**Multi-Class Detection of Eye Fundus Diseases using Deep Learning: A Comparative Study of Cataract, Diabetic Retinopathy, and Glaucoma Detection**

Dr.Sasikala devi And Parthiban C

**Abstract :**

This study aims to develop an automated system for detecting and classifying three common eye fundus diseases, namely cataract, diabetic retinopathy, and glaucoma, using deep learning techniques.It utilizes a pre-trained convolutional neural network (CNN) architecture(ResNet101)which attained an accuracy of 0.81. Deep learning algorithms have shown promising results in the medical field for image recognition tasks, and in particular, have been successfully applied to diagnose and classify eye fundus diseases.Regular eye exams are highly recommended.

**1.Introduction:**

Eye fundus conditions that impact the retina and blood vessels at the rear of the eye include cataract, glaucoma, and diabetic retinopathy. Aging, medicines, and UV rays can all contribute to cataract, which causes clouding of the eye lens and blurry vision. If ignored, glaucoma harms the optic nerve and raises eye pressure, resulting in vision loss. Its symptoms can develop quietly and it is a major cause of blindness worldwide. The main cause of blindness in working-age people is diabetic retinopathy, which is brought on by elevated blood sugar levels and damages blood vessels, resulting in vision loss. For those at risk, such as older individuals, people with diabetes, and people with high blood pressure, early detection and treatment of these eye illnesses are essential to avoid further eyesight loss.

Among the most common eye fundus disorders are cataract, glaucoma, and diabetic retinopathy. There are currently 20 million cataract sufferers, and by the year 2050, that number is anticipated to double. While one-third of diabetics develop diabetic retinopathy, which causes blindness in working-age adults, glaucoma impacts 76 million individuals worldwide. Deep learning has demonstrated tremendous promise in the analysis and diagnosis of ocular fundus illnesses, including disease detection, classification, and severity evaluation, disease segmentation, and therapy planning. For fundus images to acquire intricate patterns and connections, deep neural networks are needed. Deep learning techniques can increase diagnostic precision and speed, resulting in early treatment and less visual loss.

**2.Related work:**

**Table 1.** Deep Learning Models and their Accuracies for Various models in Recent Years

| Authors | Model | Accuracy |
| --- | --- | --- |
| (Sesikala, B., Harikiran, J., & SaiChandana, B. 2022) | CNN | 99.89% |
| (Suganyadevi, S., Renukadevi, K., Balasamy, K., & Jeevitha, P. 2022) | CNN | 85% |
| (Butt, M. M., Iskandar, D. A., Abdelhamid, S. E., Latif, G., & Alghazo, R. 2022) | GoogleNet | 97.8% |
| (Singh, L. K., Khanna, M., & Thawkar, S. 2022) | KNN | 99% |
| (Butt, M. M., Iskandar, D. A., Abdelhamid, S. E., Latif, G., & Alghazo, R. 2022) | ResNet | 89.29% |
| (Qureshi, I., Ma, J., & Abbas, Q. 2021) | ADL-CNN | 98% |
| (Akbar, S., Hassan, S. A., Shoukat, A., Alyami, J., & Bahaj, S. A. 2022) | DarkNet + DenseNet | 99.7% |
| (Gupta, S., Panwar, A., Kapruwan, A., Chaube, N., & Chauhan, M. 2022) | Inception V3 | 92% |
| (Kaushik, H., Singh, D., Kaur, M., Alshazly, H., Zaguia, A., & Hamam, H. 2021) | Stacked Convolutional Neural Network | 97.92% |
| (Sarki, R., Ahmed, K., Wang, H., & Zhang, Y. 2020) | AlexNet | 97.93% |

The dataset, preparation methods, and model design are just a few examples of the many variables that can affect how accurately deep learning models are evaluated.

Following are some notes regarding the models mentioned with their stated accuracy.

A very high precision of 99.89% reported by one model (CNN by Sesikala et al., 2022) may be a sign of overfitting or a limited, straightforward dataset.

A different model, called Inception V3 by Gupta et al., 2022, gives an accuracy of 92%, which is still a respectably high accuracy but less than some of the other models on the list.Butt et al. (2022) also used the ResNet architecture, but achieved a lower accuracy of 89.29% in detecting diabetic retinopathy. This may indicate that ResNet is not as effective for this particular task.Akbar et al. (2022) used a combination of DarkNet and DenseNet architectures, achieving an impressive accuracy of 99.7% in detecting diabetic retinopathy.Using a KNN algorithm, Singh et al. (2022) detected diabetic retinopathy with a remarkable 99% accuracy. One of the most accurate items on the list is this one.With the help of a Stacked Convolutional Neural Network, Kaushik et al. (2021) were able to identify diabetic retinopathy with an accuracy of 97.92%. This is yet another impressive outcome that demonstrates the potency of this specific paradigm.

**3.Proposed framework:**

**3.1 ResNet:**

ResNet (short for "Residual Network") is a deep neural network architecture that was introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun from Microsoft Research. The ResNet architecture was designed to address the vanishing gradient problem that occurs when training very deep neural networks.

ResNet101 is a convolutional neural network architecture that was introduced in 2015 as part of the ResNet family of models. It was created to tackle the problem of vanishing gradients, which can occur in deep neural networks and cause the network to stop learning effectively.

The ResNet101 architecture is based on a residual block structure, where the input to a block is added to the output of the block, allowing the network to learn residual functions. This approach helps to reduce the vanishing gradient problem, allowing for deeper networks to be trained effectively.The ResNet101 architecture has 101 layers, including a convolutional layer, a max pooling layer, and a fully connected layer. It also includes four stages of residual blocks, each with a different number of blocks, and ends with a global average pooling layer and a fully connected layer.

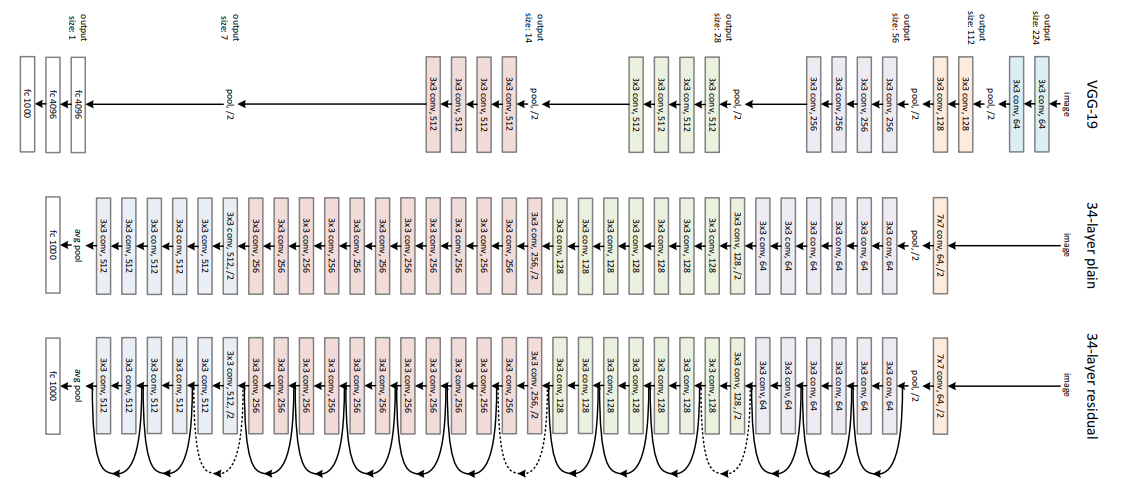
The first stage of the ResNet101 architecture consists of a single convolutional layer followed by a max pooling layer. The second stage includes three residual blocks, each with two convolutional layers. The third and fourth stages each include four and 23 residual blocks, respectively, with varying numbers of convolutional layers.One of the key features of ResNet-101 is its use of bottleneck layers, which are composed of 1x1, 3x3, and 1x1 convolutions. The 1x1 convolutions are used to reduce the number of input channels, while the 3x3 convolutions are used to extract features. The second 1x1 convolution is used to restore the number of channels to the original size. This bottleneck design helps to reduce the computational cost of the network while maintaining its accuracy.

The final global average pooling layer is used to reduce the spatial dimensions of the output feature map to a vector, which is then fed into a fully connected layer to produce the final classification output.ResNet101 has been used in a variety of computer vision tasks, including image classification, object detection, and semantic segmentation. It has achieved state-of-the-art results on many benchmark datasets, including the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) and the COCO object detection dataset.

One drawback of ResNet101 is its high computational cost, due to its large number of layers and complex structure. However, various techniques have been developed to reduce the computational burden, such as using smaller input images or using faster inference techniques.

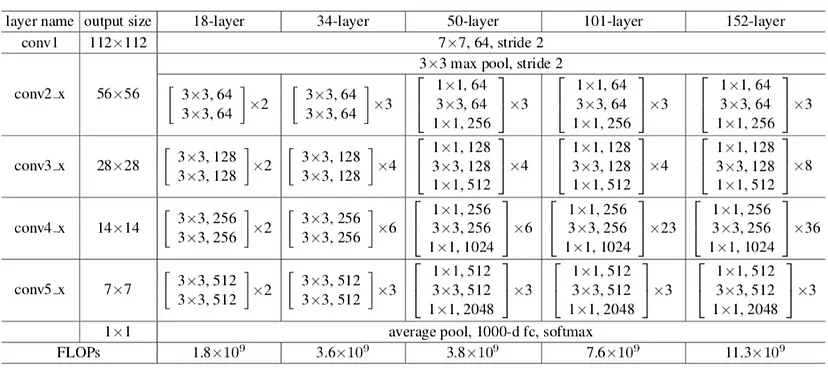
ResNet101 is a powerful and versatile convolutional neural network architecture that has made significant contributions to the field of computer vision. Its ability to effectively learn residual functions has helped to overcome the problem of vanishing gradients, allowing for deeper networks to be trained effectively.

ResNet has achieved state-of-the-art performance on many computer vision tasks, including image classification, object detection, and segmentation. Its success has led to the development of many variants, including ResNet-18, ResNet-34, ResNet-50, ResNet-101, and ResNet-152, which differ in the number of layers and the depth of the network.among these ResNet-101 was utilized in developing a deep learning model.



**Figure 1.** ResNet-101 Network architecture

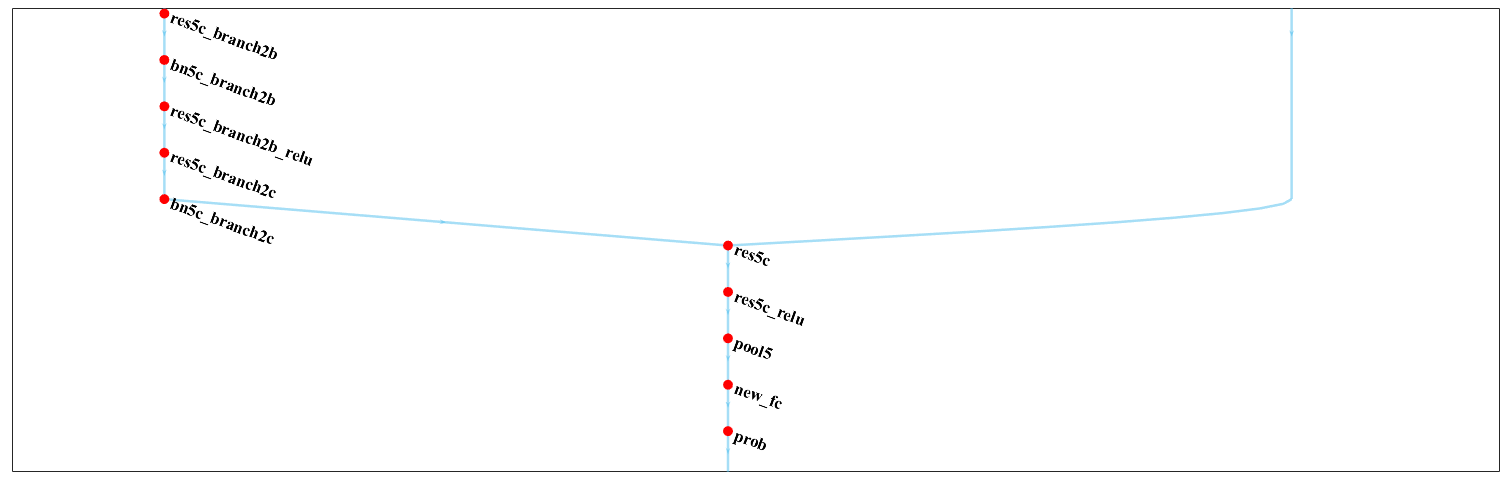
The dotted line is there, precisely because there has been a change in the dimension of the input volume (of course a reduction because of the convolution). Note that this reduction between layers is achieved by an increase on the stride, from 1 to 2, at the first convolution of each layer; instead of by a pooling operation, which we are used to see as down samplers.In the table, there is a summary of the output size at every layer and the dimension of the convolutional kernels at every point in the structure.

**Table 2.** Sizes of outputs and convolutional kernels.

In deep learning, a layer graph is a visual representation of the architecture of a neural network. It is a diagram that shows how the input data is processed through a series of layers in the network, ultimately producing an output.Each layer in the graph represents a specific mathematical operation or transformation that is applied to the input data.

The layer graph typically includes input and output nodes that represent the data being fed into the network and the final output of the network, respectively. In addition, it may include multiple layers of different types, such as convolutional, pooling, activation, and normalization layers.

Visualizing the layer graph can be helpful in understanding the architecture of a neural network and how it processes the input data. It can also be used to identify potential bottlenecks or areas for improvement in the network architecture.



**Figure 2.**layer graph depicting the newly established layers

**3.2 Transfer learning:**

Transfer learning is a technique in machine learning and deep learning that involves using knowledge gained from training a model on one task to improve the performance of a model on a different but related task.

In transfer learning, a pre-trained model is used as a starting point for a new task, instead of training a new model from scratch. The pre-trained model has already learned a set of features from a large dataset, which can be useful in solving a new task with a smaller dataset.

There are two main approaches to transfer learning:

1.Fine-tuning 2.Feature extraction

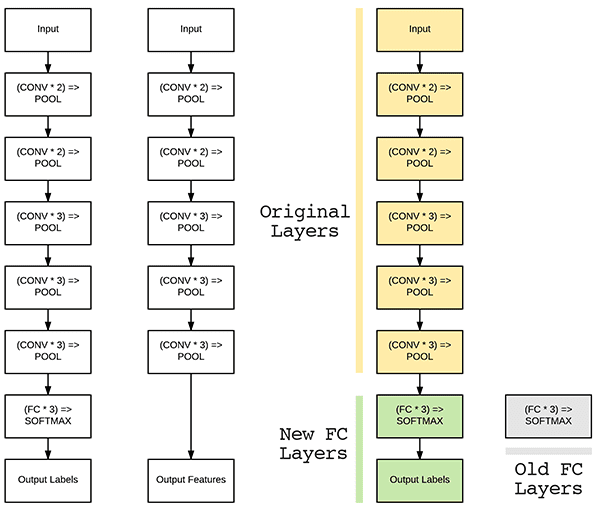
Transfer learning has many benefits, including:Faster training time,Improved accuracy,Reduced risk of overfitting,

In deep learning approaches for eye fundus disease detection, the final layer modification is a crucial step in adapting a pre-trained model to the specific task at hand. The final layer of a pre-trained deep learning model is typically a fully connected layer that outputs a vector of class probabilities.

To adapt the pre-trained model for the task of eye fundus disease detection, the final layer needs to be modified to output the desired number of classes. In this case, the final layer would need to be modified to output probabilities for the specific eye fundus diseases being targeted, such as cataract, diabetic retinopathy, and glaucoma.

There are different ways to modify the final layer, depending on the specific deep learning model being used. One common approach is to remove the final fully connected layer and replace it with a new fully connected layer that has the desired number of output neurons corresponding to the number of classes being targeted. This new layer is randomly initialized and trained along with the rest of the network using backpropagation and gradient descent.

Another approach is to keep the pre-trained final layer, freeze its weights, and add a new fully connected layer on top of it with the desired number of output neurons. This new layer is then trained along with the other layers, with the weights of the pre-trained layers fixed.

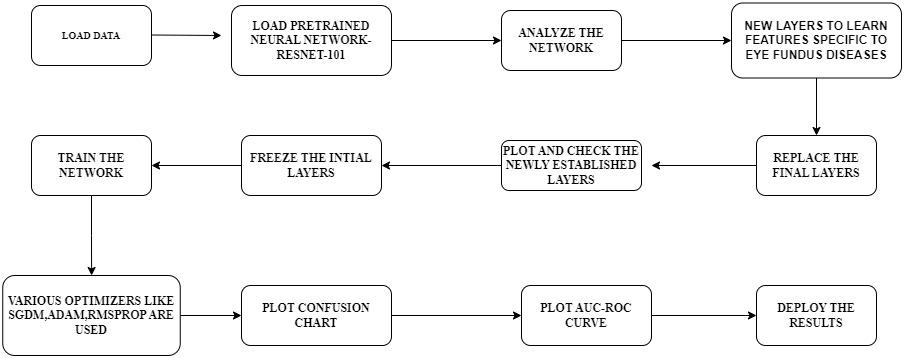


**Figure 3.** Diagram showing Transfer learning with fine tuning

Overall, modifying the final layer of a pre-trained deep learning model is a key step in adapting it for a specific task such as eye fundus disease detection, and there are various approaches to accomplish this depending on the specific model and task at hand.

**3.3 Flow diagram :**

Here is the flowchart which depicts the work flow of the study for detecting and classifying three common eye fundus diseases, namely cataract, diabetic retinopathy, and glaucoma, using deep learning techniques.



**Figure 4.** Flowchart representing the workflow of the study.

**4.Experimental analysis:**

**4.1 System configuration:**

The computer system is equipped with an Intel Core i5-10300H processor ,base clock speed of 2.50GHz,2496 Mhz,4 cores,8 Logical processors, 8GB DDR4 RAM, a 512GB NVMe solid-state drive, and an NVIDIA GeForce GTX 1650 graphics card. The operating system installed is Windows 11 Home 64-bit. It also features a Gigabit Ethernet port, Wi-Fi 6 connectivity, Bluetooth 5.1, and a range of USB 3.2 Gen 1 and Gen 2 ports. The system is housed in a sleek black aluminum chassis and features RGB lighting on the front panel. It is designed for high-performance computing tasks such as gaming, video editing, and 3D rendering.

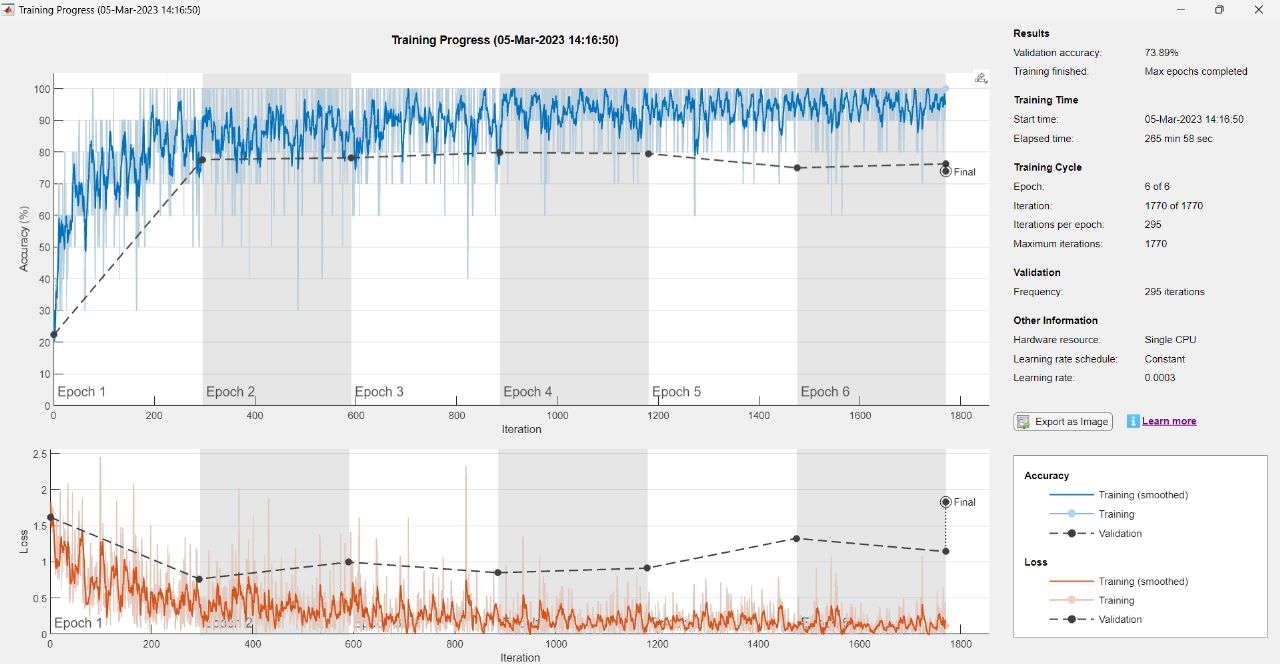
MATLAB Version: 9.13.0.2105380 (R2022b) Update 2

MATLAB License Number: 40525263

**4.2 Various optimizers used :**

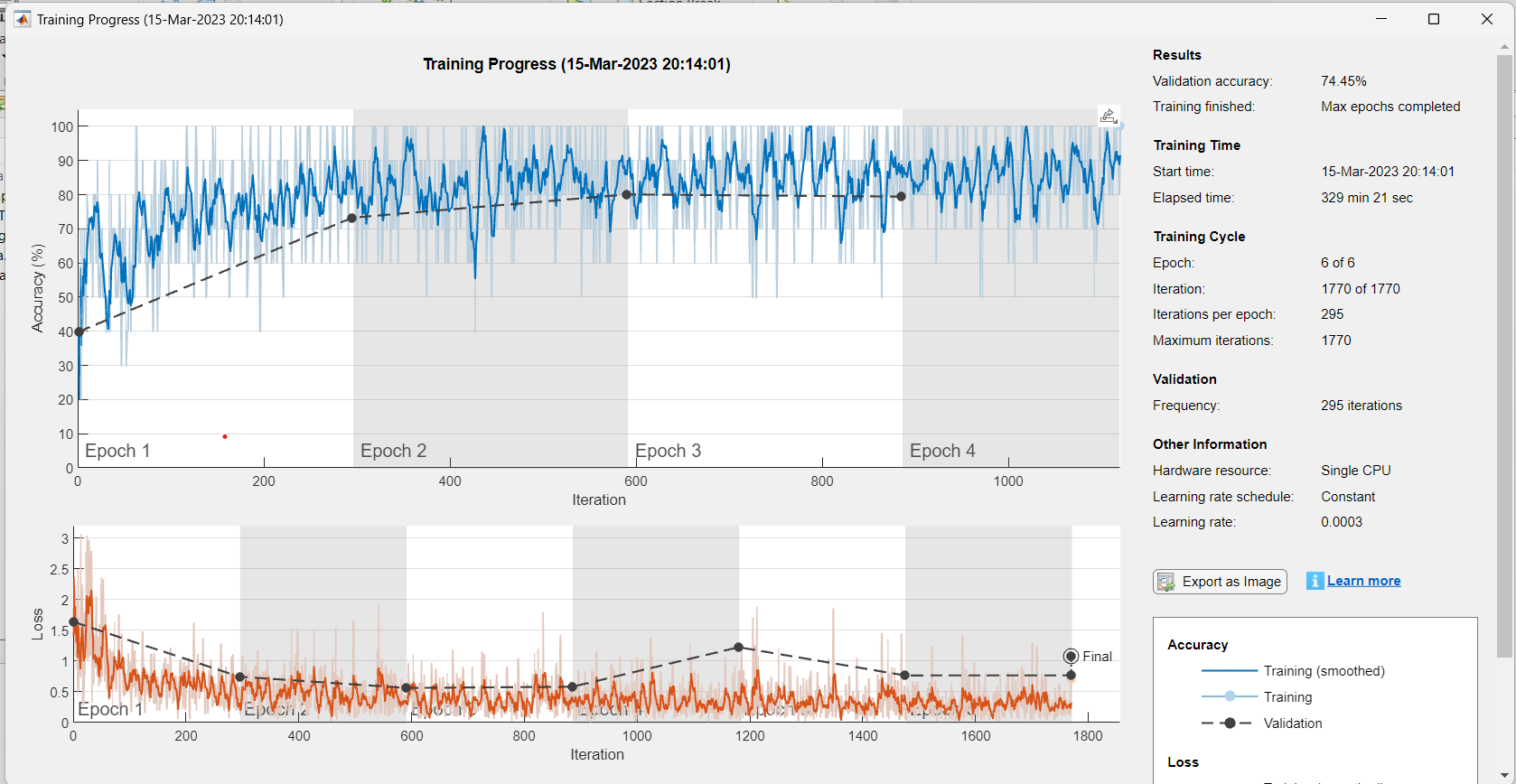
The weights and biases of the neural network are modified by optimizers during the training of deep learning models in order to reduce the loss function. The speed and precision of the training process can be significantly impacted by the optimizer selection.

There are several optimizers commonly used in deep learning, including Stochastic Gradient Descent (SGD), Adam, Adagrad, RMSprop.

**4.2.1. SGD** is a simple and widely used optimizer that updates the model parameters based on the gradient of the loss function with respect to the weights.

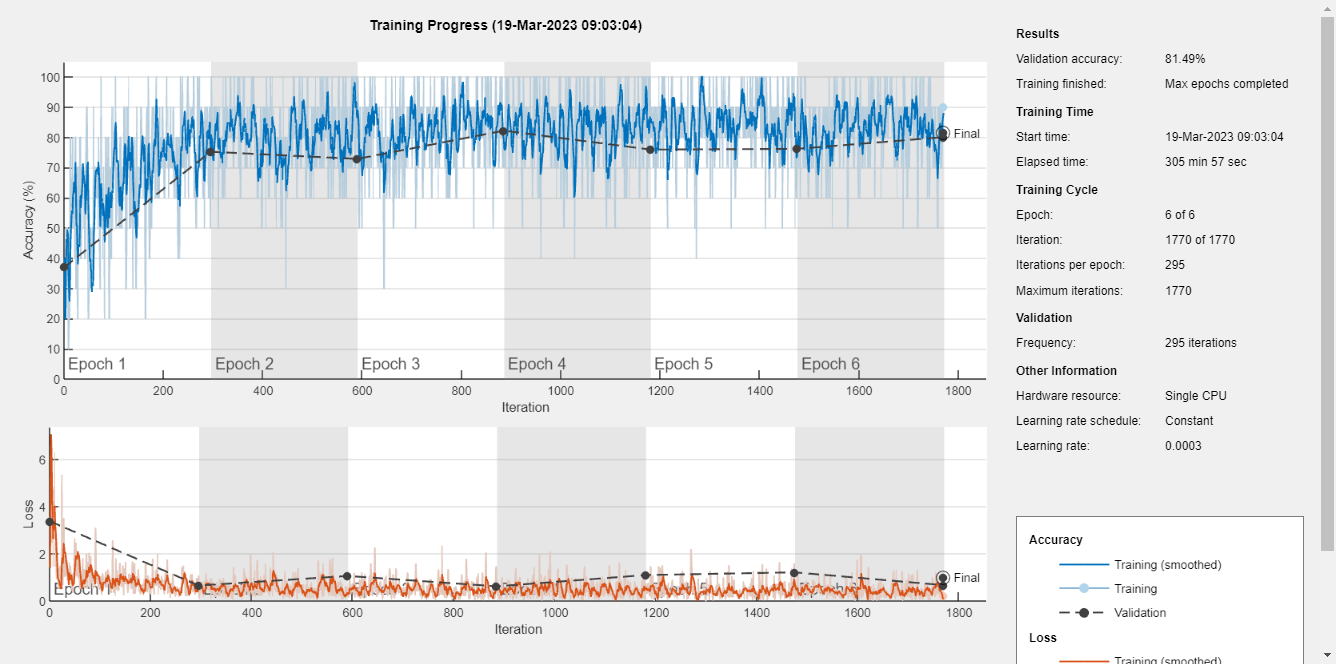
**Figure 5.** Training the network with SGDM optimizer

**4.2.2.Adam** is considered an adaptive learning rate optimizer because it uses adaptive moment estimates to scale the learning rate based on the first and second moments of the gradient



**Figure 6.** Training the network with ADAM optimizer

**4.2.3.RMSprop** uses a moving average of the squared gradient to scale the learning rate for each parameter. This means that the learning rate for each weight is adjusted according to the variance of the gradients, allowing for larger updates to be made for infrequent parameters and smaller updates for frequent parameters.



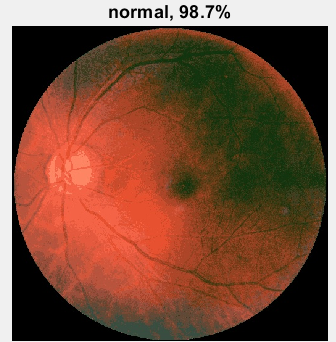
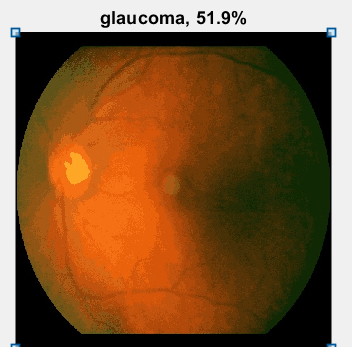
**Figure 7.** Training the network with RMSprop optimizer

Each optimizer has its own strengths and weaknesses, and the choice of optimizer often depends on the specific dataset and neural network architecture being used. By carefully selecting the optimizer and tuning its parameters, it is possible to achieve faster convergence and better performance during the training process.

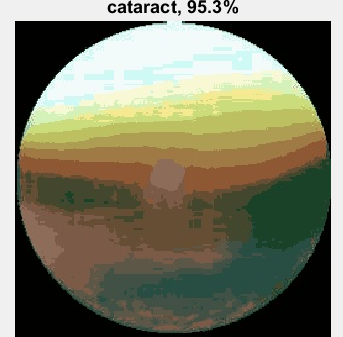
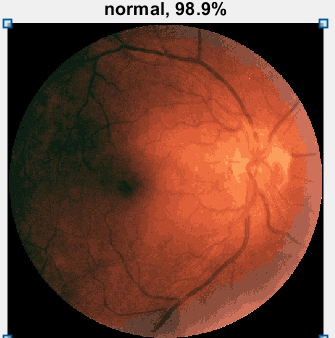
In deep learning, a validation plot is a graph that shows the performance of a model on a validation dataset during the training process. It is typically used to monitor the performance of a model and to identify when overfitting is occurring.

4.3 **Validation plot:**

The validation plot typically plots the validation loss (or error) on the y-axis and the number of training epochs on the x-axis. As the model is trained over time, the validation loss should decrease, indicating that the model is improving its performance on the validation dataset. However, if the validation loss starts to increase while the training loss continues to decrease, this indicates that the model is starting to overfit to the training data and its performance on new, unseen data may be compromised.By monitoring the validation plot, it is possible to identify when the model is starting to overfit and to adjust the model architecture or training process accordingly to prevent this from happening. Validation plots can also be used to compare the performance of different model architectures or hyperparameters, and to select the best model for a given task.

**Figure 8a-normal Figure 8b -glaucoma**

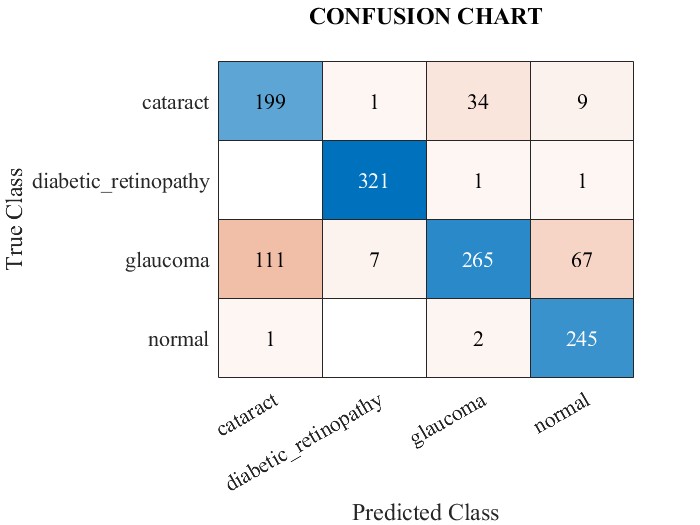
 

**Figure 8c -cataract Figure 8d-normal**

**Figure 8.** Validation images

**4.4 Confusion chart:**

The confusion matrix is a matrix used to determine the performance of the classification models for a given set of test data. It can only be determined if the true values for test data are known. The matrix itself can be easily understood, but the related terminologies may be confusing. Since it shows the errors in the model performance in the form of a matrix, hence also known as an error matrix.

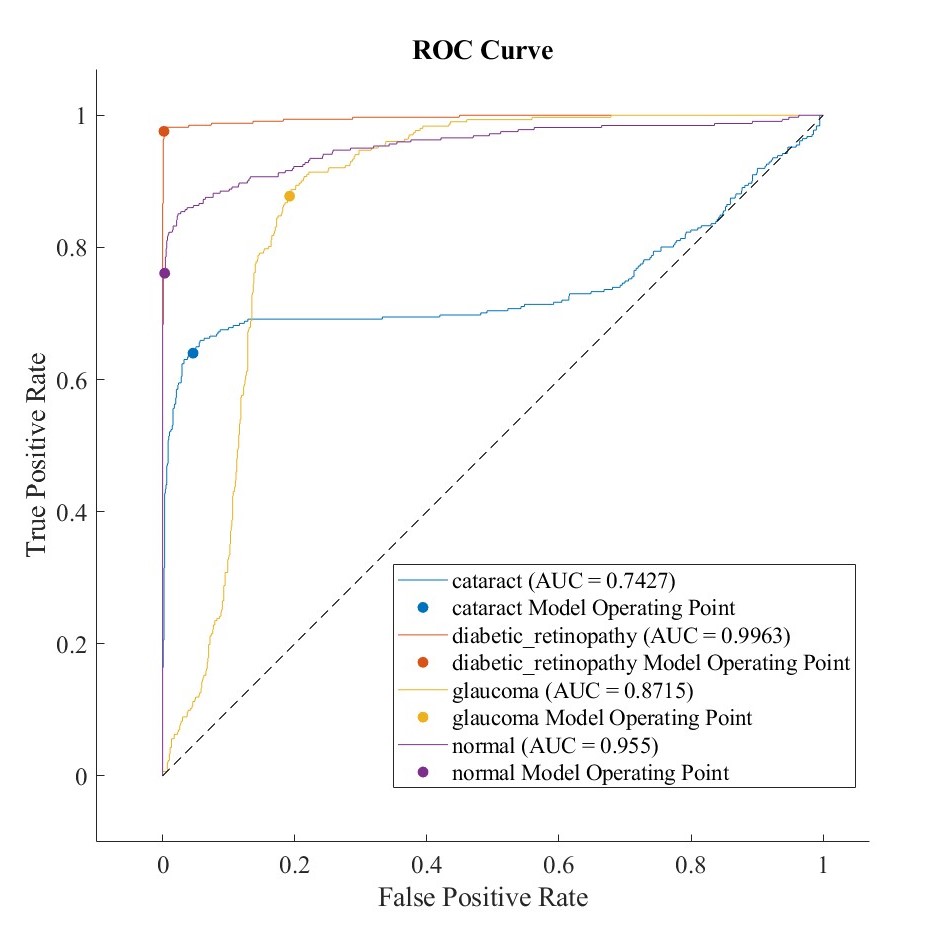


**Figure 9.** Confusion Chart

**4.5 Roc-Auc plot:**

An indicator of performance for classification issues at different threshold levels is the AUC-ROC curve. AUC stands for the level or measurement of separability, while ROC is a probability curve. It reveals how well the model can differentiate across classes. The model is more accurate at classifying 0 classes as0, and classifying 1 classes as1, the higher the AUC. Comparatively, the model is more effective at differentiating between individuals who have the condition and those who do not, the higher the AUC.

TPR is plotted against FPR on the ROC curve, with FPR on the x-axis and TPR on the y-axis.

The roc-auc plot for our approach using ResNet-101: 

**Figure 10.** AUC-ROC Curve

**5.Conclusion:**

The methods used to identify Eye Fundus Disease are given a thorough discussion in this review study. Using color fundus images, this method could distinguish between the three types of eye fundus disease (cataract, glaucoma, and diabetic retinopathy) with excellent accuracy, sensitivity, specificity, and AUC.the model has achieved an accuracy of 81.49% while using RMSprop optimizer whereas it declined while using other optimizers.the ROC plot depicts that this model is best suitable for identifying the disease diabetic retinopathy with an AUC of 0.9963.

According to experimental findings, ResNet-101 combined with the optimizer RMSprop is a suitable model for automatically identifying and classifying retinal lesions in FFA images using multi-label classification. This model forms the basis for automatic analysis of FFA images as well as thorough diagnosis and treatment decision-making for DR. It may help clinicians and the general population with widespread glaucoma screening and promptly, effectively, and affordably offer helpful advice on the diagnosis of glaucoma to the specialist.

**6. Reference :**

1.R. Sarki, K. Ahmed, H. Wang and Y. Zhang,et *al*. "Automatic Detection of Diabetic Eye Disease Through Deep Learning Using Fundus Images: A Survey," in *IEEE Access*, vol. 8, pp. 151133-151149, 2020, doi: 10.1109/ACCESS.2020.3015258.

2.Li, F., Yan, L., Wang, Y. *et al.* Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs. *Graefes Arch Clin Exp Ophthalmol* 258, 851–867 (2020). <https://doi.org/10.1007/s00417-020-04609-8>

3.Qureshi, I., Ma, J. & Abbas, Q. et *al.* Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. *Multimed Tools Appl* 80, 11691–11721 (2021). <https://doi.org/10.1007/s11042-020-10238-4>

4.Pan, X., Jin, K., Cao, J. *et al.* Multi-label classification of retinal lesions in diabetic retinopathy for automatic analysis of fundus fluorescein angiography based on deep learning. *Graefes Arch Clin Exp Ophthalmol* 258, 779–785 (2020). <https://doi.org/10.1007/s00417-019-04575-w>

5.Mule, D.B., Chowhan, S.S., Somwanshi, D.R. et *al.*(2019). Detection and Classification of Non-proliferative Diabetic Retinopathy Using Retinal Images. In: Santosh, K., Hegadi, R. (eds) Recent Trends in Image Processing and Pattern Recognition. RTIP2R 2018. Communications in Computer and Information Science, vol 1036. Springer, Singapore. <https://doi.org/10.1007/978-981-13-9184-2_28>

6.F. Alzami, Abdussalam, R. A. Megantara, A. Z. Fanani and Purwanto,et *al.* "Diabetic Retinopathy Grade Classification based on Fractal Analysis and Random Forest," 2019 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2019, pp. 272-276, doi: 10.1109/ISEMANTIC.2019.8884217.

7..L. Qiao, Y. Zhu and H. Zhou, et *al.* "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," in IEEE Access, vol. 8, pp. 104292-104302, 2020,

doi: 10.1109/ACCESS.2020.2993937.

8.Akbar, S., Hassan, S. A., Shoukat, A., Alyami, J., & Bahaj, S. A. et *al.*(2022). Detection of microscopic glaucoma through fundus images using deep transfer learning approach. *Microscopy Research and Technique*, 85( 6), 2259– 2276. <https://doi.org/10.1002/jemt.24083>

9.Singh, L. K., Khanna, M., & Thawkar, S.et *al* (2022). A novel hybrid robust architecture for automatic screening of glaucoma using fundus photos, built on feature selection and machine learning-nature driven computing. *Expert Systems*, 39( 10), e13069. <https://doi.org/10.1111/exsy.13069>

10.H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia and H. Hamam,et *al.* "Diabetic Retinopathy Diagnosis From Fundus Images Using Stacked Generalization of Deep Models," in *IEEE Access*, vol. 9, pp. 108276-108292, 2021, doi: 10.1109/ACCESS.2021.3101142.