers have suggested deep learning-based detection and classification techniques to get a final output. For example, the authors O. Musat et al. (2015) have elaborately [1] defined DME's signs, symptoms, and challenges. The author has discussed treatment methods for retinal diseases. The conditions of the diseases are observed by analyzing the OCT images, and appropriate treatments are provided. P. Romero-Aroca et al. [2] (2016) present the review to demonstrate the efficiency of the current and existing DME treatment techniques. M.A. Singer et al.[3] (2016) discussed current and existing treatment techniques and therapies, such as laser therapy, VEGF, and steroid agent to treat DME disease. The author has also discussed various new treatment techniques to find a path to provide burdenless treatment for the earlier stage of DME diseases. F. Bandello et al.[4] (2017) defined DME as the most common reason for vision loss among diabetic patients. Various existing and current diagnosing tools, therapies, and techniques are also discussed. The treatment techniques are suggested based on the characteristics of DME. The author especially presents this study to represent the morphofunctional features and clinical aspects of DME and provide a guideline for future research to handle complex conditions. G.C. Chan et al. (2017) have proposed a deep learning approach, [5] CNN, to detect and classify the OCT images to predict various diseases of DME. The input images are preprocessed and filtered using the BM3D filter and cropped using the image boundaries. And using the CNN algorithm, the input features are extracted and classified using the SVM classifier. The final result of this approach is analyzed using the 8-fold cross-validation. The result of the experiment shows that the proposed approach performed better than the existing approach with 96% accuracy.

R.M. Kamble et al. (2018) [6] have proposed a pre-trained CNN algorithm to detect DME from the given input OCT images. For this experiment, two sets of datasets are collected and evaluated. At first, the proposed model

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

experimented with publicly available datasets elucidated from the Singapore eye research center. Secondly, the proposed model has experimented with another set of datasets collected from a Chinese university Hongkong. The classification result of both experiments is produced with 100% accuracy. The proposed model accurately detects the normal and DME-infected retinal from the OCT images. A. Vahadane et al. (2018) [7] developed a two-step framework to detect DME from the OCT image. Image processing is the first step, which is performed to improve the input image quality using preprocessing techniques. The second step is performed using the CNN model to detect the label for the input data. Finally, both these steps are integrated and detect the presence of DME from the input retinal images. F. Li et al. [8] (2019) introduced a deep transfer learning method based on the VGG-16 framework for detecting various retinal diseases using OCT images. Nearly 207,130 retinal image samples were gathered from 2013 to 2017 and experimented with. Various performance metric values such as sensitivity, specificity, accuracy, and ROC value of the proposed model are evaluated in the experiment. The experiment results show that the proposed model has been predicted with 98.6%, 97.8%, 99.4%, and 100% accuracy, sensitivity, specificity, and ROC curve, respectively.

F.Li et al. (2019) proposed a deep learning-based detection technique to predict normal and DME-infected retinas from the input OCT datasets [9]. The experiment is conducted with 10-fold cross-validation; the results show that the proposed model achieved 0.973 accuracy, 0.963 sensitivity, and 0.985 specificity values. This paper also emphasizes the performance of the various detection technique using OCT images. M. Chetoui and M.A. Akhloufi (2020) suggested the CNN model [10] to automatically detect various retinal diseases such as CNV, DME, and drusen from the OCT images. The result of the proposed model demonstrates that the CNN model classifies with 98.46% and 0.989 accuracies and AUC values, respectively. T. Nazir et al. (2020) suggest this DL-based retinal detection technique.

The authors Q. Wu et al. (2021) and F.Tang et al. (2021) [11], [12] proposed deep learning-based detection techniques to detect DME diseases from OCT images. Both research works indicate that the DL-based technique effectively classifies the normal and DME-infected retinal images from OCT with more than 95% accuracy. G. Altan (2022)[13]-[16] proposed a deep CNN algorithm to detect DME diseases from the input OCT images. Using the BM3D filter, the input image is filtered, and the quality of the image is improved. Compared to the existing approaches, the proposed CNN-based model effectively detects the diseases with 99.20% accuracy, 100% sensitivity, and 98.40% specificity value. P.K. Upadhyay et al. (2022) proposed the CNN model to detect early-stage DME diseases and overcome existing approaches' challenges. The proposed approach [17] has experimented with various OCT images to detect retinal diseases. The proposed approach's result indicates that, compared to other methods, the proposed CNN model predicted retinal diseases with 97.19% accuracy.

Ji, Q., Huang (2019). proposed a model [20] called Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images. In this paper the CNN Model was used to describe the features of DME features. Chen YM, [21] is a technique for Classification of age-related macular degeneration using convolutional-neural- network-based transfer learning. Y, Shi G. [22] presents deep learning for quality assessment of retinal OCT images. Z., Zhang, X. D (2019) give a detail discussion about Deep learning-based automated detection of retinal diseases using optical coherence tomography images [23]. Nugroho, K. A. [24] (2018). Provide a comparison of handcrafted and deep neural network feature extraction for classifying optical coherence tomography (OCT) images. Sun, Y., Zhang, H., & Yao, X. (2020) develop a model [25] called automatic diagnosis of macular diseases from OCT volume based on its two-dimensional feature map and convolutional neural network with attention mechanism. paper also emphasizes the performance of the various detection technique using OCT images.

3. LIMITATION AND MOTIVATION

The ResNet50 is a widely used transfer learning model for classifying medical images. The ResNet50 architecture is widely used in computer vision. One main problem in the CNN algorithms or any other deep neural network is the vanishing gradient problem, in which the resultant output decreases in value with training which is avoided in ResNet50 by skipping the blocks. The success of any deep learning model relies heavily on the dataset's quality and size. In this case, the availability of a limited dataset can be a significant limitation, which can impact the model's performance. Although transfer learning can effectively address the limitations of a limited dataset, it may not always lead to optimal results. The pre-trained model may not be well-suited for the specific task, and fine-tuning may require extensive trial-and-error experimentation. The accuracy of diagnosing OCT images for detecting and classifying DME abnormalities depends on the quality of the human annotations. There is always a risk of human error in labeling images that could affect the model's performance. Diagnosing DME abnormalities through OCT images is a critical task in ophthalmology. Accurate diagnosis can help in the early detection and prevention of blindness, significantly improving patients' quality of life. Deep learning models have tremendously succeeded in various computer vision tasks, including medical image analysis. This approach can potentially improve the accuracy and efficiency of diagnosing DME abnormalities. Transfer learning can be a practical approach to address the limitations of a limited dataset. Using a pre-trained model can help improve the model's performance by leveraging the knowledge learned from a large dataset. ResNet-50 is a well-known deep learning model that has achieved state-of-the-art performance in various computer vision tasks. Its use in this study could improve accuracy in diagnosing DME abnormalities.

4. EXISTING METHOD

This paper proposes a method for automatically detecting DME in OCT images using transfer learning with convolutional neural networks (CNNs). The proposed method utilizes a pre-trained CNN model to extract relevant features from OCT images, followed by fine-tuning and training the model with the model dataset containing the positive and negative cases of Diabetic Macular Edema. The results show that the proposed method achieves high accuracy in DME detection through OCT images. The proposed method has the potential to be used as a screening tool for DME in areas where access to retinal testing is limited or for quick initial assessments of patients suspected of DME.

5. PROPOSED METHOD

The proposed method for detecting DME involves various steps starting from preprocessing, finding the region of interest, initial segmentation, and ResNet50, followed by a CNN algorithm for the accurate prediction of DME. The preprocessing of the images involves noise removal and sharpening of edges in the images. It is done through the Gaussian filter that removes the unwanted. Then, the region of interest in the images is selected. Only the required area is selected in this selection process, and the other areas are removed in the region selection process. It should also be noted that the processing load is reduced by selecting the region of interest. The area affected by DME is detected in the initial segmentation process, carried out through the ResNet50 model. ResNet-50 has its own architecture uses a bottleneck design for reducing the number of parameters and the multiplication process on matrices. The classification of the extracted features is carried out through the ReLU layer and SoftMax function. With this, the DME is predicted, and the prediction accuracy is compared with similar methods. The overall workflow of the proposed model is depicted in Figure-1.

Figure-1. Proposed Work Flow

6. PREPROCESSING

Before segmenting the images, it needs to be preprocessed for better classification and segmentation accuracy. In this process, the edges of the images are sharpened, and the noise in the images is removed. The bilateral filter is widely used in this process, while the non-linear smoothing filter safeguards the images' edges and blurs unwanted regions. During this smoothening process, a weighted average value is replaced in the abnormal regions. Various values, like the images' photometric similarity and the membrane's geometrical proximity, are considered in this replacement process. The bilateral filters use the following parameters: **d = 15, %sigma_c = 100**, **%sigma_s = 100**. The following image shows the difference obtained in the images. The parameters to be applied to the succeeding images are applied to the images at the predefined intervals. It is then

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

used to obtain the validation accuracy; the figure shows the results. A smoothing filter is a common approach for noise removal in OCT images. The most commonly used smoothing filter is the Gaussian filter, a linear filter that convolves the image with a Gaussian kernel. The Gaussian filter effectively removes high-frequency noise while preserving the edges and details in the image. The equation for a 2D Gaussian kernel is:

$$G(x,y) = \left(1/(2*\pi*\sigma^2)\right) * exp\left(-(x^2+y^2)/(2*\sigma^2)\right) \quad (1)$$

Where **x** and **y** are the spatial coordinates, sigma is the standard deviation of the Gaussian distribution, and π is the mathematical constant. The Gaussian filter is applied on the OCT images by the kernel is convolved with the image using the following equation:

$$I'(x,y) = \sum\left(\sum\left(I(x+i,y+j) * G(i,j)\right)\right) \quad (2)$$

where $I(x,y)$, is the original OCT image, $I'(x,y)$ is the filtered image, $G(i,j)$ is the Gaussian kernel, and the sums are taken over all i and j within the kernel size. Other filters, such as median, mean, or wavelet-based filters, can also be used for noise removal in OCT images. The filter choice depends on the specific characteristics of the image's noise and the desired smoothing level.

Finding The Region of Interest

The region of processing needs to be defined to improve the accuracy and reduce the complexity of the model's processing necessity. It helps in reducing the processing needed. The region far away from the retina is removed, and the background of the images is also removed. Three major steps are involved in identifying and selecting the region of interest. The first step involves the segmentation of the initial region. The second stage involves selecting areas with morphological operations and expanding eyes. In the final stage, the important area in the image is selected using the elimination filter, which can be seen in the following image.

Initial Segmentation

In this step, the algorithm's search area is reduced to avoid waste of time in other areas. The RPE and BM layers are separated from the ILM layers. For this process, the U-Net algorithm is used, a type of deep neural network designed specifically for the segmentation of medical images. The image segmentation is carried out through ResNet50 architecture. The ResNet-50 architecture consists of 50 deep layers trained with millions of images through the ImageNet database.

In a ResNet model, the ReLU function and batch normalization are in default. However, to skip the weights, the layers need to be trained to skip weights that are not important. The additional weight matrix used by the HighwayNet helps efficiently segments the medical images. The ResNet-50 employs convolution and average pooling with a SoftMax layer at the end of the model. It is widely used in detecting macular diseases from OCT images, as shown in Figure-2.

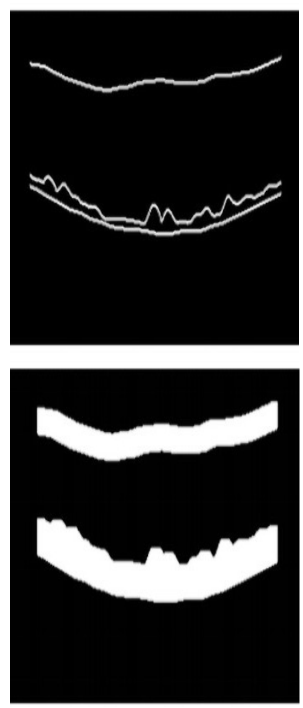


Figure-2. OCT Image Segmentation

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

ResNet50

The ResNet50 architecture (Quingge et al. [20]) mainly comprises the five conv2D layers. The convolutional and MaxPooling layers are grouped as it is how most of the residual networks are structured to function. After passing through the convolutional layer of size 64 with 2 strides length, the kernel size of the layer is 7x7. The convolutional layers are followed by a max pooling layer with 2 strides. The matrix of both the layers' kernel size is 3x3, 64x64, and 256x256, followed by another layer of kernel-size= 3x3 and filter size= 64. This combination is repeated three times, in which each MaxPooling layer of 2 strides follows.

This process is continued for the convolutional layers, the average pooling layers, and the fully connected layer. Finally, the classified results are sent to a SoftMax layer to classify the obtained features to predict the results effectively. The critical innovation of ResNet50 is the use of residual connections, which allow the network to learn better representations by mitigating the vanishing gradient problem. In traditional deep neural networks, the gradients can become very small as they propagate through many layers, making it difficult to update the weights of earlier layers. The residual connections allow the gradients to bypass (Figure-3) some of the layers, which helps to maintain the signal strength and enables the network to learn more effectively.

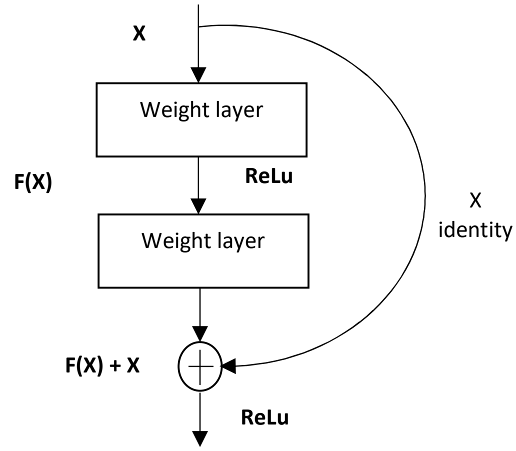


Figure-3. Bypassing The Layers

In order to reduce the computational complexity, ResNet50 follows skipping layers. While learning the data, if the feature information of the consequent data has similar weights and shapes, it skips those layers and adds to the output. This layer-skipping methodology is illustrated in Figure-3 and Figure-4. The layer connections are skipped when the output features are the same as the input features; for example, 32 x 32 -> 32 x 32, the filter mapping remains the same. If the size of the feature mapping is 32 x 32 -> 16 x 16, the depth of the feature map is doubled. In the ResNet-50 (Figure-5), the layer-connection skipping increases the performance of image processing and classification. It also speeds up the learning process more than the other deep learning models.

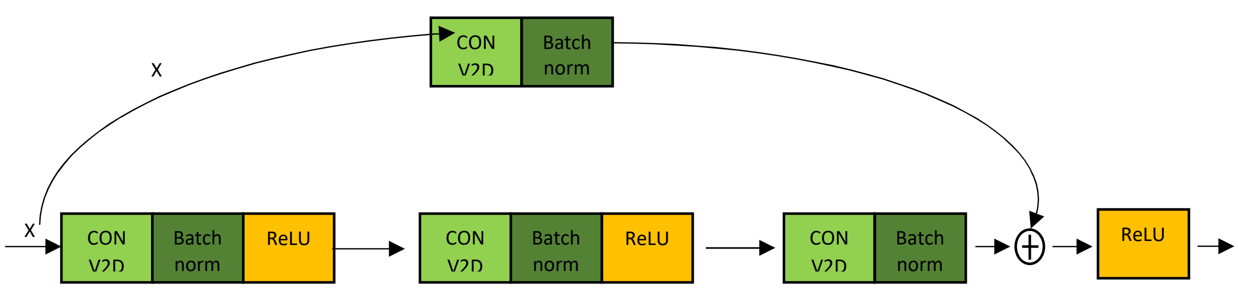


Figure-4. Skipping The Convolutional Layers

Let $\boldsymbol{X}$ be the input image to the network, and let $\boldsymbol{f}$ be a function that represents the ResNet50 architecture. Then the output of the network can be written as:

$$Y = f(X) \quad (3)$$

The function $\boldsymbol{f}$ can be decomposed into a sequence of building blocks, which are repeated multiple times to create the full network. Each building block consists of two or three convolutional layers, followed by a shortcut connection that adds the input to the output of the block. The mathematical expression of the residual connection is given as:

$$H(X) = F(X) + X \quad (4)$$

where $F(X)$, is the output of the convolutional layers in the building block, and $\boldsymbol{X}$ is the input to the block. The function $H(X)$, is the block output passed to the next block in the sequence. The mathematical expression of the building block is given as:

$$H(X) = f(W2, f(W1, X) + B1) + B2 \quad (5)$$

Where **X** is the input to the block, **W1**and **W2** are the weight matrices of the convolutional layers, ***B1*** and ***B2*** are the bias terms, and $\boldsymbol{f}$ is an activation function such as ReLU. The output of the first convolutional layer is passed through the activation function and then through the second convolutional layer. The shortcut connection adds the input $\boldsymbol{X}$ to the output of the second convolutional layer, and the result is passed through another activation function before being passed to the next block.

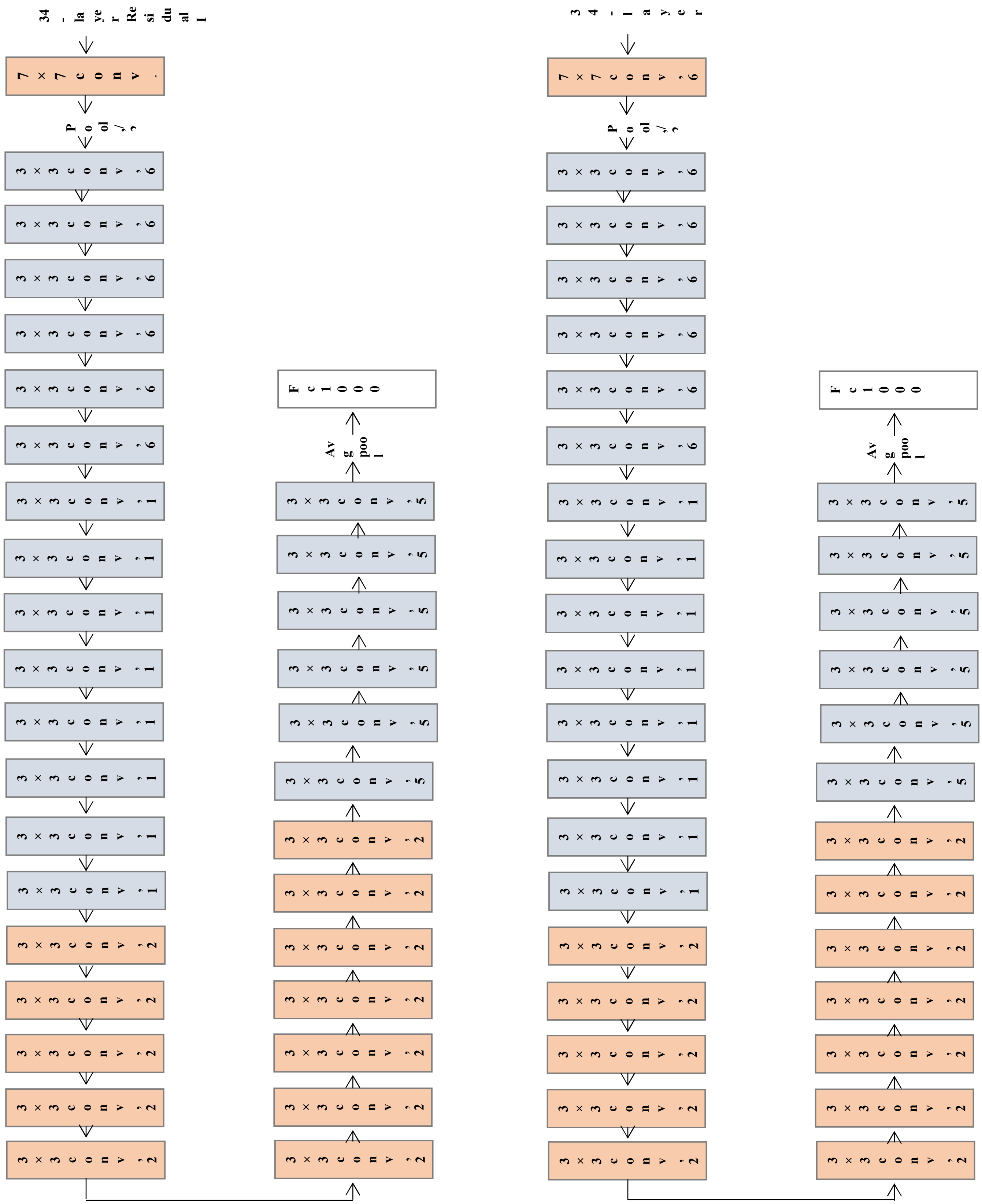


Figure-5. ResNet-50 Architecture

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Classification Using CNN Process

The image enhancement through the preprocessing steps is enough for the classification carried out through the CNN algorithm. The images' denoising and quality enhancement are obtained in the initial stage of the entire process. A typical CNN algorithm consists of the convolution, pooling, and fully connected layers followed by an activation layer to predict the results. The CNN model consists of 10 filters followed by batch normalization and dropout layers for regularisation, which is seen in the convolutional layers of the block. The activation function used is ReLU, with an FC layer and a SoftMax layer to classify the obtained features.

Input layer

The input images were resized to 225x225 pixels, the standard size for medical images. As the size varies with the devices, a standard size needs to be maintained to achieve better results from the prediction process. During the preprocessing, the input image is resized into 225x225x3, corresponding to the input layer's width, height, and number of channels

Convolutional Layer

The convolutional layer is important in the CNN algorithm, where most calculations are carried out. In this layer, the important features are obtained and saved without losing the spatial relationship of the images. It is obtained using several filters that learn the extracted features from the images without losing the relationship between them. The proposed model comprises 10 filters with 2-dimensional convolutional layers, each built using the $7x7$ filter size. The process of convolution filter is mathematically expressed as:

$$y_j = b_j + \sum_{c=0}^{n_c-1} \sum_{k=-p}^{p} x_{c,j-k} W_{c,k} \quad (6)$$

$$\frac{\delta L}{\delta x_{c,i}} = \sum_{k=-p}^{p} \frac{\delta L}{\delta y_{i+k}} w_{c,k} \quad (7)$$

$$\frac{\delta L}{\delta w_{c,k}} = \sum_{j=0}^{m-1} \frac{\delta L}{\delta y_j} x_{c,j-k} \quad (8)$$

$$\frac{\delta L}{\delta b_j} = \frac{\delta L}{\delta y_j} \quad (9)$$

The X represents the length of, **m,w** is the length of **n**, and **n** and **m** are the kernels. The result obtained is represented as y, whose size equals **x**. The **k** is the variable value present between **p** and **-p.** If the **k** exceeds any of the margins, it is termed zero. The gradients are calculated by finding the values of dl/dy. The bias term is represented by a constant value **b**. The channel index is **c** The channel index. The **a** represents the learned bias term of the network.

Batch normalization layer

Batch normalization trains the deep neural network where the mini-inputs are batched to the respective layer. It helps reduce the training epochs required by the model and standardizes and normalizes the learning of the model. In this process, the mean output of the layer is 0, and the standard deviation of the model is near 1. Additional hyperparameters like epsilon and momentum are set to 0.001 and 0.99 by default. The formula for batch normalization can be expressed as follows:

Algorithm BN()

{
Input: all the values of **x** sent for mini_batch as $B = \{x_{1\ldots m}\}$ α,β are learning parameters

Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

$\mu_B \leftarrow \frac{1}{m}\sum_{i=1}^{m} x_i$ **//mini-batch mean**

$\sigma_B^2 \leftarrow \frac{1}{m}\sum_{i=1}^{m} (x_i - \mu_B)^2$ **// mini-batch variance**

$\hat{x}_i \leftarrow \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}}$ **//normalize**

$y_i \leftarrow \gamma \hat{x}_i + \beta \equiv BN_{\gamma,\beta}(x_i)$ **//scale and shift**

}

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In the above algorithm, $\mathbf{m}$ denotes the number of mini-batches feed as input, the mean value is **μ, B** refers the present batch, and $\mathbf{x_i}$ represents the input data. The mean value and batch values are calculated from the sum of multiple inputs from a batch and divide by $\mathbf{m}$. The term σ_B^2 denotes the variance of the input values in the batch **B** $\hat{x}_i$ represents the normalizing and zer-centring input values. Finally the smoothened output value y_i is obtained from the batch normalization process.

ReLU layer

The main function of this layer is to replace the negative values with zero or make the positive values assume the convolved values and introduce the non-linearity to the feature maps of the algorithm.

$$f(x) = max(0, x) \tag{10}$$

Fully Connected layer

All the neurons of the previous layers are connected to the FC layer, which classifies the features extracted in the previous layers. It returns the features that are convolved to their respective class.

$$y_{jk}(x) = f\left(\sum_{i=1}^{nH} w_{jk} x_i + w_{j0}\right) \tag{11}$$

SoftMax Layer

Though the fully connected layers classify the features extracted from the previous layers, the probability of the extracted features needs to be calculated. For this calculation, the SoftMax function is used. The primary function of the SoftMax layer is to classify the values into either **1** or **0** based on the input.

$$\sigma(x)_i = \frac{e^{x_i}}{\sum_{j=1}^{K} e^{x_j}} \tag{12}$$

.

Output Layer

The output layer finally predicts whether the input layer consists of the DME.

Experimental Results and Discussion

The proposed ResNet-50 model is implemented in Python software, and the results are verified on OCT images. The reason for using the ResNet-50 model is because of its number of learning feature

Image Dataset

The dataset comprises OCT images from patients with normal and normal retinal disorders. The data was collected from patients aged 50 to 85 who exhibited intermediate to large drusen. Both eyes experienced the large drusen with the advanced AMD. The patients had no medical history of vitreoretinal surgery or another ophthalmic disease. The results collected under the AREDS2 protocol support the non-presence of macular druse in any eyes and the same for the proceeding years. The borders of the total retina limited the ground truth of the images, and it limited the internal membranes and Bruch's membrane between the retina. The delimitation of the epithelium that retinal pigmented also limited the drusen complex. The DOCTRAP software was used to segment the images automatically. The fovea in the images was manually marked to improve the accuracy of the segmentation process. The entire image dataset information is given in Table-1. It shows the number of images used for the training, testing, and validation process. It also shows the number of images under its classes.

The performance of the proposed model is evaluated based on the standard metrics used for the evaluation process. The batch size of the model is 64. Some widely used performance metrics are accuracy, precision, recall, and F1-Scores. The proposed model is simulated using python language in a system with an i7 7th gen CPU, 1050ti graphics and 8GB RAM, and 1TB SSD. The proposed model is simulated, and the results are tabulated and discussed in detail. Accuracy is the measure of how often a classifier correctly identifies a sample. It is defined as the ratio of the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

number of correct predictions to the total number of predictions made by the model.

$$Accuracy = (TP+TN)/(TP+TN+FP+FN) \quad (13)$$

Precision measures how often a classifier correctly identifies a positive sample. It is defined as the ratio of the number of true positives to the total number of samples predicted as positive by the model.

$$Precision = TP/(TP+FP) \quad \textbf{(14)}$$

Sensitivity or recall measures how often a classifier correctly identifies a positive sample out of all the positive ones. It is defined as the ratio of the number of true positives to the total number of positive samples.

$$Sensitivity = TP/(TP+FN) \quad \textbf{(15)}$$

Specificity measures how often a classifier correctly identifies a negative sample out of all the actual negative samples. It is defined as the ratio of the number of true negatives to the total number of negative samples.

$$Specificity = TN/(TN+FP) \quad (16)$$

F1-Score is the harmonic mean of precision and recall. It is a way to balance the trade-off between precision and recall when they have different weights in a particular task.

$$F1Score = 2*(Precision*Recall)/(Precision+Recall) \quad (17)$$

In the proposed model, ResNet50 is considered over ResNet101 because of its available resources and computational complexity. The size of the images is standardized for faster processing and classification by the neural network. The validation accuracy is comparable with the conventional approaches and pre-trained models. The proposed model is used to predict the DME in the OCT images, and the accuracy obtained is 92.03% in the earlier stage of DME. Table-1 shows the architecture of ResNet50. It consists of 5 convolutional layers. Over each layer, the image's output size decreases while the data's overall size increases. The teraflops consumed by the GPU are also given. Initially, the values of the images are set to 7x7 kernel, 64 filter size. Over each convolutional layer, the size of the filter increases, and finally, the filter size at the last convolutional layer is 1024 filter size. And the output size of the images will reduce and become 7x7.

7. TRAINING AND TESTING PROCESS

The training process involves optimizing the model's weights to minimize the loss function. You can use techniques such as gradient descent and backpropagation to accomplish this. It helps in achieving the maximum accuracy of the proposed model. For training the model, 80% of the data is taken. In terms of testing the model, the remaining 20% is used. The performance of the trained model is evaluated using the testing set. Calculate metrics such as accuracy, precision, recall, and F1 score to determine how well the model performs on unseen data. Hence, unlabelled data is used for testing the dataset. If the model's performance is unsatisfactory, you can fine-tune the model by adjusting the hyperparameters or modifying the architecture. Fine-tuning involves retraining the model using the original or an augmented training set.

Table-1. The Architecture of The Resnet50

ResNet-50 architecture				
Layer Name	**Output Size**	**18-Layer**	**34-Layer**	**50-Layer**
Conv1	224x224	7x7, 64		
Conv2	112x112			
Conv3	56x56	6x6, 64	6x6, 64	1x1, 64
		6x6, 64	6x6, 64	6x6, 64
		x2	x3	1x1, 128
Conv4	28x28	6x6, 128	6x6, 128	1x1, 128
		6x6, 128	6x6, 128	6x6, 128
		x2	x3	1x1, 256
Conv5	14x14	6x6, 256	6x6, 256	1x1, 256
		6x6, 256	6x6, 256	6x6, 256
		x2	x3	1x1, 512
Conv5	7x7	6x6, 512	6x6, 512	1x1, 512
		6x6, 512	6x6, 512	6x6, 512
		x2	x3	1x1, 1024
	1x1			
FLOPs		1.8x10^9	3.6x10^9	3.8x10^9

Overall, training and testing the classification of a dataset using ResNet50 involves a series of steps, from preparing the data to deploying the model. The process requires a good understanding of deep learning concepts and practices and can be time-consuming and resource-intensive. The first output level is verified in the experiment by applying a gaussian filter for denoising the input image. The filter value shown in Figure-6a denotes the input image with the noise value, and Figure-6b denotes the input image after denoising. The noise values degrade the image quality, which affects the pixel values.

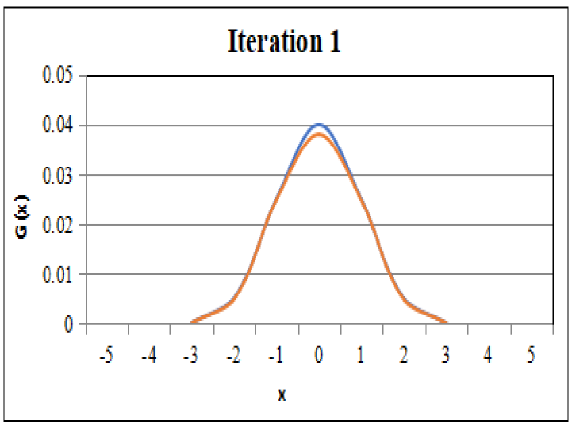


Figure-6a. Input Image With Noise (Gaussian Output)

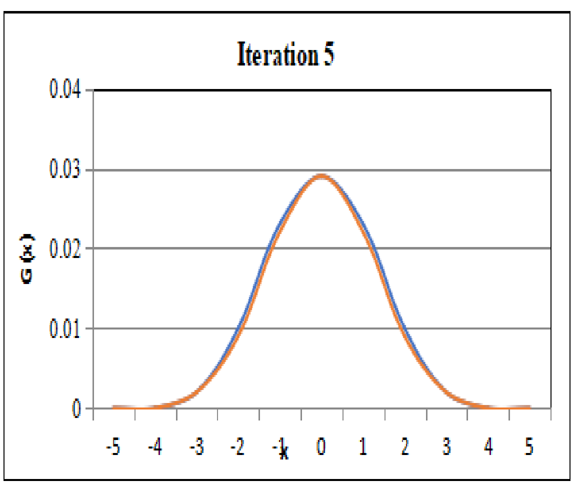
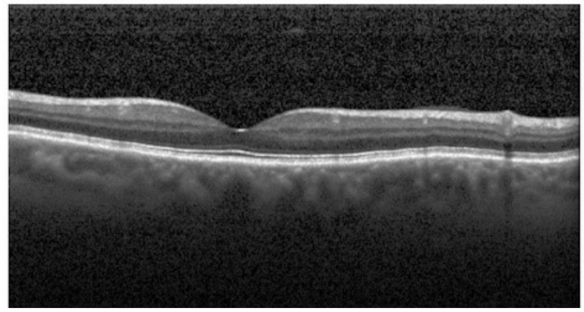


Figure-6b. Denoised Input Image (Gaussian Output)



(a). Input Image

Figure-7 Denoised input image (Gaussian output)

After preprocessing the image, the brightness level of the image is improved, and it exposes the availability of the various layers in the OCT images. The difference between the input and preprocessed images is shown in Figure-7(a) and 7b, respectively. The ResNet50 model receives 80% of the preprocessed images for the training process, which learns and classifies the given input images into four classes: normal, DME, Drusen, and CNV. DME, Drusen, and CNV are the abnormal retina cases found from the OCT images. Figure-8 shows all normal images classified by ResNet50, Figure-9 shows the DME images, Figure-10 shows the CNV images, and Figure-11 shows the Drusen images.

This paper implements a deep learning Convolutional neural network (CNN) classifier to detect the typical and DME-affected retinal image from the input OCT image. To efficiently classify the OCT image, the input data are classified into two phases: training and testing. 70% of the input data are classified for the training phase, and the remaining 30% are used for the testing phase. And to evaluate the efficiency of the proposed model, various performance metrics such as Specificity (Sc), sensitivity (Sn), and accuracy (AC) values are calculated using the following equations.

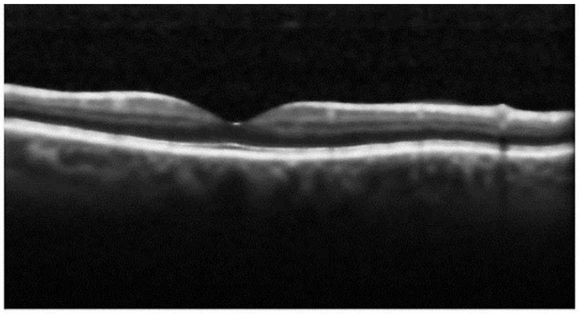
$$Sc = \frac{TN}{FP+TN} \tag{18}$$

$$Sn = \frac{TP}{TP+FN} \tag{19}$$

Journal of Theoretical and Applied Information Technology
15th August 2023. Vol.101. No 15

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$AC = \frac{(TP+TN)}{(TP=TN+FP+FN)} \quad (20)$$



(b).Denoised Image

In the equation-18, equation-19, and equation-20, TP represents the True positive, TN represents True Negative, FP represents the False positive, and FN represents the False-negative values of the proposed model.. The specificity value indicates the total amount of correctly predicted normal retinal images from the input OCT images

. The sensitivity result indicates the proposed model's ability to detect the DME-affected retinal images from the input OCT image. The Accuracy metrics define the overall accuracy result of the proposed model in detecting the normal and infected retinal images from the input OCT images

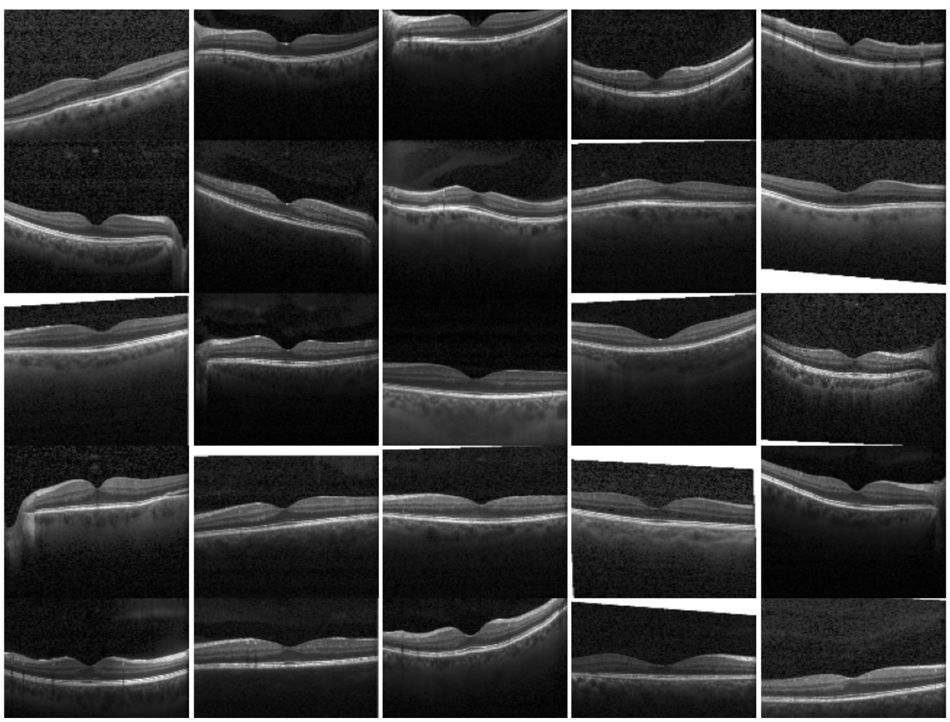


Figure-8. Normal Images Classified By ResNet-50

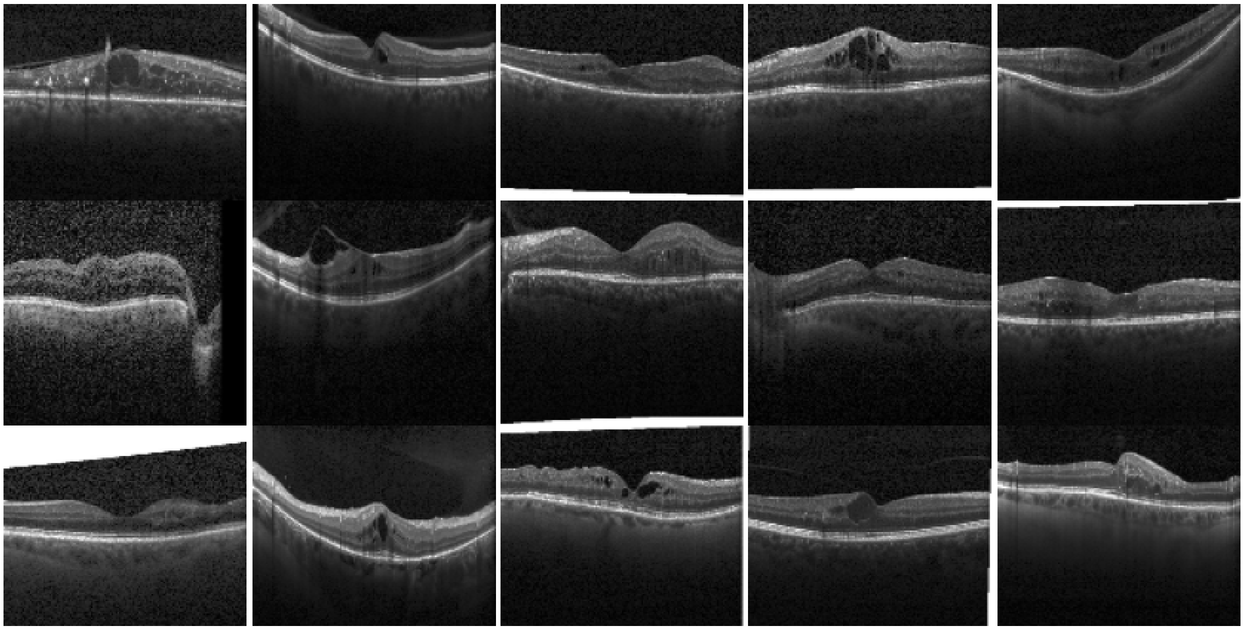


Figure-9. DME Images Classified By ResNet-50

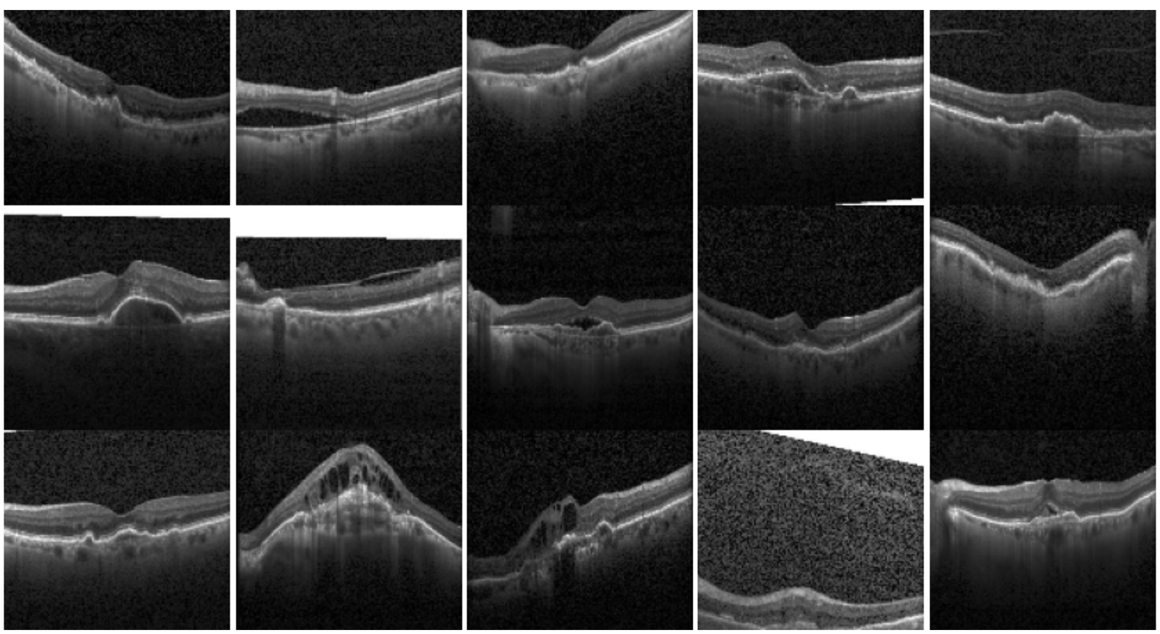


Figure-10. CNV Images Classified By ResNet-50

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

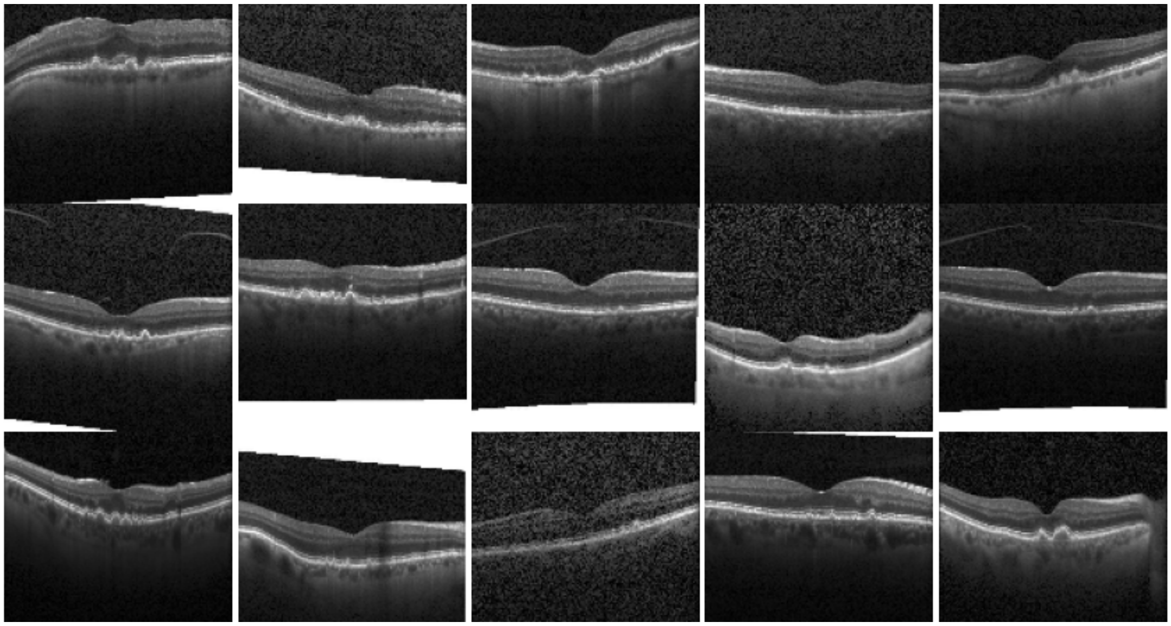


Figure-11. Drusen Images Classified By ResNet-50

Table-2. performance metric comparison result

Methods	Sensitivity (%)	Specificity (%)	Accuracy (%)
Existing CNN	93	56	75
roposed ResNet-50	90	88	97

Table-2 and Figure-13 define the performance metrics values of the proposed and existing model [1]. The classification accuracy is calculated based on the performance metrics, such as TP, TN, FP, and FN values, and these can also obtain by creating a confusion matrix shown in Figure-12. The confusion matrix shows the distribution of different retinal diseases in the dataset and the efficiency of the proposed model in detecting these different types of diseases, especially; the detection of DME, which is very important as it is used in the better prediction of the DME. It detects the important lesions in the DME dataset, as seen in the above model. Though it does not provide the required accuracy in other diseases, the accuracy in detecting the DME is very much comparable to the traditional models.

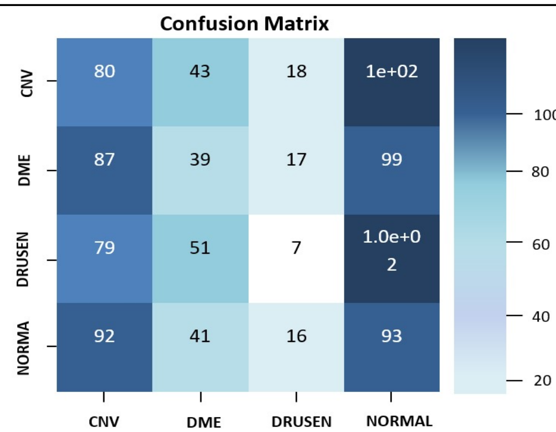


Figure-12. Confusion Matrix

The analysis results indicate that the proposed CNN-based OCT image classification technique has achieved 0.97, 0.88, and 0.9 accuracies, specificity, and sensitivity value, respectively, and the existing model classifies the input data with 93.75% sensitivity, 56.25% sensitivity, and 75% accuracy. In Figure-13, X-axis represents the various performance metrics, and the Y-axis indicates the total percentage of the metrics. Figure-14 indicates the accuracy value of the proposed CNN model. The analysis indicates that the proposed model effectively classified the standard and DME-affected retinal images with 97% accuracy. Figure-15 represents the average

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

loss value of the proposed CNN model for detecting the retinal images. The analysis indicates that the loss value decreases when the number of epochs increases. It is clearly defined that the proposed CNN-based model efficiently classifies the input sample with maximum accuracy and minimum loss value. Figure-16 illustrates the total number of normal and abnormal images in the input OCT image. The analysis indicates that the proposed CNN-based classification technique effectively classifies the input retinal images compared to others.

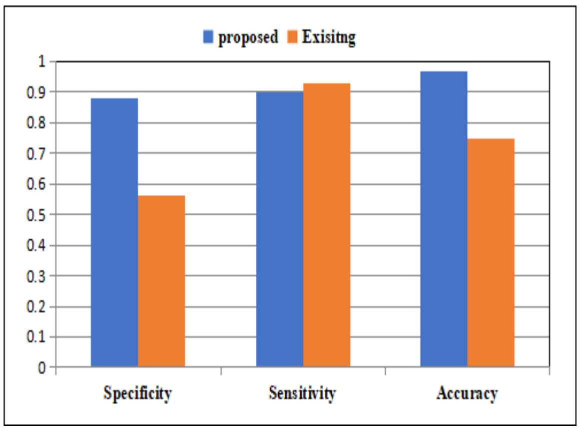


Figure-13. Performance Metrics

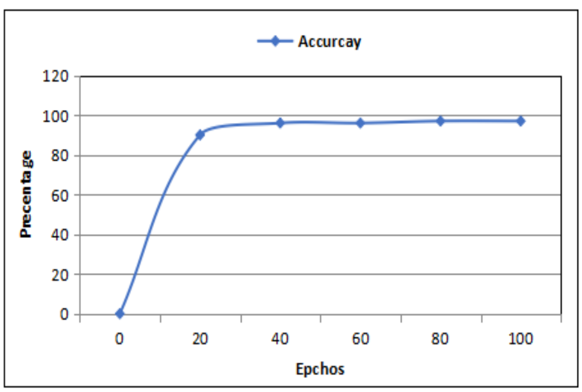


Figure-14. Accuracy Value of The Proposed Model

The overall result of this research work indicates that most patients with diabetes have highly affected by DME disease. Figure-17 depicts the age-wise classification of the DME-affected retinal images in the input OCT images. In Figure-17 age, wise data analyzed from the input OCT image is depicted. That people in the age group of 31 – 70 are highly affected due to DME. People below 30 and above 70 years are also affected by DME, but the severity and the total number of affected cases are less.

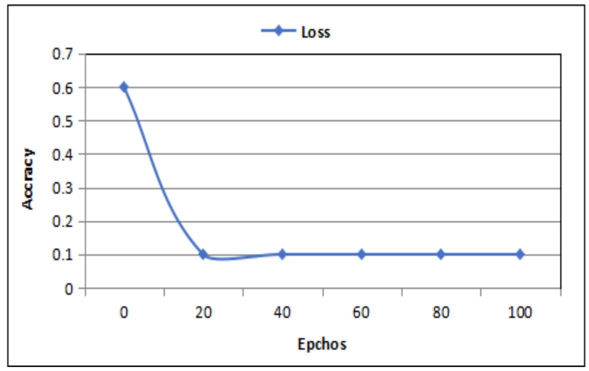


Figure-15. Loss Value of The Proposed Approach

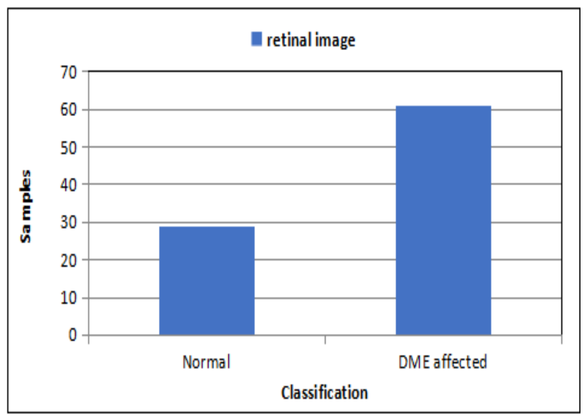


Figure-16. Normal and DME Affected Input Samples

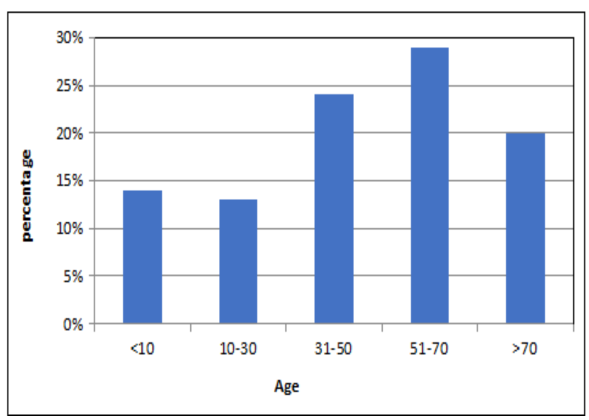


Figure-17. Age-wise Analysis

The analysis report indicates that 71% of people have diseases in both eyes. 18% were affected on the right, and 11% were affected on the left eye in the age group of 30-70. the overall result of the proposed CNN-based model proves that compared to the traditional methods, this approach performs better and accurately identifies the DME-affected retinal image from the OCT images.

In this model, the ResNet50 architecture used in the proposed model provides a better overview of OCT image processing. Thus, in this paper, the ResNet-50 is used to detect DME. The images are processed using a Gaussian filter to improve the quality of images. It also sharpens the edges of the images. This paper improves the accuracy of the DME disease classification in three ways. First, all the input images are preprocessed by removing the noise that improves the image's quality, and the brightness level of the image is enhanced, exposing the availability of the abnormal symptoms. Second, 80% of the images are used for training the transfer learning model to understand and detect any tiny differences in the input image. The model is evaluated and verified by random images and testing images. The proposed ResNet-50 stores the essential features for comparison in the residual unit and omits unwanted features that cannot help classification. The model can skip the layers to avoid gradient descent in the network.

8. CONCLUSION

The main aim of the paper is to design and implement an automatic transfer learning model for analysing and classifying the diseases of the Diabetic Macula Edema using OCT images. The complications of the OCT image analysis can be rectified by computerized programming method. Earlier research works have developed some of the computerized OCT diagnosis systems where their accuracy was less and are not reliable and scalable. Also they are semi-automatic and involves multiple sub methods to complete the diagnosis process. Thus, their complexity was high in terms of computation and time. To overcome all these kind of issues and to provide a better automatic OCT diagnosis system, this paper implemented one of the speedy and efficient deep learning algorithm, ResNet50. For improving the accuracy, this paper used pre-processing and batch normalization. The proposed ResNet50 model is experimented with python and a benchmark dataset to verify the efficacy. From the experimental and comparative results it is concluded that, the proposed ResNet50 model outperforms the others. The accuracy obtained using the proposed ResNet50 is 97%, which is higher than the existing models.

9. FUTURE WORK

In future work, the reliability and scalability is verified by experimenting the proposed model using a large volume of input image and evaluated by any random images not available in the same database.

REFERENCES:

[1]. Musat, O., Cernat, C., Labib, M., Gheorghe, A., Toma, O., Zamfir, M., & Boureanu, A. M. (2015). Diabetic macular edema. *Romanian journal of ophthalmology*, *59*(3), 133.

[2]. Romero-Aroca, P., Baget-Bernaldiz, M., Pareja-Rios, A., Lopez-Galvez, M., Navarro-Gil, R., & Verges, R. (2016). Diabetic macular edema pathophysiology: vasogenic versus inflammatory. *Journal of diabetes research*, *2016*.

[3]. Singer, M. A., Kermany, D. S., Waters, J., Jansen, M. E., & Tyler, L. (2016). Diabetic macular edema: it is more than just VEGF. *F1000Research*, *5*.

[4]. Bandello, F., Parodi, M. B., Lanzetta, P., Loewenstein, A., Massin, P., Menchini, F., & Veritti, D. (2017). Diabetic macular edema. *Macular Edema*, *58*, 102-138.

[5]. Chan, G. C., Muhammad, A., Shah, S. A., Tang, T. B., Lu, C. K., & Meriaudeau, F. (2017, September). Transfer learning for diabetic macular edema (DME) detection on optical coherence tomography (OCT) images. In *2017 IEEE international conference on signal and image processing applications (ICSIPA)* (pp. 493-496). IEEE.

[6]. Kamble, R. M., Chan, G. C., Perdomo, O., Kokare, M., Gonzalez, F. A., Müller, H., & Mériaudeau, F. (2018, December). Automated diabetic macular edema (DME) analysis using fine tuning with inception-resnet-v2 on OCT images. In *2018 IEEE-EMBS Conference on*

Biomedical Engineering and Sciences (IECBES) (pp. 442-446). IEEE.

[7]. Vahadane, A., Joshi, A., Madan, K., & Dastidar, T. R. (2018, April). Detection of diabetic macular edema in optical coherence tomography scans using patch based deep learning. In *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* (pp. 1427-1430). IEEE.

[8]. Li, F., Chen, H., Liu, Z., Zhang, X., & Wu, Z. (2019). Fully automated detection of retinal disorders by image-based deep learning. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *257*, 495-505.

[9]. Li, F., Chen, H., Liu, Z., Zhang, X. D., Jiang, M. S., Wu, Z. Z., & Zhou, K. Q. (2019). Deep learning-based automated detection of retinal diseases using optical coherence tomography images. *Biomedical optics express*, *10*(12), 6204-6226.

[10].Chetoui, M., & Akhloufi, M. A. (2020, June). Deep retinal diseases detection and explainability using OCT images. In *Image Analysis and Recognition: 17th International Conference, ICIAR 2020, Póvoa de Varzim, Portugal, June 24–26, 2020, Proceedings, Part II* (pp. 358-366). Cham: Springer International Publishing.

[11].Nazir, T., Irtaza, A., Javed, A., Malik, H., Hussain, D., & Naqvi, R. A. (2020). Retinal image analysis for diabetes-based eye disease detection using deep learning. *Applied Sciences*, *10*(18), 6185.

[12].Wu, Q., Zhang, B., Hu, Y., Liu, B., Cao, D., Yang, D., ... & Yu, H. (2021). Detection of morphologic patterns of diabetic macular edema using a deep learning approach based on optical coherence tomography images. *Retina (Philadelphia, Pa.)*, *41*(5), 1110.

[13].Tang, F., Wang, X., Ran, A. R., Chan, C. K., Ho, M., Yip, W., ... & Cheung, C. Y. (2021). A multitask deep-learning system to classify diabetic macular edema for different optical coherence tomography devices: a multicenter analysis. *Diabetes Care*, *44*(9), 2078-2088.

[14].Altan, G. (2022). DeepOCT: An explainable deep learning architecture to analyze macular edema on OCT images. *Engineering Science and Technology, an International Journal*, *34*, 101091.

[15].Upadhyay, P. K., Rastogi, S., & Kumar, K. V. (2022). Coherent convolution neural network based retinal disease detection using optical coherence tomographic images. *Journal of King Saud University-Computer and Information Sciences*, *34*(10), 9688-9695.

[16].He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[17].Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., ... & Ng, A. Y. (2017). Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning. *arXiv preprint arXiv:1711.05225*.

[18].Loey, M., Smarandache, F., & M. Khalifa, N. E. (2020). Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning. *Symmetry*, *12*(4), 651.

[19].Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and engineering sciences in medicine, 43, 635-640.

[20].Ji, Q., Huang, J., He, W., & Sun, Y. (2019). Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images. Algorithms, 12(3), 51.

[21].Chen YM, Huang WT, Ho WH, Tsai JT. Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning. BMC Bioinformatics. 2021 Nov 8;22(Suppl 5):99. doi: 10.1186/s12859-021-04001-1. PMID: 34749641; PMCID: PMC8576967.

[22].Wang J, Deng G, Li W, Chen Y, Gao F, Liu H, He Y, Shi G. Deep learning for quality assessment of retinal OCT images. Biomed Opt Express. 2019 Nov 4;10(12):6057-6072. doi: 10.1364/BOE.10.006057. PMID: 31853385; PMCID: PMC6913385.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[23].Loey, M., Smarandache, F., & M. Khalifa, N. E. (2020). Within the lack of chest COVID-19 X-ray dataset: a novel detection model based on GAN and deep transfer learning. Symmetry, 12(4), 651.

[24].Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Physical and engineering sciences in medicine, 43, 635-640.

[25].A.VictorIkechukwu, S. Murali, R. Deepu, R.C . Shivamurthy, (2021), "ResNet-50 vs VGG-19 vs training from scratch: A comparative analysis of the segmentation and classification of Pneumonia from chest X-ray images", Global Transitions Proceedings, Elsevier, https://doi.org/10.1016/j.gltp.2021.08.027.

[26].Li, F., Chen, H., Liu, Z., Zhang, X. D., Jiang, M. S., Wu, Z. Z., & Zhou, K. Q. (2019). Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomedical optics express, 10(12), 6204-6226.

[27].Nugroho, K. A. (2018, October). A comparison of handcrafted and deep neural network feature extraction for classifying optical coherence tomography (OCT) images. In 2018 2nd International Conference on Informatics and Computational Sciences (ICICoS) (pp. 1-6). IEEE.

[28].Sun, Y., Zhang, H., & Yao, X. (2020). Automatic diagnosis of macular diseases from OCT volume based on its two-dimensional feature map and convolutional neural network with attention mechanism. Journal of Biomedical Optics, 25(9), 096004-096004.