Institute of Research and Publication Indonesia (IRPI)
Public Research Journal of Engineering, Data Technology and Computer Science
Journal Homepage: https://journal.irpi.or.id/index.php/predatecs
Vol. 3 Iss. 1 July 2025, pp: 42-58
ISSN(P): 3024-921X | ISSN(E): 3024-8043

Analysis Comparison Classification Image Disease Eye Using the CNN Algorithm, Inception V3, DenseNet 121 and MobileNet V2 Architecture Models

Nasya Amirah Melyani[1*], Ayuni Fachrunisa Lubis[2], Aghnia Tatamara[3], Ryando Rama Haiban[4], Muhammad Iltizam[5], Muhammad Aufi Rofiqi[6], Sakhi Hasan Abdurrahman[7], Nitasnim Samae[8], Bilal Shahid[9], Muhammad Habibullah[10], Muhammad Ibrara Ismail[11]

[1,2]Department of Information System, Faculty of Science and Technology,
Universitas Islam Negeri Sultan Syarif Kasim Riau, Indonesia
[3]İşletme Çalışmaları Programı, Iktisadi Ve Idari Bilimler Fakültesi, Dokuz Eylül Üniversitesi, Turkey
[4]Department of Ushuluddin, Faculty of Sharia, Yarmouk University, Jordan
[5]Sarjana Pendidikan, Fakultas Pendidikan, Universiti Pendidikan Sultan Idris, Malaysia
[6]Department of Hadith Science, Faculty of Ushuluddin, Al-Azhar University, Egypt
[7]Department of Tafsir and Science of the Koran, Faculty of Ushuluddin, Al-Azhar University, Egypt
[8]Department of Teaching Islamic Studies, Faculty of Islamic Science,
Prince of Songkla University, Thailand
[9]Department of Sales and Marketing, Faculty of Management Sciences,
Shaheed Zulfikar Ali Bhutto Institute of Science and Technology, Pakistan
[10]Department of Fiqh and Usul, Faculty of Syariah wal Qonun, Al-Wasatiyah University, Yamen
[11]Department of Fiqh Tahawwulat, Faculty of Ushuluddin, Al-Wasatiyah University, Yamen

E-Mail : [1]12150323859@students.uin-suska.ac.id, [2]12150322141@students.uin-suska.ac.id,
[3]aghniiatatamara04@gmail.com, [4]ryandohaibaan@gmail.com, [5]ijampetra@gmail.com,
[6]Aufirofiqi@gmail.com, [7]udelgateltv@gmail.com, [8]nitasnimsamae12@gmail.com,
[9]bilalshahid1226@gmail.com, [10]mhdhabibullah0406@gmail.com, [11]alwasatiah.indo@gmail.com

Received Jul 17th 2024; Revised Mar 26th 2025; Accepted Apr 27th 2025; Available Online Jul 06th 2025, Published Jul 31th 2025
Corresponding Author: Nasya Amirah Melyani

Abstract

Eye disease is a significant global health problem, with more than two billion people experiencing vision impairment. Some of the main causes of visual impairment include cataracts, glaucoma, diabetic retinopathy, and age-related macular degeneration. Early detection of eye disease is very important to prevent blindness. The fundus of the eye, which includes the retina and blood vessels, is an important area in the diagnosis of retinal diseases. Fundus disease can cause significant vision loss and is one of the leading causes of blindness. Automated analysis of fundus images is used to diagnose common retinal diseases, ranging from easily treaTable to very complex conditions. This research discusses eye disease image classification using several Convolutional Neural Network (CNN) architectures, namely Inception V3, DenseNet 121, and MobileNet V2. The dataset used is 4217 fundus images categorized based on the patient's health condition. Data is processed through normalization and augmentation to improve model performance. Experimental results show that MobileNet V2 has the highest accuracy of 81.3%, followed by Inception V3 with 77.3%, and DenseNet 121 with 76.7%. The use of appropriate CNN models in the classification of eye fundus images can help in early detection of eye diseases, thereby preventing further visual impairment.

Keyword: Cataract, Convolutional Neural Network, Diabetic Retinophaty, Eye Fundus, Glaukoma

1. INTRODUCTION

The eye is a sense organ that transmits visual stimuli to the brain. The World Health Organization (WHO) reports that more than two billion people worldwide suffer from near vision impairment far and one billion people experience disturbance in vision eye [1]. In the year 2013, 64.3 million people between the ages of 40 and 80 suffered from glaucoma, and estimates suggest that this number will increase to 76 million in 2020, then to 111.8 million in 2040. The World Health Organization highlights the importance of early detection of eye disease to effectively prevent and treat visual impairment [2]. The main causes of visual impairment include a variety of conditions, including Cataracts, Glaucoma, Diabetic retinopathy, AMD, and Myopia [3]. In addition to these eye fundus diseases, there are normal classes.

The normal class of fundus disease refers to a state where the back of the eye, including the retina, blood vessels, and optic nerve, shows no signs of abnormality or disease. On a normal fundus examination, the retina will appear healthy with consistent color, no bleeding, exudates, or other pathological changes. Blood vessels will be clear and regular, and the optic disc (where the optic nerve enters the eye) will have clear borders and normal color. This examination is important for detecting various eye conditions such as diabetic retinopathy, macular degeneration, and glaucoma at an early stage [4].

The fundus is the back portion of the eye, encompassing the retina, choroid, photoreceptor cells, blood vessels, and nerve optics [5]. Disease fundus can cause lost vision which is the reason for major blindness [6]. Analysis picture fundus automatically used as a tool for diagnosis of disease retina which general [7]. Retinal diseases range from fairly common and easily treaTable diseases to quite rare diseases and complex [8]. The condition of blood vessels on the retina is an important factor used to diagnose the presence of several eye diseases [9]. Chronic Eye Diseases (COD) like myopia, diabetic retinopathy, age-related macular degeneration, glaucoma, and cataracts can impact the eyes and potentially lead to severe visual impairment or blindness [10].

Diabetic Retinopathy (DR) is a disease with retina secondary consequences of damaged capillary retina caused by impaired glucose tolerance, and is known as an inflammatory neurovascular complication with neurological disorders/ dysfunction, which finally very influences vision and even results in blindness [11]. DR is marked by disturbance vessels blood progressive in the retina and it is caused by hyperglycemia and can happen in all patients with diabetes, regardless of the level of its severity [12]. Retinopathy diabetes until currently has been tested manually by an ophthalmologist. Manual diagnosis of DR is time-consuming [13]. The disease remains largely undetected until the late stages of the disease. Therefore, early detection is necessary to prevent vision loss [14]. It is mainly pathologically characterized by changes vascular retina [15]. Initially, the symptoms are not seen, that's why an inspection routine and early detection by a specialist are very important. In addition, retinopathy accelerates the onset of other eye diseases such as glaucoma or cataracts [16].

Cataract is turbidity lenticular which obscures the lens transparent on the eye man. Cataracts can occur consequence of a hydration lens (additional fluid) or denaturation proteins on the lens. Detection of early cataracts plays an important role in treatment and in a way significantly can reduce the risk of blindness [17]. The impact which is given by cataracts can influence the productivity and mobility of the sufferer which impacts the decrease in people's quality of life [4]. Cataracts can be anticipated by carrying out early detection when the eyes begin to experience problems [18].

Glaucoma (GL) is an eye disease that damages the optic nerve that connects the eye to the brain [19]. Glaucoma is a disease eye triggered by increasing pressure fluid on the eye thereby damaging the optic nerve and causing partial or complete loss of vision [20]. Glaucoma is an asymptomatic condition, and patients do not require medical attention until it is in an advanced stage, so diagnosis is often too late to prevent blindness [21]. Glaucoma is currently the leading retinal disease, which damages eye because pressure intraocular (IOP) on the eye [22]. In the disease glaucoma, the optic disc enlarges from its original size due to an abnormal condition in the eye. Signals resulting from vision are sent to the brain via the optic nerve. The region of the nerve that does not have rods and cones is called the blind spot [19]. This disease begins with an irregular behavior of the eye's drainage flow which ultimately causes an increase in intraocular pressure, which in severe stages of the disease worsens the condition of the optic nerve head and causes vision loss [23]. Glaucoma progressively damages the optic nerve with degeneration of the nerve fibers, leading to visual impairment leading to blindness [24]. Existing glaucoma detection processes have not consistently provided satisfactory results. Manual glaucoma detection (i.e. assessment of the optic nerve head) is subjective, expensive, time consuming, and variability in physician performance is very high. the aforementioned automatic glaucoma detection using fundus images [25]. Diabetic Eye Disease (DED) comprises a group of eye conditions, which include Diabetic Retinopathy, Diabetic Macular Edema, Glaucoma, and Cataracts. Advances in Artificial Intelligence (AI) offer lots of profit in the detection of DED automatically compared to the approach manually. Matter This includes reduced human error, time efficiency, and easier detection of minor abnormalities. DED detection system can assembled through technique processing picture combined use of technique Machine Learning (ML) or Deep Learning (DL) [26].

Artificial Intelligence is a technology that is currently developing rapidly. Almost all midwives have to utilize technology, starting from the field industry, banking, health, etc. AI or what is usually called artificial computing is a science that has many branches. One of the branches which famous is Machine Learning (ML). ML is an algorithm deep mathematics system computer that adopts learning machine use data And do predictions in the future. ML has been very developed and has been used, including in the health sector [27].

The new branch of ML is Deep Learning (DL) which contains more complex data. DL is an architecture which more complex and uses more lots amount layers, so in its application, it is hoped that it can handle complex problems and more data [28]. With its ability to extract features important from data without intervention man which significantly, DL has demonstrated outstanding performance in various fields, including image recognition, natural language processing, and medical data analysis. This makes DL a very promising technology in supporting and improving AI-based diagnostic capabilities, including early detection and classification of eye diseases [29].

One of the DL algorithms that has good performance in the field of image classification is Convolutional Neural Network (CNN). CNN is basically an arrangement of many layers consisting of convolution layers, pooling layers, and fully connected layers [30]. CNNs are specifically designed to process data in the form of a grid like an image, by automating the extraction of important features from the image. In the convolution layer, filters are applied to detect local patterns such as edges, textures, or specific objects in the image [31]. Pooling layers function to reduce data dimensions and minimize overfitting by summarizing important information. The fully connected layer then integrates the extracted features to perform classification or prediction [32].

In the context of image classification, CNNs are very effective due to their ability to automatically detect disease [33]. Use of CNNs in the medical categorization of,nervous systems. This diagnosis increases accuracy detection, and speeds up process analysis. For example, CNN can trained with datasets big for automatic detection minimize optical interference [34]. Once trained, a CNN can automatically analyze new images and provide an accurate diagnosis [35].

In study this uses the method CNN because its ability to detect eye disease accurately [36]. CNN is often used because the algorithm is able to load the whole information with various type of scale and can classify pictures with more accuracy [35]. CNN has lots architecture in processing data picture which own amount layers which different and the level of depth is not the same for each architecture [37]. We will use three models, namely Inception V3, DenseNet 121, and MobileNet V2 to classify eye disease images. We will compare results from third model the and determine where which better in classify images of eye diseases.

CNN have various architectures designed to handle diverse tasks in image recognition and classification [32]. Each CNN architecture offers specific advantages in accordance with type data and problem certain [38]. In context classification of disease eye using imagery fundus eye, election architecture which appropriate very important for reaching accuracy which high and efficient of the diagnosis process [15]. Three CNN architectures that are popular and proven to be effective in medical image classification are Inception V3, DenseNet 121, and MobileNet V2 .

DenseNet121 was chosen in this study because it has several significant advantages. First, this architecture ensures robust gradient flow, which helps overcome the vanishing gradient problem that often arises in deep neural networks. Second, DenseNet121 offers efficient computing by connecting each layer directly to all previous layers, reducing the complexity of the extracted features. Additionally, this architecture allows smoother error signal propagation through the network, which contributes to more effective and efficient model training. With these advantages, DenseNet121 shows strong potential in feature extraction, segmentation, and object detection, making it a solid choice for various applications in image recognition and classification [39].

Inception V3 was chosen for this study because of its outstanding efficiency and performance. This architecture is an enhanced version of the basic framework of Inception V1, which was developed by the Google team and introduced in GoogLeNet in 2014. Inception V3 has a total of 42 layers, which increases the depth and capabilities of the network. Its main advantages include splitting into smaller convolutions, use of irregular convolutions via spatial factorization, incremental approximation, and effective reduction of grid size. This optimization produces high accuracy, precision, recall, and F-measure values, making it very effective for complex image classification tasks [40].

MobileNet V2 was chosen for this research because MobileNetV2 has a number of advantages that make it superior in applications on mobile devices. This architecture uses depth-wise separable convolutions which reduces the amount of computation significantly, making it more efficient compared to its predecessor, MobileNetV1. Additionally, MobileNetV2 introduces building blocks with residual connections and expansion and projection layers that help gradients flow through the network, improving training efficiency. MobileNetV2 also shows improved computational efficiency by performing only 300 million multiply-accumulate operations (MACs) for a 224x224 RGB image, compared to 569 million MACs on MobileNetV1, and has almost 20% fewer parameters. This efficiency makes it ideal for mobile devices with limited memory access and low computing power. Model training was selected using these three architectures using the ADAM optimizer [41].

In research by Dogo et al. (2020), combining different optimizations including vanilla SGD, with momentum, and with Nesterov momentum, RMSProp, ADAM, ADAGrad, ADADelta, ADAMax, and NADAM were trained and evaluated using three benchmark image classification datasets. The research results show that the Adam optimizer achieves excellent performance and performance compared to several other algorithms [42]. Then Irfan et al. (2020) in their research on the comparison of the SGD, RMSProp, and ADAM optimizers in animal classification using the CNN algorithm. The research results show that the CNN algorithm with the ADAM optimizer produces higher accuracy when compared to the SGD and RMSprop methods. The model trained using the optimal Adam function achieved 89.81% accuracy in the experiment, indicating the validity of the hypothesis [43].

In research conducted by Mohanty et al. (2023) regarding DL detecting and classifying diabetic retinitis using the DenseNet 121 architecture. Results from the DenseNet model 121 show accuracy classification excellence, ability classification which fast, so suiTable for real-time medical applications [44]. Research by Junayed et al. (2021) regarding an automatic cataract detection system using DL for fundus images. The experimental results show that with five trained CNN models, namely MobileNet, VGG-16, VGG-19, Inception-v3, and ResNet-50, eye disease reaches performance competitive. Accuracy which is high, efficiency cost And time possible doctor eyes detect disease cataracts in a way appropriate time more appropriate [17]. Another study by Guo et al. (2021) used the MobileNet V2, AlexNet, and Inception V3 architectures in predicting various eye diseases using fundus photography through transfer learning. The results showed that the MobileNet V2 architecture was superior to other architectures with an accuracy of 96.2%, a sensitivity of 90.4%, and a specificity of 97.6% through five trials on test data [45].

Based on previous research that has been described, this research only carried out one class classification, or what is called binary classification, and on average used a batch size of 32. Therefore, this research proposes the Inception V3, DenseNet 121, and MobileNet V2 architectures using 4 classes, namely normal, cataracts, glaucoma, and diabetic retinopathy, and used a batch size of 256. This research used a data sharing technique, namely Holdout Validation with a percentage of 70:30, and used the Adam optimizer. The objective of this research is to compare and evaluate performance from CNN architecture in eye diseases by using eye fundus images to find architecture that is capable of producing accuracy which high. Process classification data on study It was done by comparing three models namely Inception V3, Dense Net 121, and Mobile Net V2 with pre-trained models. Therefore, this study will use Inception V3, DenseNet 121, and MobileNet V2 to classify eye diseases in eye fundus images. It is hoped that by exploiting the advantages of each architecture, we can achieve the best performance in detecting eye diseases, thereby supporting more efficient screening and diagnosis processes in the medical field.

2. MATERIALS AND METHODS

The steps of this research can be seen in the methodology study which is outlined in figure 1. This research stage begins with a review of literature studies from scientific articles and other sources of information related to the topic study. Then stage collection datasets image disease eye, After the data is collected, it goes to the initial data processing and data sharing stage using CNN architectural modeling, namely the Inception V3, DenseNet 121, and MobileNet V2 classification, then goes to the evaluation stage of the research results that have been carried out.

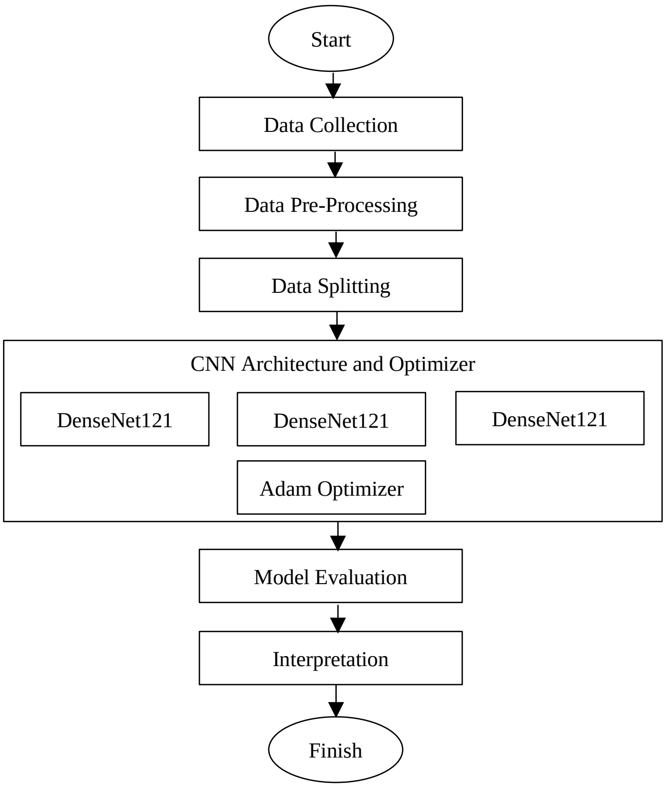

Figure 1. Research Methodology

2.1 Disease Eye

Diabetic Eye Disease (DED) is a group of eye problems that affect diabetes patients. Identifying DED is an important activity in retinal fundus imaging because early diagnosis and treatment can ultimately minimize the risk of visual impairment. Retinal fundus images play an important role in the classification and initial identification of DED [46].

2.2 Architecture CNN

Convolutional Neural Network (CNN) is a type of architecture network nerve imitation designed specifically for processing data that has a grid-like topology in images. CNN consists of several layers, including layer convolution which extracts features important from input via filter, layer unification which reduces data dimensions to reduce computational complexity, and fully connected layers that connect every neuron from one layer to the next to make the final decision. CNNs are very effective in tasks such as image recognition, face recognition, and object detection due to their ability to capture spatial patterns and feature hierarchies from input data [47]. At this stage, Convolutional Neural Network (CNN) used for image classification in eye disease was designed and implemented. Process This involve a number of step important. First, the election model is done by choosing the architecture CNN which in accordance, like Inception V3, DenseNet 121, and MobileNet V2 has proved to be effective in image classification tasks. Next, several convolution and pooling layers are built to extract features from the picture. Layer convolution works to detect feature local to picture, while the pooling layer is used to reduce feature dimensions and control overfitting [38]. After that, fully connected layers are added to perform classification based on the features extracted by the convolution and pooling layers. Finally, an output layer with softmax activation is used to classify images into cataract, diabetic retinopathy, glaucoma, and normal categories, generating probabilities for each category to enable decision making based on model predictions [48]. Figure 2 shows the convolution layer of CNN.

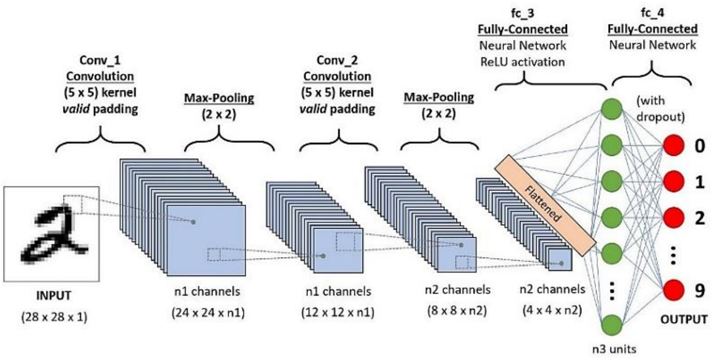

Figure 2. Convolution Layers CNN

2.3 InceptionV3

Inception V3 is a pre-trained convolutional neural network model with 48 layers, that is capable of learning and recognizing complex patterns and features in medical images. One of the key features of Inception V3 is its ability to scale to large datasets and handle images of various sizes and resolutions. Inception V3 Connected Convolutional Network is applied for the early detection of eye diseases including diabetic retinopathy. The model classifies fundus images based on their severity. The dataset used is Eye Disease Classification. The design of Inception V3, consisting of several convolution stages, is depicted in Figure 3 [49].

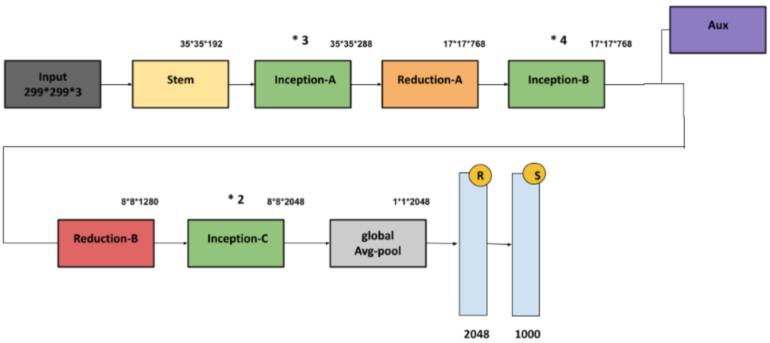

Figure 3. Architecture CNN Inception V3

2.4 DenseNet121

DenseNet is a type of CNN that enables deeper network architectures by connecting each layer to every other layer in a sequential manner. The DenseNet 121 model stands out for its unique connectivity patterns and efficient feature extraction capabilities, achieving impressive results with an average accuracy of 91.2%, specificity of 69%, sensitivity of 96%, and more. The architecture consists of basic convolution and pooling layers, dense blocks, and transition layers. It begins with a convolution block that applies a 7x7 sliding window to the input image, producing 64 output layers. A global average pooling layer is also used to perform spatial pooling on the feature map, resulting in a fixed-length vector representation. The DenseNet 121 architecture is shown in Figure 4 [39].

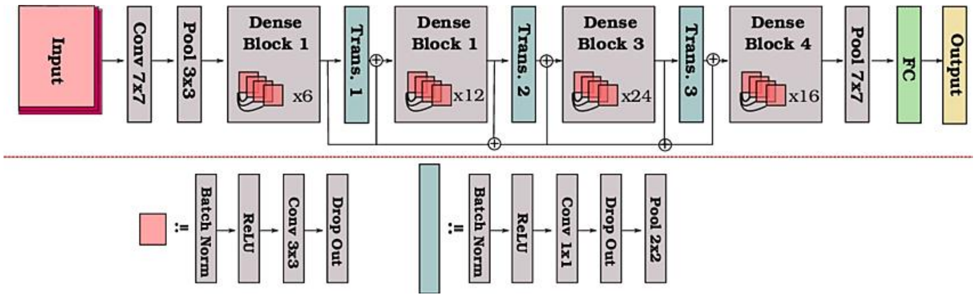

Figure 4. Architecture DenseNet 121

2.5 MobileNet V2

MobileNet V2 is an improvement from the previous version which uses depthwise separable techniques convolution (DSP), which is also known as convolution separated. The goal is to create powerful neural network architectures by tuning parameters [50]. Differences between MobileNet -V2 and the version previously including additional feature connection bottleneck and shortcut. Feature this can reduce the number of calculations performed by the machine compared to the original cellular network. Version 2 of mobile net allows input images larger than 32x32, providing better performance for larger image sizes. Because the mobile net is designed for mobile access, this architecture must be able to interoperate with a power computer which minimal, However still can give a level accuracy which high [51]. Figure 5 is the Convolution Layer of the MobileNet V2 architecture.

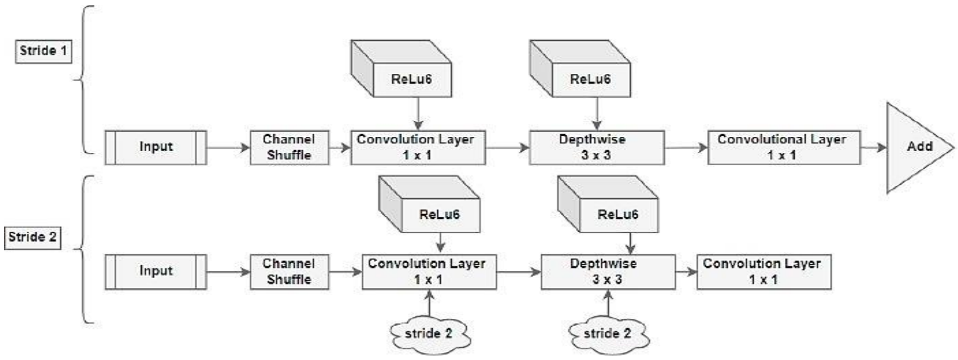

Figure 5. Convolution Layers Architecture MobileNet V2

2.6 Adam Optimizer

Adam's Optimizer is an adaptive moment estimation algorithm that combines the advantages of AdaGrad and RMSProp. It adjusts the learning rate for each parameter based on the first and second moments of the gradient, making it efficient and effective for large datasets and non-stationary objectives. Adam is computationally efficient, requires less memory, and is not affected by rescaling the diagonal of the gradient. It is especially useful for problems with sparse gradients and noisy objectives, and its hyperparameters have intuitive interpretation and usually require little customization [52].

2.7 Evaluation Metrics

To evaluate the performance of a classifier, various metrics are employed. This study utilizes a range of evaluation metrics, starting with the confusion matrix, which provides a numerical representation of classification accuracy [53]. The confusion matrix is a widely used technique in machine learning that compares actual class labels with predicted class labels. It consists of two dimensions: actual class and predicted class, with each row representing an actual class instance and each column representing a predicted class state. The matrix includes four key components: true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The trained model is then assessed using multiple criteria, including recall, accuracy, precision, and F1 score, to provide a comprehensive evaluation of its performance [54].

2.7.1 Accuracy

This equation is used to measure the accuracy of the classifier. To calculate accuracy, divide the total amount of data by the amount of correctly classified data. Equation 1 is the formula for accuracy.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

2.7.2 Precision

The precision metric indicates the proportion of positive data that is accurately predicted. In other words, high precision results in fewer false positive results. Equation 2 is the precision formula.

$$Precision = \frac{TP}{TP+FP} \tag{2}$$

2.7.3 Recall

The parameter used to assess the completeness of a classifier is called recall. Recall is a measure of false negatives; lower recall is a measure of false negatives. An increase in recall is often accompanied by a decrease in precision. Equation 3 is the recall formula.

$$Recall = \frac{TP}{TP+FN} \tag{3}$$

2.7.4 F1-Score

The result of multiplying recall and precision is divided by the sum of recall and precision to get the F1 score. Equation 4 is the F1-Score formula.

$$F1 - Score = 2\,x\,\frac{Recall\ x\ Precision}{Recall+Precision} \tag{4}$$

3. RESULTS AND DISCUSSION

3.1 Data Collection

The first step in this study is collection of data. The dataset used in this research is a dataset of 4217 records about diabetes mellitus obtained from Kaggle (https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification/code) which provides datasets on eye diseases. The eye_diseases_classification dataset on Kaggle is a collection of 4,217 eye images divided into four disease categories: cataract (1,038 images), diabetic retinopathy (1,098 images), glaucoma (1,007 images), and normal condition (1,074 images). This dataset offers significant variation for analysis using the Convolutional Neural Network (CNN) algorithm in detecting eye diseases. While there is no specific information on lighting conditions and image quality, this dataset was collected from credible sources such as IDRiD, Ocular Recognition, and HRF, so it can be assumed that the quality is adequate for research purposes. As an open data repository, this dataset can be accessed by researchers without restrictions, supporting transparency and reproducibility in research on eye health.

3.2 Data Preprocessing

After data was collected, so step next is data preprocessing or process data preprocessing. The stage in data preprocessing carried out is augmentation by scaling the image pixel values. Then rotate the image up to 45 degrees, and flip the image horizontally and vertically. Exposure to normalization and augmentation displayed in Table 1 . Data augmented is training data, while validation data and test data cannot be augmented. Figure 6 is the result of augmentation which has done. Figure 7 is the appearance data figure on data training.

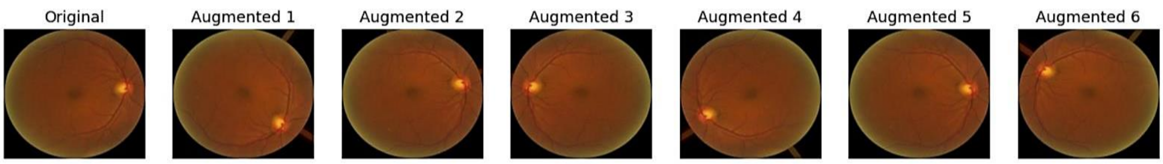

Figure 6. Results Augmentation

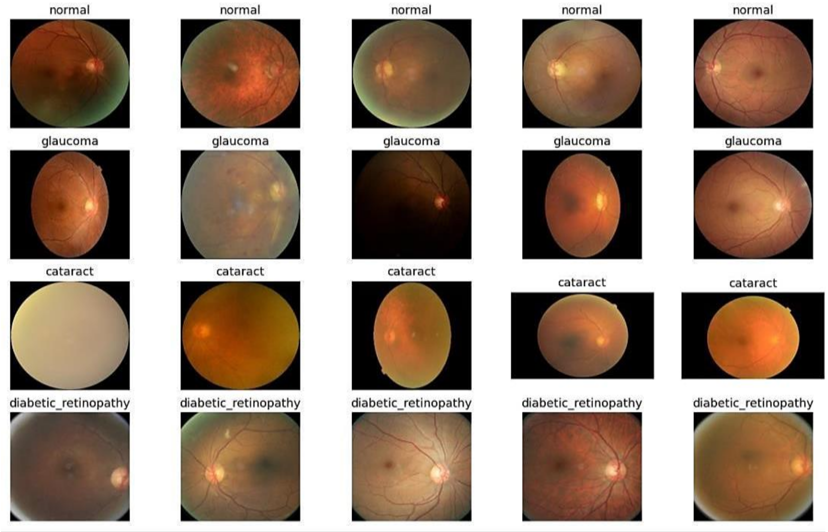

Figure 7. Displaying Data Training

Table 1. Augmentation which Used

Parameter	Mark	Explanation
rescale 1./255	1./255	Scales the image pixel values from range [0, 255] become [0, 1]. This helps normalize data so makes it easier model in training process
rotation_range	45	Randomly rotates the image up to 45 degrees clockwise or counterclockwise. This helps the model to be more robust to image rotation.
horizontal_flip	True	Randomly flips images horizontally. For example, an image showing the left direction will be flipped to the right direction.
vertical_flip	True	Randomly flips the image vertically. For example, an image showing the top will be flipped to become the bottom

3.3 Data Splitting

Based on the data collected, eye fundus data has four classes, namely normal, glaucoma, cataract, and diabetic_retinopathy. The amount of whole data on the fourth class is 4217 fundus images. In the deep learning concept, there is a division of data, namely training, test, and validation data. Data sharing uses techniques including cross-validation and hold-out. In this research, the data used the holdout technique with a ratio of 70:30, namely 70% of the training data and test data (testing). Figure 1 is data training for every class, Table 1 shows the amount of data for each class that has been divided using the holdout technique. The next stage is to set hyperparameters using the CNN algorithm and several architectures as in Table 2.

Table 2. Distribution Datasets Use Technique Holdout

Holdout	Class	80:20		Testing Data
		Training Data	Validation Data	
70:30	Cataract	580	146	312
	Glaucoma	563	141	303
	Retinopathy Diabetic	614	154	330
	Normal	600	151	323

3.4 Modeling Architecture CNN

The data is trained using a CNN architecture and uses MobileNet V2, DenseNet 121, and Inception V3 architectures. In the CNN model, custom layers are added to suit your needs. Table 3 is exposure to layer customization which added to the model. After customizing the model, then the next stage is determining the optimizer, callbacks and starting training using MobileNet V2, DenseNet 121and Inception V3. In Table 4, the hyperparameter settings used are presented.

Table 3 . Layer Customization CNN

Layer	Function
base_model	It is a basic model used as an architectural basis (for example, MobileNet V2, DenseNet 121, or Inception V3). It works as a feature extractor beginning.
GlobalAveragePooling2D()	Takes the global average of each extracted feature by base_model, reduce

Layer	Function
	dimensions data while retaining important information.
Dense(256, activation='relu')	Layer is completely connected (dense) with 256 neurons and function Activation ReLU, which works For study complex patterns of extracted features.
Dropout(0.3)	Dropouts layers with level dropout 30%, used to prevent overfitting by disabling in a way random 30% from neurons during training.
Dense(256, activation='relu')	Layer completely connected second with 256 neurons and function Activation ReLU, which add capacity learn a model to map features to output classes.
Dense(4, activation='softmax')	Layer output with 4 neurons and function activation softmax, used to classify input into wrong one from four class which desired.

Table 4. Arrangement Hyperparameters

Name Parameter	Arrangement Parameter	Mark
Adam	Learning Rate	0.0001
Early Stopping	Patience	10
Model Checkpoints	Save_best_only,	True
	Save_weights_only	True
ReduceLROnPlateau	Factor	0.3
	Patience	4
	Min_lr	0.00001
	Batches size	256
	Epoch	50

3.5 Evaluation

The trained model is then evaluated using the Classification Report, Confusion Matrix, and ROC Curve. Classification Report is presented in Tables 5, 6, and 7. Table 5 shows the results of the classification report using the MobileNet V2 architecture. The results of the classification report by the Inception V3 architecture are presented in Table 6 and the DenseNet 121 architecture in Table 7. In Table 5 using MobileNet V2 shows very good performance in the Cataract class which has recall, precision, f1-score and support as big as 0.85577, 0.94014, 0.89597, and 312 with accuracy 0.81309 occupy accuracy highest compared to other architectures. Likewise Table 6 using Inception V3 for normal class classification shows very good performance with recall, precision, f1-score and support values, namely 0.84615, 0.85993, 0.86076 and 312 with an accuracy of 0.77366. The model in Table 7 also trains the normal class well with recall, precision, f1-score and support values of 0.88782, 0.87658, 0.88217 and 312 with an accuracy of 0.76735. The best accuracy is 0.81309 by MobileNet V2 architecture.

Table 5. Classification Reports Architecture MobileNet V2

Name Class	Recall	Precision	F1-Score	Support
Cataract	0.85577	0.94014	0.89597	312
Diabetic Retinopathy	0.76667	0.92674	0.83914	330
Glaucoma	0.68317	0.86250	0.76243	303
Normal	0.94118	0.64544	0.76574	323
Accuracy			0.81309	1268

Table 6. Classification Reports Architecture Inception V3

Name Class	Recall	Precision	F1-Score	Support
Cataract	0.84615	0.85993	0.86076	312
Diabetic Retinopathy	0.78182	0.82958	0.76661	330
Glaucoma	0.73927	0.70886	0.64336	303
Normal	0.72755	0.70359	0.69799	323
Accuracy			0.77366	1268

Table 7. Classification Reports Architecture DenseNet 121

Name Class	Recall	Precision	F1-Score	Support
Cataract	0.88782	0.87658	0.88217	312
Diabetic Retinopathy	0.73939	0.88406	0.80528	330
Glaucoma	0.64356	0.75000	0.69272	303
Normal	0.79567	0.61779	0.69553	323
Accuracy			0.76735	1268

An analysis of model errors or failures, such as those caused by blurry, noisy, or poorly lit images, could enhance the evaluation by providing insights into the model's limitations and potential areas for improvement.

Discussing challenging cases or examples of misclassification would add significant value to the study. As shown in figure 8, a Confusion Matrix was used to evaluate the MobileNet V2 model's performance on a testing dataset of 1,268 images. Out of these, 1,008 images were correctly classified, while 260 were misclassified. Incorporating an analysis of these errors, particularly under challenging conditions, could provide deeper insights into the model's robustness and areas where enhancements are needed.

Figure 9 is the Confusion Matrix from Inception V3. The results of the confusion matrix from Inception V3 explain the amount of testing data that is proven correct as much as 981 data, and as much as 287 testing data is still detected incorrectly during classification. Based on figure 10 is a confusion matrix from DenseNet 121 which shows that 941 testing data were correctly detected and 327 data were still incorrectly detected. So the best Confusion Matrix is obtained by the MobileNet V2 architecture.

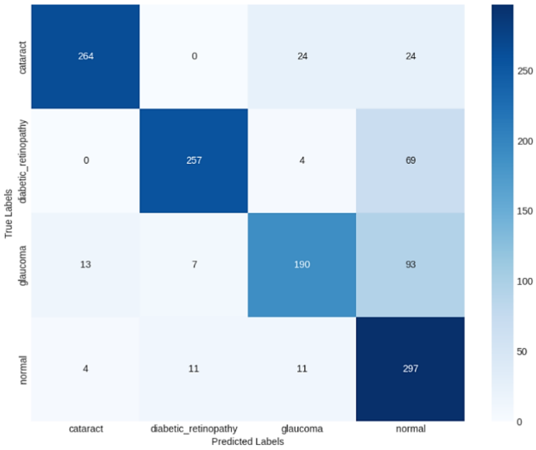

Figure 8. Confusion Matrix MobileNet V2

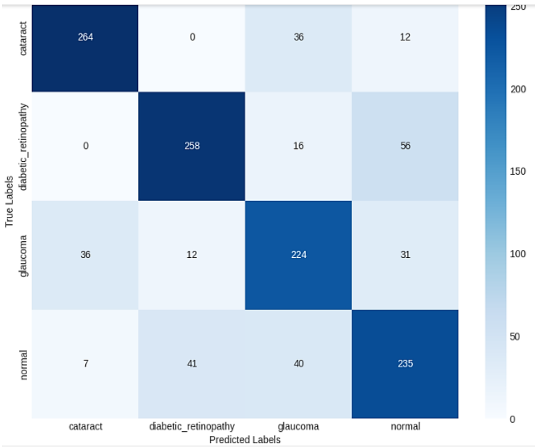

Figure 9 . Confusion Matrix Inception V3

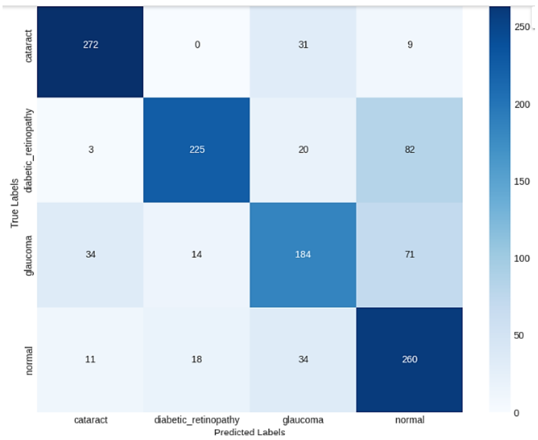

Figure 10. Confusion Matrix DenseNet 121

The model for the eye fundus dataset was trained with 70% of the data and tested with 30%. The training process spanned 50 epochs. Figures 11, 12, and 13 depict the training and validation curves for the MobileNet V2, Inception V3, and DenseNet 121 architectures, respectively. These figures illustrate the relationship between accuracy, loss, and the number of epochs. Accuracy shows a positive correlation with more epochs, while loss shows a negative correlation, decreasing as epochs increase. Among the architectures, MobileNet V2 performed the best, with training and validation accuracies around 83% and 80%, and losses around 0.4 and 0.5, indicating sTable and superior performance.

Figure 11 shows the training and validation accuracy curves, indicating that the MobileNet V2 architecture model achieves a training accuracy of about 83% and a validation accuracy of about 80%. The model shows good performance, with training and validation accuracy consistently increasing over time. The loss training and validation curves show that the model achieves a training loss of about 0.4 and a validation loss of about 0.5. The training and validation losses consistently decreased over time, indicating that the model learned well and generalized well to the validation data.

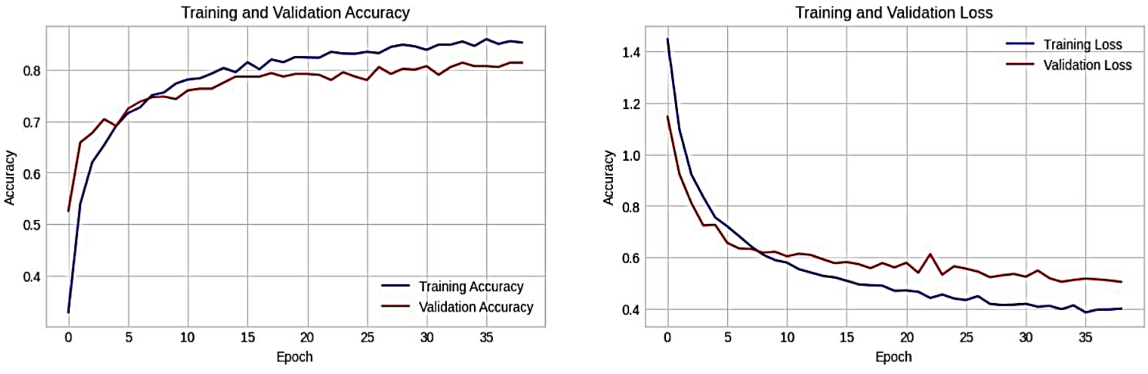

Figure 11. Curve Accuracy and Loss MobileNet V2.

The graph in figure 12 shows that the Inception V3 architecture model achieved a training accuracy of about 79% and a validation accuracy of about 75% after 18 epochs. The graph also shows that the training loss decreased to about 0.55, while the validation loss decreased to about 0.62. This shows that the model learned the training data well, but experienced slight overfitting, as the validation loss was higher than the training loss. Overfitting can occur when the model learns the training data too well and becomes too specific to the training data, so it cannot generalize well to new data.

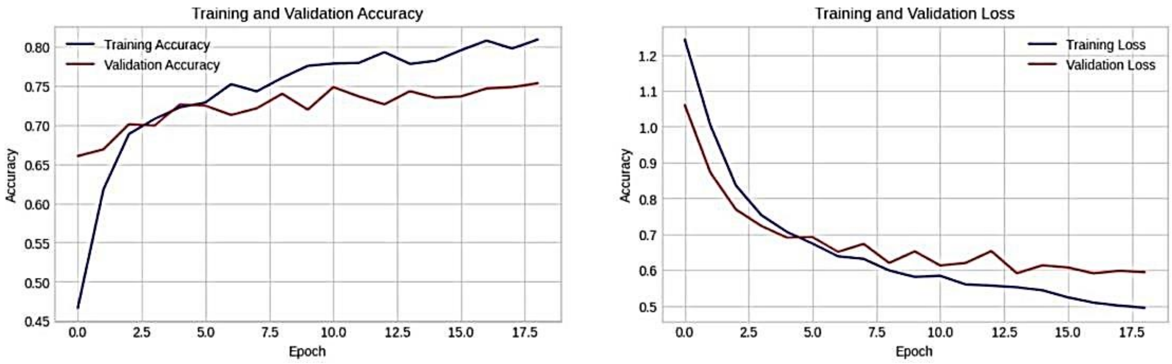

Figure 12. Curve Accuracy and Loss Inception V3

The graph in figure 13 shows that the DenseNet 121 architecture model achieves high training and validation accuracy, with training accuracy peaking at around 0.78 and validation accuracy peaking at around 0.75. This shows that the model learns well and is able to generalize to data that it has never seen before. The loss graph shows that the training and validation loss drops significantly during training, reaching around 0.6 for training loss and around 0.7 for validation loss. This shows that the model becomes better at predicting the correct output as training progresses. Although there is a slight difference between the training and validation accuracy, and the training and validation loss, it shows that the model is not overfit and can generalize well to new data.

The Receiver Operating Characteristic (ROC) curve is a probability curve that measures the area under the curve using Area Under the Curve (AUC). The maximum value of AUC is 1, which has the most perfect classification. The best AUC value among the three architectures is MobileNet V2 with an average AUC of 0.9594.

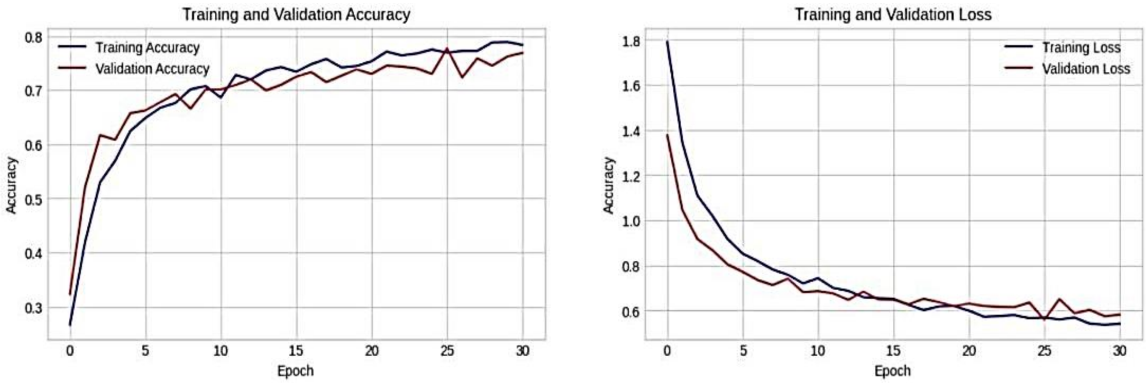

Figure 13. Curve Accuracy and Loss DenseNet 121.

The AUC-ROC curve in figure 14 shows that the MobileNet V2 architecture achieves an AUC value above 0.93 for each eye fundus disease class, with the following details: AUC for cataract is 0.9857, diabetic retinopathy 0.9728, glaucoma 0.9431, and normal condition 0.9360. The average AUC-ROC value of 0.93 indicates that MobileNet V2 has an almost perfect classification ability in detecting various types of eye diseases. However, adding a discussion of more challenging image conditions or examples of classification failures can add significant value. It is important to understand the limitations of the model and the situations where classification accuracy can be affected, so that further improvements and developments can be made in future research.

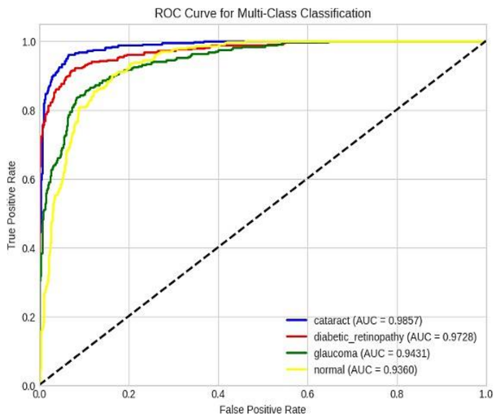

Figure 14. Curve AUC-ROC MobileNet V2

Then in figure 15 using the Inception V3 architecture achieves an AUC value above 0.91 in each class. The AUC value in each class, including AUC cataract = 0.9804, AUC diabetic_retinopathy = 0.9637, AUC glaucoma = 0.9402 and AUC normal = 0.9101 value in each class and the average AUC-ROC in Figure 15 worth 0.91 can be concluded that Inception V3 has a near perfect classification. The third architecture, DenseNet 121 in figure 16, has an AUC value above 0.88. The AUC value in each class, including AUC cataract = 0.9804, AUC diabetic_retinopathy = 0.9447, AUC glaucoma = 0.8923 and AUC normal = 0.8844 value in each class and the average AUC-ROC in Figure 15 is 0.88, it can be concluded that Inception V3 has a good classification.

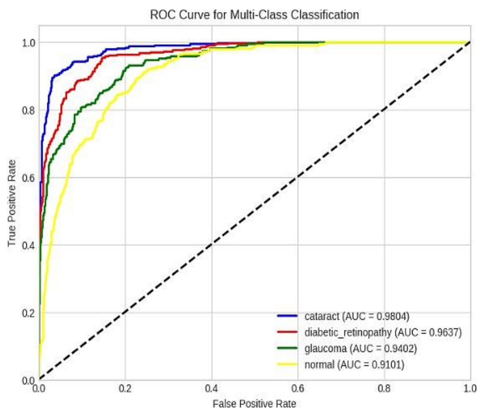

Figure 15. Curve AUC-ROC Inception V3

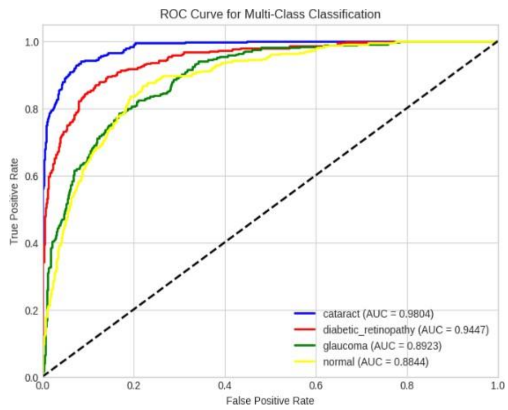

Figure 16. Curve AUC-ROC DenseNet 121

The accuracy results of the experiment are shown in figure 17. Figure 17 shows that the experiment using the MobileNet V2 architecture produced the highest accuracy value of 81.3%. On position accuracy best second achieved by architecture Inception V3 as big as 77.3%. DenseNet 121 is at position lowest with accuracy as big as 76.7%. Figure 18 shows the successful architecture MobileNet V2 in that predicted class on fundus eye. Figure 18 chosen is a class cataract. Results the test shows the cataract class as the prediction result with the highest probability of 0.775954.

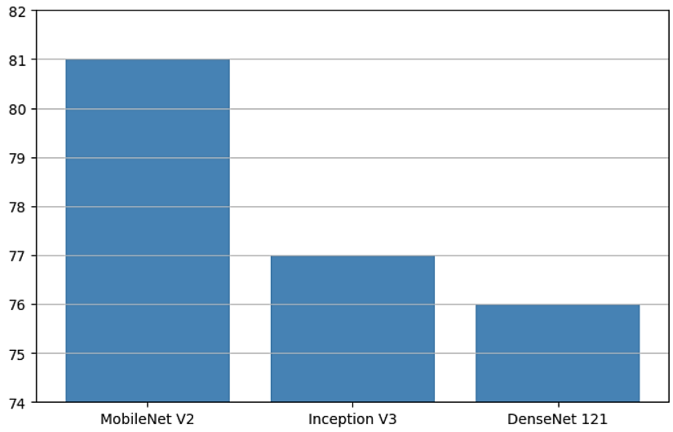

Figure 17. Diagram Comparison Architecture

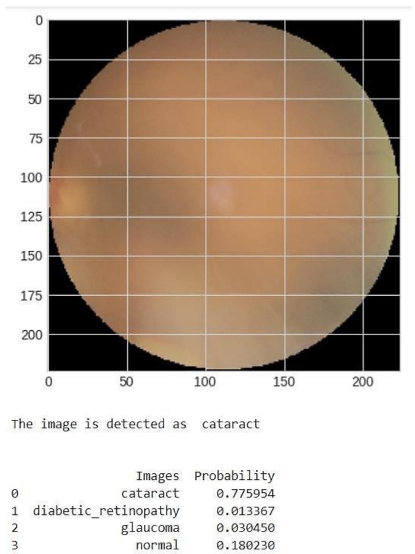

Figure 18. Test Try Image

Given the growing use of deep learning in medical diagnostics, including eye disease classification, real-world implementation faces several challenges that need to be considered. Larger models such as Inception V3 and DenseNet 121 offer impressive accuracy but come with significant computational requirements, making them less suiTable for real-time applications or use in resource-limited settings. These models demand powerful hardware, high memory capacity, and longer processing times, which can be impractical for quick diagnostics in clinical environments or on porTable devices. In contrast, lighter models like MobileNet V2, which utilize depthwise separable convolutions, offer a more efficient alternative by significantly reducing the computational load while maintaining competitive accuracy. This trade-off between computational cost and model performance is particularly important for medical applications where quick, real-time processing is essential, especially in mobile or remote diagnostic settings. Understanding these practical limitations and the hardware capabilities required to run such models is crucial when selecting an appropriate architecture for real-world use in medical fields, ensuring that the model can perform efficiently without compromising on diagnostic speed or accuracy.

4. CONCLUSION

The conclusion of this research shows that the MobileNet V2 architecture has the best performance in eye disease image classification with the highest accuracy of 81.3%, followed by Inception V3 with 77.3% accuracy, and DenseNet 121 with 76.7% accuracy. MobileNet V2 also shows the best performance in detecting cataract classes with the lowest detection error rate. Increasing the number of epochs has proven to be effective in increasing accuracy and reducing loss values in the model. The preprocessing process which includes image normalization and data augmentation such as image rotation and flipping also plays an important role in improving model robustness. Overall, this study shows that MobileNet V2 is superior in eye disease classification in fundus images compared to other architectures.

For future research, algorithms such as EfficientNet, ResNet, and Vision Transformers (ViT) are recommended. EfficientNet offers high efficiency and accuracy by utilizing optimal scaling techniques, making it ideal in clinical environments with limited resources. ResNet is known for its residual architecture that can overcome the problem of vanishing gradients, crucial for the classification of complex medical images. Meanwhile, Vision Transformers (ViT) utilizes the attention mechanism to capture spatial relationships in medical images, which can improve the detection of important features that may be missed by traditional CNN architectures. The combination of these algorithms can improve the early detection of eye diseases and support more efficient and accurate diagnosis. In addition to the algorithms mentioned, future research could incorporate CLAHE and ESRGAN for image enhancement. CLAHE improves local contrast and enhances the visibility of retinal features, which is crucial for precise diagnosis. ESRGAN, an advanced generative adversarial network, further refines image resolution and detail, enhancing the model's ability to detect subtle abnormalities. This combination can significantly improve the quality of input images, leading to more accurate and reliable classification of eye diseases, ultimately supporting early detection and better patient outcomes.

REFERENCES

[1] M. C. S. Tang, S. S. Teoh, H. Ibrahim, and Z. Embong, "Neovascularization detection and localization in fundus images using deep learning," *Sensors*, vol. 21, no. 16, Aug. 2021, doi: 10.3390/s21165327.

[2] F. Du *et al.*, "Recognition of eye diseases based on deep neural networks for transfer learning and improved D-S evidence theory," *BMC Med Imaging*, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12880-023-01176-2.

[3] A. E. Ilesanmi, T. Ilesanmi, and G. A. Gbotoso, "A systematic review of retinal fundus image segmentation and classification methods using convolutional neural networks," *Healthcare Analytics*, vol. 4. Elsevier Inc., Dec. 01, 2023. doi: 10.1016/j.health.2023.100261.

[4] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia, and H. Hamam, "Diabetic Retinopathy Diagnosis from Fundus Images Using Stacked Generalization of Deep Models," *IEEE Access*, vol. 9, pp. 108276–108292, 2021, doi: 10.1109/ACCESS.2021.3101142.

[5] Y. Pan *et al.*, "Fundus image classification using Inception V3 and ResNet-50 for the early diagnostics of fundus diseases," *Front Physiol*, vol. 14, 2023, doi: 10.3389/fphys.2023.1126780.

[6] J. Wang, L. Yang, Z. Huo, W. He, and J. Luo, "Multi-Label Classification of Fundus Images with EfficientNet," *IEEE Access*, vol. 8, pp. 212499–212508, 2020, doi: 10.1109/ACCESS.2020.3040275.

[7] B. Bhatkalkar, A. Joshi, S. Prabhu, and S. Bhandary, "Automated fundus image quality assessment and segmentation of optic disc using convolutional neural networks," *International Journal of Electrical and Computer Engineering*, vol. 10, no. 1, pp. 816–827, 2020, doi: 10.11591/ijece.v10i1.pp816-827.

[8] M. Stanojević, D. Drašković, and B. Nikolić, "Retinal disease classification based on optical coherence tomography images using convolutional neural networks," *J Electron Imaging*, vol. 32, no. 03, Nov. 2022, doi: 10.1117/1.jei.32.3.032004.

[9] A. O. Joshua, F. V Nelwamondo, and G. Mabuza-Hocquet, "Blood Vessel Segmentation from Fundus Images Using Modified U-net Convolutional Neural Network", doi: 10.18178/joig.8.1.21-25.

[10] V. Mayya, S. K. S, U. Kulkarni, D. K. Surya, and U. R. Acharya, "An empirical study of preprocessing techniques with convolutional neural networks for accurate detection of chronic ocular diseases using

fundus images," *Applied Intelligence*, vol. 53, no. 2, pp. 1548–1566, Jan. 2023, doi: 10.1007/s10489-022-03490-8.
[11] X. Yao *et al.*, "Distribution of diabetic retinopathy in diabetes mellitus patients and its association rules with other eye diseases," *Sci Rep*, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-96438-w.
[12] N. Tsiknakis *et al.*, "Deep learning for diabetic retinopathy detection and classification based on fundus images: A review," *Computers in Biology and Medicine*, vol. 135. Elsevier Ltd, Aug. 01, 2021. doi: 10.1016/j.compbiomed.2021.104599.
[13] G. Ali, A. Dastgir, M. W. Iqbal, M. Anwar, and M. Faheem, "A Hybrid Convolutional Neural Network Model for Automatic Diabetic Retinopathy Classification From Fundus Images," *IEEE J Transl Eng Health Med*, vol. 11, pp. 341–350, 2023, doi: 10.1109/JTEHM.2023.3282104.
[14] M. M. Butt, D. N. F. A. Iskandar, S. E. Abdelhamid, G. Latif, and R. Alghazo, "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features," *Diagnostics*, vol. 12, no. 7, Jul. 2022, doi: 10.3390/diagnostics12071607.
[15] S. Duan *et al.*, "Semi-supervised classification of fundus images combined with CNN and GCN," *J Appl Clin Med Phys*, vol. 23, no. 12, Dec. 2022, doi: 10.1002/acm2.13746.
[16] A. Skouta, A. Elmoufidi, S. Jai-Andaloussi, and O. Ouchetto, "Hemorrhage semantic segmentation in fundus images for the diagnosis of diabetic retinopathy by using a convolutional neural network," *J Big Data*, vol. 9, no. 1, Dec. 2022, doi: 10.1186/s40537-022-00632-0.
[17] M. S. Junayed, M. B. Islam, A. Sadeghzadeh, and S. Rahman, "CataractNet: An automated cataract detection system using deep learning for fundus images," *IEEE Access*, vol. 9, pp. 128799–128808, 2021, doi: 10.1109/ACCESS.2021.3112938.
[18] R. Sarki, K. Ahmed, H. Wang, Y. Zhang, J. Ma, and K. Wang, "Image Preprocessing in Classification and Identification of Diabetic Eye Diseases," *Data Sci Eng*, vol. 6, no. 4, pp. 455–471, Dec. 2021, doi: 10.1007/s41019-021-00167-z.
[19] R. Sarki, K. Ahmed, H. Wang, and Y. Zhang, "Automatic Detection of Diabetic Eye Disease through Deep Learning Using Fundus Images: A Survey," *IEEE Access*, vol. 8. Institute of Electrical and Electronics Engineers Inc., pp. 151133–151149, 2020. doi: 10.1109/ACCESS.2020.3015258.
[20] R. Mahum, S. U. Rehman, O. D. Okon, A. Alabrah, T. Meraj, and H. T. Rauf, "A novel hybrid approach based on deep cnn to detect glaucoma using fundus imaging," *Electronics (Switzerland)*, vol. 11, no. 1, Jan. 2022, doi: 10.3390/electronics11010026.
[21] A. Neto, J. Camara, and A. Cunha, "Evaluations of Deep Learning Approaches for Glaucoma Screening Using Retinal Images from Mobile Device," *Sensors*, vol. 22, no. 4, Feb. 2022, doi: 10.3390/s22041449.
[22] H. N. Veena, A. Muruganandham, and T. Senthil Kumaran, "A novel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 6187–6198, Sep. 2022, doi: 10.1016/j.jksuci.2021.02.003.
[23] M. J. M. Zedan, M. A. Zulkifley, A. A. Ibrahim, A. M. Moubark, N. A. M. Kamari, and S. R. Abdani, "Automated Glaucoma Screening and Diagnosis Based on Retinal Fundus Images Using Deep Learning Approaches: A Comprehensive Review," *Diagnostics*, vol. 13, no. 13. Multidisciplinary Digital Publishing Institute (MDPI), Jul. 01, 2023. doi: 10.3390/diagnostics13132180.
[24] J. Latif, S. Tu, C. Xiao, S. Ur Rehman, A. Imran, and Y. Latif, "ODGNet: a deep learning model for automated optic disc localization and glaucoma classification using fundus images," *SN Appl Sci*, vol. 4, no. 4, Apr. 2022, doi: 10.1007/s42452-022-04984-3.
[25] W. T. Song, I. C. Lai, and Y. Z. Su, "A Statistical Robust Glaucoma Detection Framework Combining Retinex, CNN, and DOE Using Fundus Images," *IEEE Access*, vol. 9, pp. 103772–103783, 2021, doi: 10.1109/ACCESS.2021.3098032.
[26] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, Jun. 2021, doi: 10.3390/s21113704.
[27] A. Saber, M. Sakr, O. M. Abo-Seida, A. Keshk, and H. Chen, "A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique," *IEEE Access*, vol. 9, pp. 71194–71209, 2021, doi: 10.1109/ACCESS.2021.3079204.
[28] Prof. N. P. Tembhare, Prof. P. U. Tembhare, and Prof. C. U. Chauhan, "Chest X-ray Analysis using Deep Learning," *Int J Res Appl Sci Eng Technol*, vol. 11, no. 1, pp. 1441–1447, Jan. 2023, doi: 10.22214/ijraset.2023.48637.
[29] A. A. Ali and F. A. A. Dawood, "Deep Learning of Diabetic Retinopathy Classification in Fundus Images," *Journal of Engineering*, vol. 29, no. 12, pp. 139–152, Dec. 2023, doi: 10.31026/j.eng.2023.12.09.
[30] A. Bin Tufail *et al.*, "Diagnosis of Diabetic Retinopathy through Retinal Fundus Images and 3D Convolutional Neural Networks with Limited Number of Samples," *Wirel Commun Mob Comput*, vol. 2021, 2021, doi: 10.1155/2021/6013448.

[31] P. D. Rinanda, D. N. Aini, T. A. Pertiwi, S. Suryani, and A. J. Prakash, "Implementation of Convolutional Neural Network (CNN) for Image Classification of Leaf Disease In Mango Plants Using Deep Learning Approach," *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 2, pp. 56–61, Feb. 2024, doi: 10.57152/predatecs.v1i2.872.

[32] A. Raj, N. A. Shah, A. K. Tiwari, and M. G. Martini, "Multivariate Regression-Based Convolutional Neural Network Model for Fundus Image Quality Assessment," *IEEE Access*, vol. 8, pp. 57810–57821, 2020, doi: 10.1109/ACCESS.2020.2982588.

[33] B. J. Bhatkalkar, D. R. Reddy, S. Prabhu, and S. V. Bhandary, "Improving the Performance of Convolutional Neural Network for the Segmentation of Optic Disc in Fundus Images Using Attention Gates and Conditional Random Fields," *IEEE Access*, vol. 8, pp. 29299–29310, 2020, doi: 10.1109/ACCESS.2020.2972318.

[34] A. O. Asia *et al.*, "Detection of Diabetic Retinopathy in Retinal Fundus Images Using CNN Classification Models," *Electronics (Switzerland)*, vol. 11, no. 17, Sep. 2022, doi: 10.3390/electronics11172740.

[35] C. L. Lin and K. C. Wu, "Development of revised ResNet-50 for diabetic retinopathy detection," *BMC Bioinformatics*, vol. 24, no. 1, Dec. 2023, doi: 10.1186/s12859-023-05293-1.

[36] G. Bogacsovics, J. Toth, A. Hajdu, and B. Harangi, "Enhancing CNNs through the use of hand-crafted features in automated fundus image classification," *Biomed Signal Process Control*, vol. 76, Jul. 2022, doi: 10.1016/j.bspc.2022.103685.

[37] S. R. Shah, S. Qadri, H. Bibi, S. M. W. Shah, M. I. Sharif, and F. Marinello, "Comparing Inception V3, VGG 16, VGG 19, CNN, and ResNet 50: A Case Study on Early Detection of a Rice Disease," *Agronomy*, vol. 13, no. 6, Jun. 2023, doi: 10.3390/agronomy13061633.

[38] D. Hastari, S. Winanda, A. R. Pratama, N. Nurhaliza, and E. S. Ginting, "Application of Convolutional Neural Network ResNet-50 V2 on Image Classification of Rice Plant Disease," *Public Research Journal of Engineering, Data Technology and Computer Science*, vol. 1, no. 2, Feb. 2024, doi: 10.57152/predatecs.v1i2.865.

[39] A. M. Mutawa, S. Alnajdi, and S. Sruthi, "Transfer Learning for Diabetic Retinopathy Detection: A Study of Dataset Combination and Model Performance," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095685.

[40] Radhika KP and Vinay S, "Prediction of Diabetic Retinopathy Using InceptionV3 Model," *International Journal of Advances in Engineering and Management (IJAEM)*, vol. 4, p. 1327, 2022, doi: 10.35629/5252-040713271331.

[41] S. Sheikh and U. Qidwai, "Using MobileNetV2 to Classify the Severity of Diabetic Retinopathy," *International Journal of Simulation Systems Science & Technology*, Mar. 2020, doi: 10.5013/ijssst.a.21.02.16.

[42] E. M. Dogo, O. J. Afolabi, and B. Twala, "On the Relative Impact of Optimizers on Convolutional Neural Networks with Varying Depth and Width for Image Classification," *Applied Sciences (Switzerland)*, vol. 12, no. 23, Dec. 2022, doi: 10.3390/app122311976.

[43] D. Irfan and T. Surya Gunawan, "Comparison Of SGD, RMSProp, And Adam Optimation In Animal Classification Using CNNs," *2nd International Conference on Information Science and Technology Innovation (ICoSTEC)*, pp. 1–7, 2023, doi: 10.35842/icostec.v2i.35.

[44] C. Mohanty *et al.*, "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," *Sensors*, vol. 23, no. 12, Jun. 2023, doi: 10.3390/s23125726.

[45] C. Guo, M. Yu, and J. Li, "Prediction of different eye diseases based on fundus photography via deep transfer learning," *J Clin Med*, vol. 10, no. 23, Dec. 2021, doi: 10.3390/jcm10235481.

[46] Y. Pan *et al.*, "Fundus image classification using Inception V3 and ResNet-50 for the early diagnostics of fundus diseases," *Front Physiol*, vol. 14, 2023, doi: 10.3389/fphys.2023.1126780.

[47] D. Bhatt *et al.*, "Cnn variants for computer vision: History, architecture, application, challenges and future scope," *Electronics (Switzerland)*, vol. 10, no. 20. MDPI, Oct. 01, 2021. doi: 10.3390/electronics10202470.

[48] M. Abdullah, F. berhe Abrha, B. Kedir, and T. Tamirat Tagesse, "A Hybrid Deep Learning CNN model for COVID-19 detection from chest X-rays," *Heliyon*, vol. 10, no. 5, Mar. 2024, doi: 10.1016/j.heliyon.2024.e26938.

[49] Y. Pan *et al.*, "Fundus image classification using Inception V3 and ResNet-50 for the early diagnostics of fundus diseases," *Front Physiol*, vol. 14, 2023, doi: 10.3389/fphys.2023.1126780.

[50] A. Souid, N. Sakli, and H. Sakli, "Classification and predictions of lung diseases from chest x- rays using mobilenet v2," *Applied Sciences (Switzerland)*, vol. 11, no. 6, Mar. 2021, doi: 10.3390/app11062751.

[51] P. N. Srinivasu, J. G. Sivasai, M. F. Ijaz, A. K. Bhoi, W. Kim, and J. J. Kang, "Classification of skin disease using deep learning neural networks with mobilenet v2 and lstm," *Sensors*, vol. 21, no. 8, Apr. 2021, doi: 10.3390/s21082852.

[52] A. Taner, Y. B. Öztekin, and H. Duran, "Performance analysis of deep learning cnn models for variety classification in Hazelnut," *Sustainability (Switzerland)*, vol. 13, no. 12, Jun. 2021, doi: 10.3390/su13126527.

[53] M. Stanojević, D. Drašković, and B. Nikolić, "Retinal disease classification based on optical coherence tomography images using convolutional neural networks," *J Electron Imaging*, vol. 32, no. 03, Nov. 2022, doi: 10.1117/1.jei.32.3.032004.

[54] L. Ali, F. Alnajjar, H. Al Jassmi, M. Gochoo, W. Khan, and M. A. Serhani, "Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures," *Sensors*, vol. 21, no. 5, pp. 1–22, Mar. 2021, doi: 10.3390/s21051688.