Machine Learning Approach to Detect Annotate Eye Diseases using Retinal Images

Perera H. A. N. S.
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20166106@my.sliit.lk

Lakshith G. P. R.
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20165666@my.sliit.lk

Muthukumarana M. W. A. N. C.
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20227890@my.sliit.lk

Kariyawasam K. G. P. C.
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
it20172978@my.sliit.lk

Devanshi Ganegoda
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
devanshi.g@sliit.lk

Jeewaka Perera
Department of Computer Science & Software Engineering
Sri Lanka Institute of Information Technology
Malabe, Sri Lanka
jeewaka.p@sliit.lk

***Abstract*—This research addresses the escalating global health concerns of Diabetic Retinopathy (DR) and Age-Related Macular Degeneration (AMD), prevalent causes of visual impairment and blindness. The investigation pivots around developing advanced diagnostic tools through image processing and machine learning techniques, with a focus on their applicability in resource-limited settings in many regions of the globe. This research aims to develop deep learning models that accurately detect and categorize the severity of DR using retinal fundus images. These models offer the potential to facilitate early detection and enhance the effectiveness of treatments by improving diagnosis efficiency. Simultaneously, work is being carried out on automated diagnostic tools that utilize retinal Optical Coherence Tomography (OCT) images for the identification and classification of AMD's severity. Current methodologies are time-consuming and prone to observer variations, thus necessitating a more precise and automatic diagnostic approach. By the end of this research, a cross-platform mobile application is proposed, integrating the developed deep learning technology. This application is designed for the early detection and annotation of DR and AMD from retinal images, with the aim of revolutionizing the current clinical processes. It is anticipated that this research will expedite the diagnostic procedure, thereby contributing to the reduction of the global burden of blindness and visual impairment.**

***Index Terms*—Diabetic Retinopathy, Age Related Macular Degeneration, Eye Diseases, Retinal Fundus Images, Retinal OCT Images, EfficientNet, VGG16**

I. Introduction

Visual impairment and blindness are life-altering conditions affecting millions globally. Two predominant causes are Diabetic Retinopathy (DR) and Age-Related Macular Degeneration (AMD), constituting significant health problems, especially in regions like Sri Lanka. This research aims to tackle these eye diseases through advanced image processing and machine learning techniques, bridging the gap between high-tech healthcare and practical clinical applications.

Diabetic Retinopathy, the fourth leading cause of blindness and the fifth most common cause of visual impairment, is a notable complication of diabetes that affects the eye's retina. According to past research, one-third of Sri Lankan adults with self-reported diabetes have retinopathy. Despite the high prevalence, resources and expertise for screening and treating DR remain insufficient in the region, highlighting the urgent need for service expansion and mid-level HR training. More importantly, there is a call for advanced techniques to facilitate early detection and continuous monitoring of DR, ultimately reducing the disease's impact.

One of the primary components of this research focuses on developing deep learning models to detect DR using retinal fundus images. Additionally, our work involves the classification of the severity of DR based on the same images. These models will enable a more efficient and accurate diagnosis, increasing the odds of early detection and effective treatment.

Age-Related Macular Degeneration, another debilitating eye condition, causes progressive vision loss, impacting millions worldwide. Optical Coherence Tomography (OCT) is a vital tool in detecting and monitoring AMD. Despite its utility, the current manual process for AMD detection from OCT images is time-consuming and suffers from observer variations, potentially leading to misdiagnosis or delayed diagnosis.

This research also aims to develop automated diagnostic tools for detecting symptoms of AMD using retinal OCT images and classifying AMD's severity. By providing an automatic and accurate diagnostic approach, it can be helped to address the significant challenge to early detection and treatment of AMD. Current AMD classification methods, mostly developed for web applications, rely on multiple OCT images, making it challenging to distinguish between dry and wet AMD in the same patient's eye. By developing an efficient and precise method for diagnosing and classifying AMD, it can significantly enhance the clinical process.

In conclusion, this research proposes a cross-platform mobile application integrating deep learning technology for the early detection and annotation of DR and AMD using retinal images. The developed tools will streamline and expedite the diagnostic process, bringing us one step closer to reducing the burden of blindness and visual impairment globally.

II. Literature Review

There are several research papers which proposed a deep learning and image processing-based approach in the detection and classification of two common eye diseases, Diabetic Retinopathy and Age-Related Macular Degeneration. These research studies have used fundus images and Optical Coherence Tomography (OCT) images as the imaging modalities for DR and AMD respectively.

A. Detect Symptoms of Diabetic Retinopathy using Retinal Fundus Images

Akanksha Soni and Dr. Avinash Rai, "A Novel Approach for the Early Recognition of Diabetic Retinopathy using Machine Learning" proposed a deep learning approach using k-mean clustering, SVM and Random Forest classification algorithms [1]. The ocular image is pre-processed using a histogram equalization procedure and segmented using the k-means clustering algorithm to separate the usual and unusual areas of the ocular image. The segmented image regions are then classified using Support Vector Machine (SVM) and Random Forest. The dataset used in the study consisted of 89 color images, including 84 images with mild non proliferative signs of DR and 5 images without any DR symptoms. The model achieved recognition rates of 94.38% and 96.62% for SVM and Random Forest classifiers respectively.

Qomariah D U N et al., "Classification of Diabetic Retinopathy and Normal Retinal Images using CNN and SVM" in 2019 presented a deep learning approach to extracting features & classification using SVM CNN [2]. The proposed approach is tested on 77 retinal images from Messidor's base 12 and 70 retinal images from base 13 databases. Alexnet, VggNet, InceptionNet, GoogleNet, DenseNet, and Resnet were used in this study to gain the feature vector for classification. The extracted feature vectors are then input to a Support Vector Machine (SVM) classifier to find an optimal hyperplane to separate into classes - normal and Diabetic Retinopathy. The result of the experiment shows 95.83% & 95.24% as the highest accuracy values for base 12 & 13 respectively.

M Asiful Huda et al. in their research "An Improved Approach for Detection of Diabetic Retinopathy Using Feature Importance and Machine Learning Algorithms" proposed improved machine learning feature importance algorithms for detection of Diabetic Retinopathy [3]. Decision Tree, Logistic Regression, & SVM are used in their proposed system to improve the overall performance and robustness of the Diabetic Retinopathy detection system. The model achieved a precision of 97% and a recall of 92%, which shows a notable contrast with respect to existing results, 72% and 63% outcome.

B. Grade Severity of Diabetic Retinopathy using Retinal Fundus Images

Narayana et al. proposed an approach to detect Diabetic Retinopathy using Convolutional Neural Network (CNN) with the VGG-16 model. EyePACS dataset which consists of 35,126 images of both left and right eyes is utilized to train the model [4]. The system used normalization, centering, and cropping to 512 x 512 pixels to preprocess the retinal images. The research addresses the class imbalances in the dataset through data augmentation techniques. In this study, CNN architecture incorporates the VGG-16 model with trainable convolutional layers in Block-5 and an optimized dense layer to grade the severity of Diabetic Retinopathy. The study achieved an Average Class Accuracy (ACA) of 74%, sensitivity of 80%, and specificity of 65%, with an area under the curve (AUC) of 0.80.

A research study conducted by Jyostna proposed an optimized model for predicting the severity level of Diabetic Retinopathy (DR) [5]. This study used features extracted from pre-trained models. The researchers use activation filter values from convolution blocks 3, 4, and 5 of the VGG-16 model to acquire feature representation. Different pooling methods and fusion techniques were designed to represent retinal images. Kaggle APTOS 2019 contest dataset is used to train the proposed model. The proposed method achieved an accuracy of 84.31% and an AUC of 97. The outcomes outperformed existing models, particularly for severe and proliferate stage DR images.

C. Detect Symptoms of Age-related Macular Degeneration using Retinal OCT Images

In their study, "Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images", Lee et al. in 2018 investigated the effectiveness of deep-learning models for classifying normal and Age-related Macular Degeneration using OCT images [6]. The researchers use an optimized VGG16 convolutional neural network as the deep learning model. The model was adapted to suit the specific requirements of the classification task at hand. The obtained results demonstrated promising performance, with an ROC curve achieving 93.83% and an accuracy of 88.98%. These results indicate the effectiveness of the proposed deep-learning model for accurately classifying normal and AMD OCT images.

Srivastava et al. proposed a system on automated detection of age-related macular degeneration (AMD) using Optical Coherence Tomography (OCT) images [7]. Their research aimed at incorporating the choroid layer to improve the performance of the deep learning model in detecting AMD. The proposed system utilizes the ResNet50 deep learning model. The methodology involves two variations of the ResNet model. ResNet1 employed preprocessed and cropped images for both training and classification tasks. ResNet2, in addition to the preprocessed and cropped images, removed the choroid layer by referencing the location of the Bruch's Membrane (BM) as a reference point before performing training and classification. The ResNet1 model achieved accuracy of 96.78% and ResNet2 achieved accuracy of 95.82% in detecting age-related macular degeneration (AMD) from OCT images.

D. Classification of Age-related Macular Degeneration using Retinal OCT Images

Ali Serener et al. has proposed an automated approach for classifying dry and wet Age-related Macular Degeneration using deep learning Convolutional Neural Networks [8]. The study employed pre-trained AlexNet and ResNet models. The model is trained using ImageNet dataset which encompasses 8000 OCT images of four classes - healthy (normal), dry AMD, wet AMD, and diabetic macular edema (DME). The implementation of the proposed approach utilizes an optimized eighteen-layer ResNet architecture, trained using a GeForce GTX 1080 Ti GPU and the Caffe deep learning framework. The ResNet model achieves an area under the Receiver Operating Characteristic curve of 94% and 63% for dry AMD and wet AMD, respectively.

Govindaiah et al., in their study proposed a novel deep learning framework for automated screening to identify individuals at risk of developing Age-related Macular Degeneration (AMD) [9]. The proposed system utilizes the Age-Related Eye Disease Study (AREDS) dataset, composed of 150000 images graded by expert graders and ophthalmologists. Inception-ResNet-V2 and Xception deep neural networks were used to screen AMD. The authors identified two experiments for the proposed study. The first experiment categorized the images into two classes based on their clinical significance as none or early AMD and intermediate or advanced AMD. The second experiment categorized the images into four classes as no AMD, early AMD, intermediate AMD, and advanced AMD. Over 95.3% of accuracy is achieved for the first experiment and the second experiment demonstrates accuracy of 86%.

Sertkaya et al., in their study, use various CNN models, LeNet, AlexNet, and Vgg16 architectures for the diagnosis of Neovascularization, Diabetic Macular Edema, Drusen, and healthy eye conditions using OCT images [10]. The study shows successful results, especially with Vgg16 and AlexNet architectures in classifying Age-related Macular Degeneration (AMD) stages. The dropout layer structure in AlexNet model minimized the training loss. The VGG-16 architecture achieved a classification result of 93.01%.

III. Methodology

The proposed solution intends to provide a deep learning approach, to detect and classify Diabetic Retinopathy and Age-related Macular Degeneration. As depicted in Figure 1, the users need to register to the system. If the user is already registered to the system, the user can login to the system. The user can select the image type, fundus image or OCT image from the options provided by the system. The users can capture or upload the retinal image image of the patient. The main purpose of these components is to screen and classify for eye disease when the image is uploaded. To reach this goal, deep neural networks have been used. This proposed system detects DR and AMD and then classifies according to their severity by utilizing distinct models.

A. Detect Symptoms of Diabetic Retinopathy using Retinal Fundus Images

a) Data Collection and Preprocessing: For the proposed research study, the retinal fundus image dataset is fetched from open-source image data repository, Asia Pacific Tele-Ophthalmology Society 2019 Blind Detection (APTOS 2019 BD). The dataset consists of 3662 samples of high-resolution fundus images, classified into five classes (no DR, mild DR, moderate DR, severe DR, and proliferative DR). The dataset comprises three sets, training, validation and test. The training set is used to train the deep learning model of the proposed system. To ensure the dataset has the optimal data quality for the model, preprocessing is carried out prior to training the deep learning algorithm. Initially, the dimension of the entire dataset is resized to a standardized resolution of 224*224 pixels. Secondly, pixel values of retinal images are normalized to a range between 0 and 1 to enhance gradient-based optimization during model training. To enhance the quality and reduce noise in the retinal fundus images, contrast enhancement, noise reduction, and histogram equalization techniques are utilized. Data augmentation, which is a crucial tactic to increase the diversity of data accessible for training models, without actually gathering new data in the model. In the proposed research, random rotations, flips, and translations techniques are utilized to preprocess the dataset.

b) Model Architecture Design: The proposed system utilizes TensorFlow and Keras frameworks as the deep learning model architecture. The model is based on EfficientNetB3 which is a pre-trained Convolutional Neural Network (CNN). The pre-trained weights enable the model to efficiently extract high-level features from images. The top classification layers of EfficientNet are not included because they are tailored to the ImageNet dataset. It provides a better trade-off between computational efficiency and performance. Since EfficientNetB3 pre-trained on the ImageNet dataset, it optimizes the training process and model performance. More importantly, EfficientNet is utilized in the proposed system as the performance of the model makes it more suitable for mobile applications.

Additional layers are added on top of the EfficentNetB3 base model to enhance the capacity to learn complex features from retinal fundus images of the proposed model. In order to improve convergence during training, Batchnormalization is

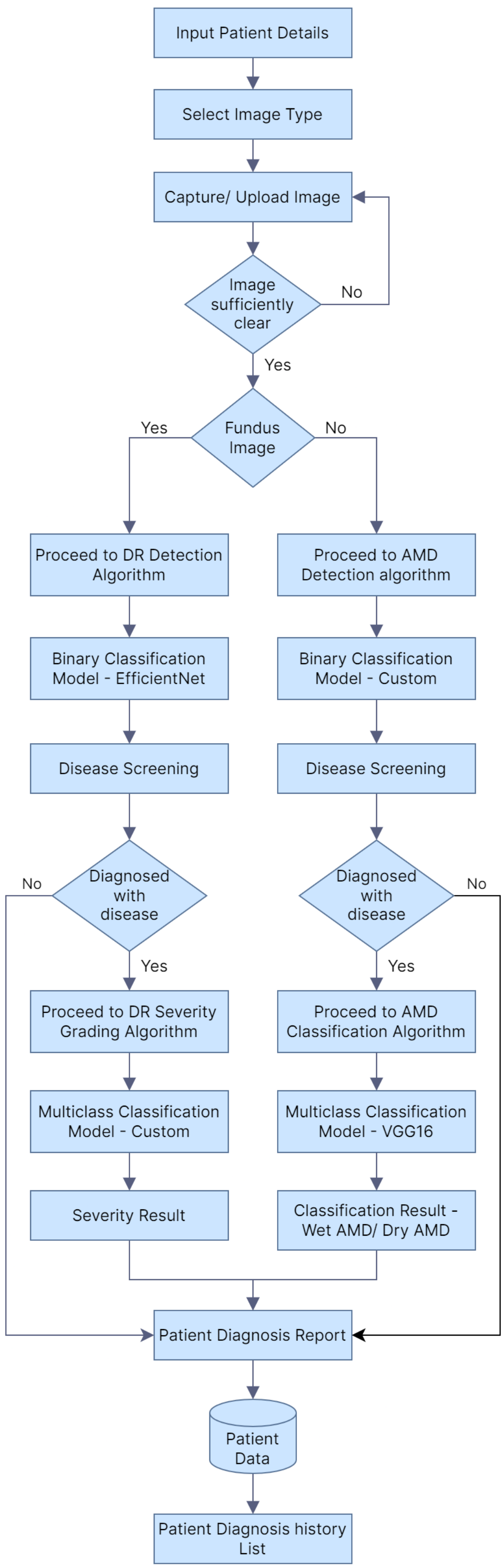

Fig. 1. Overall System Diagram

used to normalize the activations of the preceding layer. The bottleneck layer, a Dense layer with 256 units, is introduced and extracts key features from the output of the base model. To avoid overfitting and encourage the model to learn more reliable representations, regularization techniques, such as L1 and L2 regularization, are incorporated into the Dense layer. A Dropout layer with a rate of 0.45 is used to intensify over fitting by randomly setting a fraction of input units to 0 during training. The output layer is then added and contains a Dense layer with a softmax activation function, which generates predictions for the detection of Diabetic Retinopathy.

c) Model Training and Optimization: The proposed model is fed with preprocessed retinal fundus images during the training phase. To maximize the performance of the model, its parameters are iteratively changed to minimize the loss function. With a learning rate of 0.001, the Adamax optimizer is used to effectively update the weights and biases of the model. The model is trained to accurately detect symptoms of Diabetic Retinopathy from retinal fundus images by iteratively adjusting the modelś parameters, optimizing the loss function, and utilizing regularization techniques.

d) Model Evaluation: The trained model is tested to detect the symptoms of diabetic retinopathy. The evaluation is conducted on separate validation and test datasets, which were not used during the training phase. To gauge the overall effectiveness of the model, the accuracy matrix is computed. Additional evaluation metrics, such as precision, recall, and F1-score are utilized to provide a thorough evaluation of the capacity of the proposed model. To assess the precision of the diagnosis, the model's predictions are contrasted with the ground truth labels

B. Grade Severity of Diabetic Retinopathy using Retinal Fundus Images

a) Data Collection and Preprocessing: For this research study, the retinal fundus image dataset was sourced from Asia Pacific Tele-Ophthalmology Society 2019 Blind Detection (APTOS 2019 BD) too. The dataset was split into two sets: training and validation, with the split set at 80% for training and 20% for validation. The split was determined using the ImageDataGenerator class from TensorFlow with a rescale factor of 1/255 to normalize pixel values, enhancing the model's training process.

To adapt the dataset to the required model input, the images were resized to a standard resolution of 48x48 pixels, maintaining a grayscale color mode to retain essential features while reducing computational complexity.

b) Model Architecture Design: The deep learning model for this research was meticulously designed using TensorFlow and Keras, two powerful open-source platforms that are widely used for machine learning and deep learning tasks. The model's architecture was constructed with a blend of simplicity and efficiency, making it adaptable and potent in tackling the problem at hand. This architecture included a set of two convolutional layers, each trailed by a max pooling layer, a dropout layer for regularization, a flattened layer to restructure the data, and ultimately a dense layer that serves as the output layer.

The Convolutional layers were pivotal in the architecture for feature extraction, which is crucial in image processing tasks. These layers were designed with a kernel size of 3x3, a typical and effective size that balances the trade-off between capturing sufficient image details and limiting computational cost. Convolutional layers are capable of identifying local features in images, like edges and shapes, by sliding over the image with the defined kernel and computing dot products. They translate the raw pixel data into a more meaningful representation without any loss of volume dimensions.

Following the convolutional layers, MaxPooling layers were introduced. The objective of these layers is to downscale the representation and reduce its dimensions. This helps in decreasing the computational load, making the model more efficient. Additionally, max pooling serves the purpose of providing a form of translational invariance, which essentially means the model becomes more robust to the position of features in the image, thereby controlling overfitting. MaxPooling operates by sliding a window across the input and choosing the maximum value within that window as the representative, effectively reducing the spatial dimensions.

A Dropout layer was incorporated after the max pooling layers as a regularization technique. Overfitting, a situation where the model performs exceptionally well on training data but poorly on unseen data, is a common issue in machine learning. To mitigate this, the dropout layer randomly turns off a fraction of neurons during each training update, precisely 20% in this case. By doing this, it forces the model to learn more robust and generalized features since it can no longer rely on any one feature or combination of features, thus enhancing the model's generalization capabilities.

Following the dropout layer, the model architecture included a Flatten layer. This layer simply transforms the 2D matrix (height and width, depth) output from the previous layer into a 1D array or vector. The transformation is crucial because the dense layer that follows only accepts input in vector form.

The final piece of the architecture was the Dense layer. This layer is fully connected, meaning each neuron in the dense layer is connected to every neuron in the previous layer. It was designed with five units, each corresponding to one of the five severity classes of diabetic retinopathy. A softmax activation function was applied to this layer to enable it to output a probability distribution over the five classes. The class with the highest probability is considered as the model's prediction.

In summary, the model architecture has been thoughtfully designed with layers and techniques to effectively perform the task of diagnosing the severity of diabetic retinopathy. Each layer plays a specific role in transforming the input images into a form that the model can use to learn and make accurate predictions.

c) Model Training and Optimization: The model was compiled with the Adam optimizer, known for its efficient computation, little memory requirement, and suitability for problems with large amounts of data and parameters. A sparse categorical cross-entropy loss function was used given the multi-class nature of the problem.

The model was then trained on the preprocessed dataset for five epochs. The model parameters were adjusted iteratively to minimize the loss function, optimizing the model's capability to accurately detect symptoms of diabetic retinopathy from the retinal fundus images.

d) Model Evaluation: The trained model's performance was assessed utilizing a separate validation set, distinct from the data used in the training phase. The model's evaluation was executed using the built-in function 'model.evaluate', which computed the loss and accuracy values for the model on the validation set.

The 'model.evaluate' function inherently accounts for the class imbalances, if any, and delivers a holistic measure of the model's performance. By providing the loss and accuracy metrics, it offers an overall quantification of the extent to which the model's predictions conform to the actual labels in the validation set. The loss value signifies how well the model managed to optimize its parameters during training, while the accuracy score illustrates the percentage of the validation set images the model classified correctly.

Moreover, the development of loss and accuracy values over the course of training epochs was plotted for both the training and validation sets. This visual representation served as an effective tool for understanding the model's learning process, indicating any signs of underfitting or overfitting. The plot also provided insights into how the model's performance improved or deteriorated over time, which could influence decisions for potential adjustments in the model's architecture or training process. For a more granular evaluation of the model's performance, a confusion matrix can be generated. This statistical tool provides a comprehensive view of the model's performance, showing the number of correct and incorrect predictions made for each class. It's a powerful method to understand not only where the model is making correct predictions, but also which classes are getting confused with each other.

Finally, once the training process was concluded and the performance deemed satisfactory, the trained model was saved into an HDF5 file using the 'model.save' function. This saved model file comprises all necessary information about the model, including its architecture, learned parameters (weights and biases), the loss function, the optimizer, and its current state. This enables not only the recreation of the exact model later but also the possibility to resume training from where it was left off, providing flexibility and efficiency in the model's utilization.

C. Detect Symptoms of Age-related Macular Degeneration using Retinal OCT Images

a) Data Collection and Preprocessing: The dataset, obtained from Kaggle, consisted of 15,900 OCT images, with 7,950 images classified as AMD and 7,950 as normal. The images were resized into 224x224 pixels, then made grayscale, and had their pixel values normalized to a range of 0 to 1 to preprocess the data. We obtained a balanced representation by choosing 2,300 samples per class for the training set, 300 samples per class for the validation set, and 500 samples per class for the test set. Mentioned samples came from the related class directories and were chosen at random. To enhance model generalization, we utilized the ImageDataGenerator from the Keras library for data augmentation. This involved applying techniques such as rotation, shifts, zooming, and flipping to the training data. The ImageDataGenerator was fitted to the training data, enabling real-time augmentation during model training. During the training process, we used a batch size of 32. We generated augmented training and validation data using the fitted ImageDataGenerator, and for the test data, no augmentation was applied.

b) Model Architecture Design: Convolutional neural network (CNN) was utilized to detect symptoms of age-related macular degeneration (AMD) in retinal OCT images. The CNN architecture consisted of multiple layers: three 2D convolutional layers with increasing filter sizes (32, 64, and 128), each followed by max-pooling layers to reduce spatial dimensions. The feature maps were flattened and fed into a fully connected layer with 128 neurons and ReLU activation. To prevent overfitting, a dropout layer with a dropout rate of 0.5 was employed. The output layer had a single neuron with sigmoid activation for binary classification. The model was compiled with the Adam optimizer and binary cross-entropy loss, and accuracy was used as the evaluation metric. This model design aimed to effectively capture relevant patterns and features in retinal OCT images to accurately detect AMD symptoms.

c) Model Training and Optimization: We utilized a deep learning approach to train a convolutional neural network (CNN) on a dataset of preprocessed and augmented retinal OCT images. The dataset was divided into training and validation sets, with an 80:20 ratio. The model was trained for 20 epochs using the Adam optimizer and specified batch size. We monitored the model's performance using training and validation loss, as well as metrics like accuracy. The training process involved adjusting the model's architecture and hyperparameters to optimize performance. By iteratively refining the model, we aimed to create an accurate and reliable system for detecting AMD symptoms from retinal OCT images.

d) Model Evaluation: In the dataset, the test set was prepared by reshaping the images to match the required input dimensions of 224x224 pixels with a single channel. The model was then evaluated and calculated the test loss and test accuracy. The test loss represented the overall error of the model on the test set, while the test accuracy measured the percentage of correctly classified images. These metrics provided quantitative insights into the model's effectiveness in identifying AMD symptoms based on retinal OCT images. The test loss and test accuracy served as crucial indicators to gauge the model's performance and its suitability for detecting AMD symptoms.

D. Classification of Age-related Macular Degeneration using Retinal OCT Images

a) Data Collection and Preprocessing: The research employed a dataset of optical coherence tomography scans categorized into three distinct classes choroidal neovascularization (CNV), Drusen, and Normal retina. A total of 72136 OCT scans were acquired, comprising 37205 scans indicative of CNV, 8616 scans indicative of Drusen, and 26315 scans indicative of a normal retina.

The data were subjected to various preprocessing procedures to standardize their format and enhance quality. Accordingly, images were adjusted to a uniform size of 150x150 pixels to maintain consistency, and the RGB color space was implemented to acquire the chromatic data within the images. The normalization process was implemented to mitigate the influence of pixel intensity fluctuations on the model's learning ability, thereby enhancing its effectiveness in data analysis.

The Keras ImageDataGenerator was utilized to implement data augmentation techniques to enhance the dataset's diversity and resilience. The methods employed encompassed zooming, flipping, rotation, shifting, brightness modification, and shear. Implementing these augmentations in real-time during training broadens the model's exposure to various model variations, resulting in improved generalization and performance.

b) Model Architecture Design: The base model selected for this study was the VGG16 architecture, a deep CNN pre-trained on the ImageNet dataset. The utilization of pre-existing weights of the VGG16 model was employed to capitalize on its robust feature extraction capabilities. The VGG16 model's fully connected layers were eliminated, and a novel classification layer structure was explicitly developed to categorize Dry and Wet AMD.

The modified architecture consisted of a flattening layer to transform the extracted features into a 1-dimensional vector. Subsequently, a fully connected layer comprising 256 units and a Rectified Linear Unit (ReLU) activation function was employed, facilitating non-linear transformations and capturing intricate patterns in the dataset. A dropout layer was introduced to address the overfitting issue with a dropout rate 0.5. This layer randomly deactivates a portion of the units during training, thereby reducing the risk of the model becoming overly dependent on particular features. A dense layer activated by the softmax function, consisting of three units, was utilized to represent the three distinct categories: CNV, Drusen, and Normal retina.

c) Model Training and Optimization: The Adam optimizer was employed to train the model, owing to its suitability for deep neural network training. The learning rate was established at 0.0001 to ensure stable convergence during training. A categorical cross-entropy loss function was deemed appropriate for multi-class classification tasks.

The class weights were computed using the reciprocal of the class frequencies to rectify the class imbalance problem in the dataset. The approach used in this study involved assigning greater weights to the underrepresented classes, namely CNV and Drusen while assigning comparatively lower weights to the overrepresented class, Normal retina. During the training process, the loss function was adjusted to include the class weights, which resulted in the model assigning equal significance to all classes and effectively acquiring knowledge from the underrepresented classes.

Two methodologies were implemented to mitigate the issue of overfitting and improve the overall generalization of the model. Initially, the technique of early stopping was employed, whereby the validation loss was closely monitored, and the training process was terminated if the loss did not improve for a specific number of epochs. This prevented the model from continuing to train when no further improvement was observed. Furthermore, a callback for reducing the learning rate was implemented, which decreased the learning rate if the validation loss failed to fall within a predetermined number of epochs. Implementing an adaptive learning rate adjustment strategy facilitated the convergence of the model toward an improved solution.

The model's training process was conducted for 20 epochs, utilizing a batch size 64. During the training process, the training dataset underwent on-the-fly augmentation using the ImageDataGenerator. This resulted in the generation of augmented images created using pre-defined augmentation techniques. It incorporated augmented images furnished supplementary variations to the model, facilitating the acquisition of resilient features and enhancing its capacity to extrapolate to novel data.

d) Model Evaluation: The model's performance was evaluated on an independent testing dataset not utilized during the training or validation phases. The sci-kit-learn library was used to compute evaluation metrics: accuracy, precision, recall, and F1-score. The abovementioned metrics have furnished valuable insights regarding the model's all-around performance and proficiency in accurately classifying the distinct AMD categories.

Additionally, using the AUC-ROC curve comprehensively analyzed the model's discriminatory power across various classification thresholds. A higher Area Under the Curve (AUC) value indicates superior class separability and more robust classification performance.

e) Comparative Analysis and Interpretation: A comparative analysis was performed to evaluate the efficacy of the proposed model in the domain of Dry and Wet AMD classification to established state-of-the-art methods or benchmarks. The assessment centered on each class's classification accuracy, precision, recall, and F1 score. The inherent merits and demerits of the model could be discerned through comparative analysis of the proposed model's performance relative to alternative approaches.

Qualitative analysis techniques were implemented to acquire additional comprehension regarding the model's decision-making mechanism and construe its categorizations. Activation maps were generated through methods such as Grad-CAM, which effectively identified the salient areas within the OCT images that played a crucial role in determining the classification outcomes of the model. The interpretability and trustworthiness of the model can be improved by better understanding its reasoning and feature selection mechanisms through the visualization of these crucial regions.

IV. RESULTS AND DISCUSSIONS

Figure 6 illustrates the overview of the model performance metrics for the complete system.

Component of the System	Accuracy	Loss
Detect DR	97.4%	4.52%
Grade Severity of DR	71.06%	7.76%
Detect AMD	99.1%	1.61%
Classification of AMD	95.42%	13.05%

A. Detect Symptoms of Diabetic Retinopathy using Retinal Fundus Images

The EfficientNet model was used in classifying Diabetic Retinopathy from retinal fundus images. The reason behind selecting this model was its proven success in image classification tasks. EfficientNet models have demonstrated outstanding performance with relatively fewer parameters, making them suitable for tasks where computational resources are limited.

The model structure consisted of multiple convolutional layers, leveraging the efficient architecture to learn hierarchical features from retinal fundus images. It also incorporated various pooling and regularization layers, ensuring the model's robustness and generalization capabilities.

After training the EfficientNetB3 model on our dataset, we achieved an accuracy of approximately 97.4%. This indicates the effectiveness of the proposed model in accurately diagnosing Diabetic Retinopathy using retinal fundus images.

B. Grade Severity of Diabetic Retinopathy using Retinal Fundus Images

A custom model was developed to grade the severity of Diabetic Retinopathy since the need for a fine-grained classification.

The custom model architecture included two convolutional layers, each followed by a max pooling layer to downsample the spatial dimensions. A dropout layer was incorporated to prevent overfitting, and a flattened layer was used to reshape the data for feeding into the subsequent dense layer for classification.

Upon training the custom model on a labeled dataset of retinal fundus images, the model achieved an accuracy of approximately 71% in grading the severity of Diabetic Retinopathy, showcasing its ability to accurately classify the different stages of the disease based on retinal fundus images.

C. Detect Symptoms of Age-related Macular Degeneration using Retinal OCT Images

For the detection of Age-related Macular Degeneration symptoms, a custom model was developed which was designed specifically for retinal OCT images.

The custom model for OCT images comprised three 2D convolutional layers with increasing filter sizes (32, 64, and 128) to capture different levels of details. Each convolutional layer was followed by max-pooling layers, reducing spatial dimensions and retaining the most salient features.

During the training process on a labeled dataset of retinal OCT images, the custom model achieved an accuracy of approximately 99%. This demonstrates the effectiveness of the CNN architecture in capturing relevant patterns and features for accurate detection of AMD symptoms.

D. Classification of Age-related Macular Degeneration using Retinal OCT Images

The proposed model employed the VGG16 model for the classification of Age-related Macular Degeneration using retinal OCT images. VGG16 is a widely adopted architecture with a deep stack of convolutional layers, making it suitable for capturing intricate features from complex images.

The modified architecture removed the fully connected layers of VGG16 and introduced a novel classification layer structure. It included a flattening layer, a fully connected layer with 256 units and ReLU activation, and a dropout layer with a rate of 0.5 for regularization. The output layer consisted of three units activated by the softmax function for multi-class classification.

After training VGG16 on the OCT image dataset, we obtained an accuracy of approximately 95.42%. This showcases the efficacy of the VGG16-based architecture for accurate classification of AMD based on retinal OCT images

V. Conclusion and Future Works

In this research, we addressed the pressing issue of visual impairment and blindness caused by Diabetic Retinopathy (DR) and Age-Related Macular Degeneration (AMD). Our goal was to leverage advanced image processing and machine learning techniques to develop efficient and accurate diagnostic tools for early detection and continuous monitoring of DR and AMD. Our proposed cross-platform mobile application, integrating these advanced technologies, has the potential to significantly impact clinical practices, especially in regions where resources are scarce. The research focused on developing deep learning models for detecting DR and classifying the severity of DR using retinal fundus images. By harnessing deep learning, we aimed to enhance the chances of early detection and timely treatment of DR. Additionally, we developed automated diagnostic tools to detect symptoms of AMD and classify its severity. By automating the diagnostic process, we aimed to reduce misdiagnosis and delays in patient care.

As the future work, data augmentation and transfer learning techniques can be used to enhance the performance and generalization capabilities of the deep learning models. Also, integrating clinical data such as patient history and other medical parameters, to improve the diagnostic accuracy and personalize treatment plans for individual patients.

References

[1] A. Soni and A. Rai, "A Novel Approach for the Early Recognition of Diabetic Retinopathy using Machine Learning," in 2021 International Conference on Computer Communication and Informatics, ICCCI 2021, Institute of Electrical and Electronics Engineers Inc., Jan. 2021. doi: 10.1109/ICCCI50826.2021.9402566.

[2] D. U. N. Qomariah, H. Tjandrasa, and C. Fatichah, "Classification of Diabetic Retinopathy and Normal Retinal Images using CNN and SVM," 2019 12th International Conference on Information & Communication Technology and System (ICTS), Jul. 2019, doi: https://doi.org/10.1109/icts.2019.8850940.

[3] S. M. A. Huda, I. J. Ila, S. Sarder, Md. Shamsujjoha, and Md. N. Y. Ali, "An Improved Approach for Detection of Diabetic Retinopathy Using Feature Importance and Machine Learning Algorithms," IEEE Xplore, Jun. 01, 2019. https://ieeexplore.ieee.org/abstract/document/8843676 (accessed Mar. 25, 2022).

[4] Narayana Bhagirath Thota and Doshna Umma Reddy, "Improving the Accuracy of Diabetic Retinopathy Severity Classification with Transfer Learning," Aug. 2020, doi: https://doi.org/10.1109/mwscas48704.2020.9184473.

[5] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Deep convolution feature aggregation: an application to diabetic retinopathy severity level prediction," Signal, Image and Video Processing, vol. 15, no. 5, pp. 923–930, Jan. 2021, doi: https://doi.org/10.1007/s11760-020-01816-y.

[6] C. S. Lee, D. M. Baughman, and A. Y. Lee, "Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images," Ophthalmology Retina, vol. 1, no. 4, pp. 322–327, Jul. 2017, doi: https://doi.org/10.1016/j.oret.2016.12.009.

[7] R. Srivastava, Ee Ping Ong, and Beng Hai Lee, "Role of the Choroid in Automated Age-related Macular Degeneration Detection from Optical Coherence Tomography Images," Jul. 2020, doi: https://doi.org/10.1109/embc44109.2020.9175809.

[8] A. Serener and S. Serte, "Dry and Wet Age-Related Macular Degeneration Classification Using OCT Images and Deep Learning," IEEE Xplore, Apr. 01, 2019. https://ieeexplore.ieee.org/document/8741768?arnumber=8741768

[9] A. Govindaiah, R. T. Smith, and A. Bhuiyan, "A New and Improved Method for Automated Screening of Age-Related Macular Degeneration Using Ensemble Deep Neural Networks," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2018, doi: https://doi.org/10.1109/embc.2018.8512379.

[10] Mehmet Sertkaya, B. Ergen, and Mesut Toğaçar, "Diagnosis of Eye Retinal Diseases Based on Convolutional Neural Networks Using Optical Coherence Images," 2019 23rd International Conference Electronics, Jun. 2019, doi: https://doi.org/10.1109/electronics.2019.8765579.