Senior Design Project Report

CSE/EEE/ETE 499

Enhancing Ocular Disease Diagnosis in Fundus Images with CNN Models and Deep Learning approaches

Submitted By:

Md. Marop Hossain	2013982042
Tasnia Tasnim	1912533642
Zunayed Islam Pranto	1921609642

Supervisor

Dr. Mohammad Ashrafuzzaman Khan

Assistant Professor

ELECTRICAL AND COMPUTER ENGINEERING

NORTH SOUTH UNIVERSITY

SUMMER 2023

Introduction:

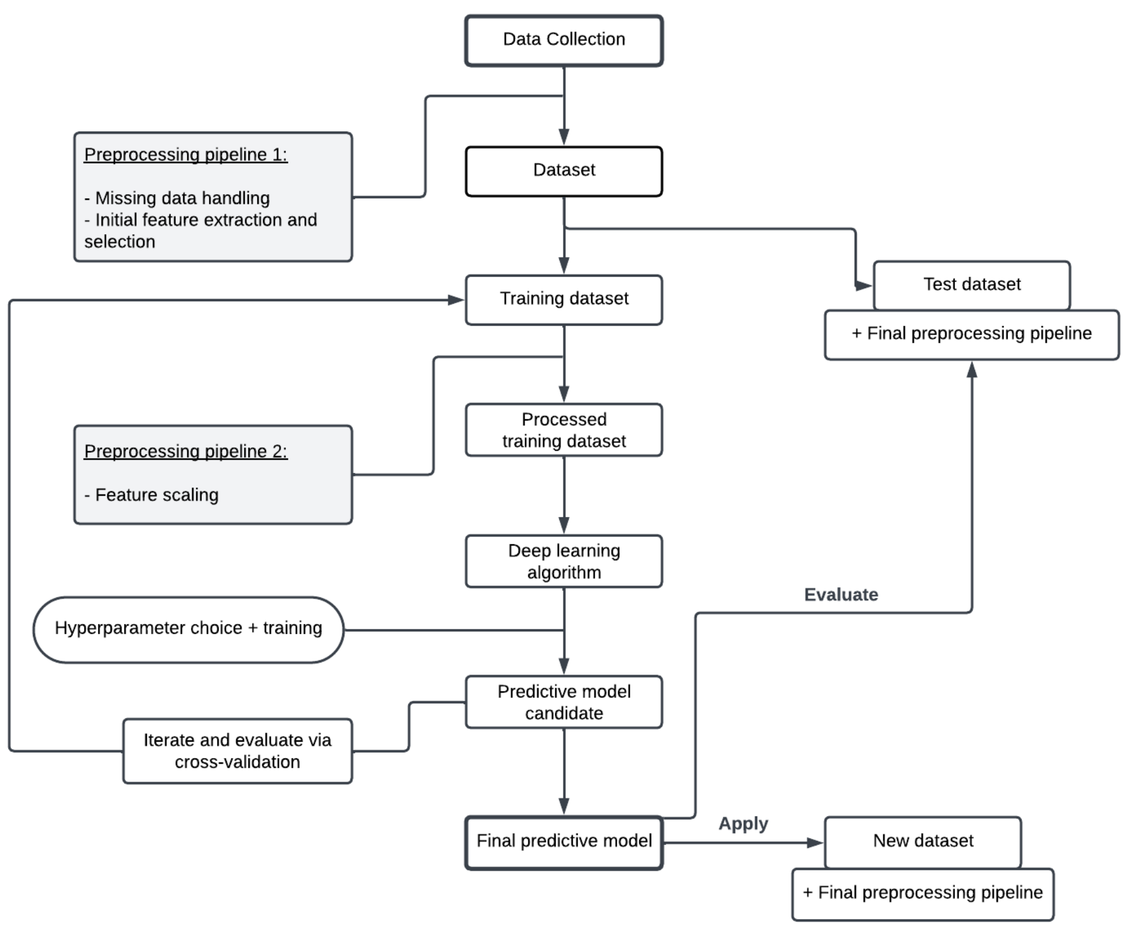
The eye is a vital organ for a human being that plays a critical role compared to other organs of the human body. Eye diseases can cause vision loss or blindness, which can have a profound impact on a person's quality of life. There are a lot of eye diseases, such as **diabetes, glaucoma, cataracts, age-related macular degeneration, hypertension, pathological myopia, and abnormalities.** To prevent such eye diseases, early detection and timely treatment can help reduce the risk of great loss of vision.

Most eye care institutions are not well off in many underdeveloped nations. Furthermore, reliable medical treatment and ophthalmologists are scarce in rural areas. So, it becomes quite tough for the people of rural areas to carry the expenses of better treatment. As the population is increasing, the number of patients with eye diseases is also rapidly increasing. So, it's a community's or government's obligation to improve eye care facilities for its citizens.

A system that can detect eye diseases from retinal fundus images can be developed using digital image processing and machine learning. The system will take the retinal fundus images as input. Then, from the fundus images, the system can extract and simplify the features of specific eye diseases.

In this project, our main focus is to **develop a system using deep learning through CNN that will use the retinal fundus image to identify, extract, and evaluate disease-specific characteristics.** The system will help to early detect the diseases, which allows patients to maintain a good quality of vision while avoiding serious vision loss and blindness.

Model Training and Evaluating process:



Overview of the retinal diseases we work on

Cataract

A cataract is a cloudy area in the lens of your eye. Cataracts are more prone to develop as you become older. In fact, more than half of people in their eighties and nineties have cataracts or have had cataract surgery. At first, you may be unaware that you have a cataract. Cataracts, on the other hand, may cause your vision to become hazy, clouded, or less coloured as time passes. It's possible that you'll have trouble reading or doing other normal chores. The good news is that cataracts can be removed surgically. Cataract surgery is a safe technique that improves vision in patients who have cataracts. The majority of cataracts are age-related and develop as a result of changes that occur in your eyes as you age. Cataracts can form as a result of a variety of factors, such as eye surgery or injury to treat another eye condition (like glaucoma).

Glaucoma

Glaucoma is a series of illnesses affecting the visual cortex, a nerve in the back of the eye that causes vision loss and blindness. Because the symptoms appear gradually, it's likely that you won't even notice them. The only way to tell if you do have glaucoma is to have a full-dilated eye exam. Although there is no cure for glaucoma, early treatment can often prevent future visual loss and protect your vision. Other varieties of glaucoma exist, but open-angle glaucoma is the most prevalent and is what most people hear when they hear term glaucoma. Angle-closure glaucoma and congenital glaucoma are less common kinds of glaucoma.

Diabetic retinopathy

Diabetic retinopathy is an infection of the eye that can cause diabetics to lose their vision and eventually go blind. The blood vessels in the eyes are damaged (the light-sensitive area of tissue located behind of your eye). If you have diabetes, you should have a fully dilated eye test at least once a year. Diabetic retinopathy might manifest itself without signs at first, but catching it early can help you keep your eyesight. Physical activity, a healthy diet, and medication compliance can all prevent you from getting or delay loss of vision.

CNN in Medical Imaging

A retinal eye condition is often diagnosed through a Fundus image taken with specialized equipment. The abnormality may be detected in the eye fundus image as some abnormal impact on the retina. With the development of computer vision and computer-assisted image identification, breakthroughs in medical science have occurred. Through the use of machine learning and neural networks in image recognition, improvements in image categorization and identification have been possible. CNN will ensure nearly as much precision in image recognition as any existing identification technology. In cases where a patient cannot obtain professional guidance for identification of disease or the hospital needs automated assistance, those images can be utilized to identify specific eye diseases. Using CNN-based classification methods, the need for manually segmenting retinal disease zones is eliminated, resulting in a completely automatic classifier. CNN's extensive integration with medical image identification brings up a whole new set of possibilities for progress.

Computer Vision and Implementation of CNN

The ever-expanding breadth of technical progress has resulted in an incalculable number of computational process-based tasks. One of the most prominent examples is computer vision. This method teaches the machine to recognize photos and videos. The emergence of machine learning, sometimes known as "deep learning," has revolutionized computer vision applications by allowing for large-scale computer vision and pattern recognition. The rise of public picture repositories (ImageNet) and improved computing performance in graphics processing units can be credited for this boom in popularity (GPUs). Convolutional neural networks (CNN), a machine learning-based computer vision technology, are swiftly gaining traction among other classic machine learning techniques for image recognition. The fundamental reason for this is that CNN is faster and more accurate than other approaches like Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and Artificial Neural Networks (ANN) (ANN). To emphasize this point, CNN has a high degree of accuracy in detecting altered or manipulated photos, which is the most important advantage in real-world circumstances. The accuracy of multiple convolutional neural network models has recently increased to the point that, in extreme circumstances, CNN can statistically surpass even humans in terms of identifying a specific breed or species. Computational power and data collection will continue to expand at an exponential rate. As a result, the seemingly endless options for improving CNN and deep learning have been significantly explored.

Research Problem

The classification of retinal eye images is a fascinating computer vision topic with many applications in the medical field. Understanding the retinal blood vessels, for example, is critical for eye doctors to assess eye illnesses like glaucoma and hypertension, which can cause vision problems and blindness if left untreated. Much prior research has concluded that comprehending vascular anomalies from retinal images aids medical physicians in diagnosing and treating stroke, brain injury, carotid atherosclerosis, artery disease, and cerebral amyloid angiopathy early. Some data from these trials suggests that having a specialist examine the retinal eye on a regular basis can improve a patient's quality of life.

The human visual retina is a light-sensitive layer that is critical to human vision. The retina of the eye shares certain anatomical similarities with the central nervous system. The fact that the brain's capillary condition affects the retinal blood vessels produces damage to the retinal eye, indicating systemic microvascular damage linked with disorders like hypertension or diabetes.

Retinal image categorization has gotten a lot of attention in the last decade, which has resulted in a lot of papers. Although various approaches have been proposed, recognizing retinal ocular pictures remains a difficult calculation challenge. The fundus diversity of each patient is one of the challenges.

With technological advancements and image analysis, the procedure of disease detection can be automated, and the patient can be referred to a doctor for additional evaluation. Using developments in electronic computer vision and machine learning, a number of clinical decision support systems are developed specifically to identify diabetic retinopathy and age-related macular disease. Although most of these algorithms are capable of performing as well as human experts, many are focused on diagnosing a specific retinal condition. Most of these models use the retinal picture to identify, extract, and analyze disease-specific characteristics. This necessitates a thorough understanding of the condition as well as time spent developing characteristics for the classifiers. We propose to construct a universal classification algorithm that can identify good retinal images from sick retinal images in this research. The proposed system is based on deep learning and can automatically discover characteristics at various levels from a retinal image training dataset. A broad model like this could be beneficial as a first-level screening tool, particularly in rural areas. This would allow for early diagnosis of retinal illnesses, as well as the avoidance of costly travel and testing for those who do not require further consultation. This is beneficial to both the medical world and rural residents. The diagnostic instrument may be handled by semi-skilled technicians, thanks to an improved user interface, solving the problem of a shortage of skilled medical personnel in rural areas. A fundus camera captures the retinal pictures. Despite the fact that fundus cameras are indeed costly, low-cost cameras are being created that are now inexpensive. This is yet another setup fee that will benefit society as a whole.

Machine Learning

Machine Learning is a data analysis method used by Insights. Machine Learning techniques learn from data to uncover hidden patterns without having to be instructed to look for them. The use of machine learning in disease diagnosis has exploded in recent years. These software algorithms perform by identifying patterns in data in photos at various levels and matching them to diseases that are known to exist. As evidenced by the academic literature, supervised learning is being employed for the early recognition and characterization of eye illnesses such as cataracts, conjunctivitis, and diabetic retinopathy. We describe various research findings in the part on related work where machine prognosis is equivalent to that of human experts for particular eye conditions.

Neural Network

A ***CNN (Convolutional Neural Network)*** is a form of neural network which can recognize structural characteristics in a picture. By allowing filtration to slide through the image sequence and perform pattern matching, the CNN is able to capture the sequence at any place across the retina. The stride determines how far the filtration must move across the image as it matches the image pattern. CNN models are made up of self-learning weight matrices that are built into the processing units. Every neuron receives some inputs, does a dot product with weights and biases, and optionally performs an activation function. From raw picture frames on one end to category scores on the other, the entire network employs a single algorithm as follows. Because the inputs are images, the CNN design can encode specific attributes. This minimizes the number of variables in the network and makes the forward algorithm more efficient to implement.

Objective

The goal of this study is to create a proper ensemble learning model which can determine between healthy and diseased retinal pictures. The proposed system is based on deep learning and can automatically discover characteristics at various levels from a retinal image training dataset.

The objectives of this project are:

- ✓ *To evaluate the deep learning model using the ensemble learning technique.*
- ✓ *To offer recommendations on improving the deep learning model.*
- ✓ *To detect multiple eye diseases using a single model.*

Proposed Model:

Input Data

Data Collection:

For this project, we have used the **Ocular Disease Intelligent Recognition (ODIR) dataset**. It is one of the largest publicly available multiclass ocular disease detection datasets in the world. This dataset was compiled by Shanggong Medical Technology Co. by collecting fundus images from different hospitals in China. The fundus images in **this dataset are split into eight different ocular disease classification categories**. These categories include **seven disease classes**: **diabetes (D), cataract (C), glaucoma (G), age-related muscular degeneration (A), myopia (M), hypertension (H),** and **other abnormalities or diseases (O)**. In total, this dataset **contains 5000 cases of color fundus photographs (CFPs)**, and it is split into training and testing subsets. Roughly *8713* cases (left eye images and right eye image) are used for training, and the rest are used for testing.

For this experiment, we used 4 classes:

- **Normal**
- **Glaucoma**
- **Cataract**
- **Diabetic Retinopathy**
  - ✓ *mild non-proliferative retinopathy*
  - ✓ *proliferative retinopathy*
  - ✓ *diabetic retinopathy*
  - ✓ *moderate non proliferative retinopathy*

Dataset Sample:

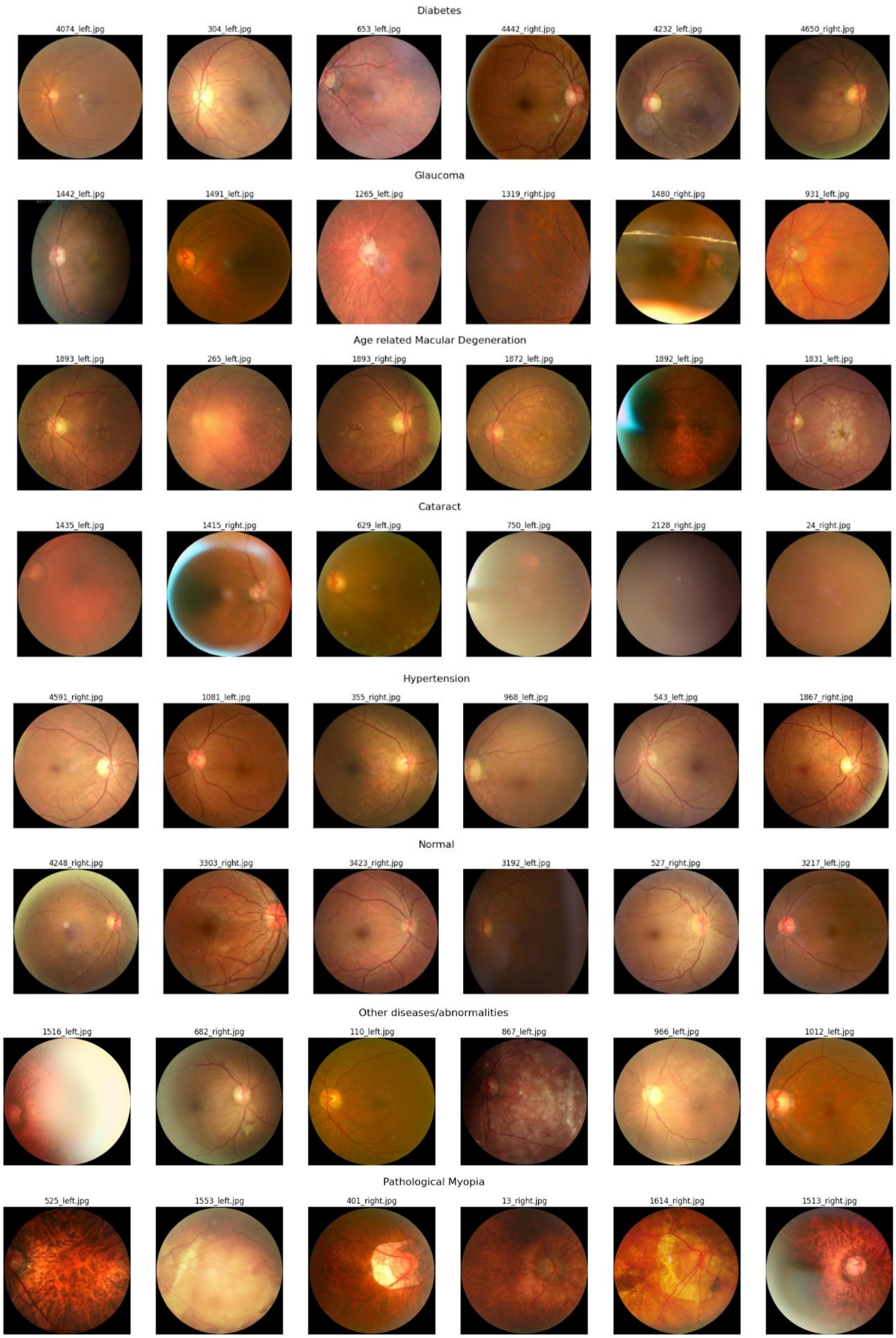


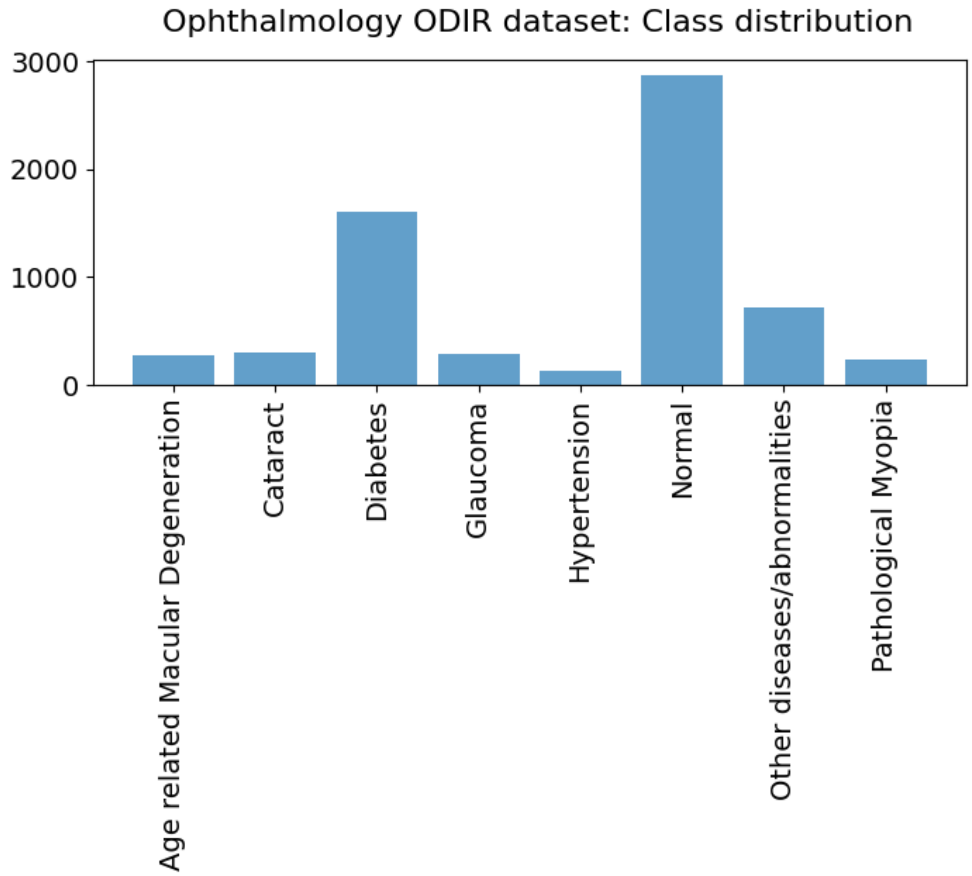
Figure: A look at the fundus images of the ODIR dataset

Our datasets were created by combining four different illnesses, and each class included **8713 photos**. Thus, it is quite difficult to display an appropriate sample from the database. A total of **48 images** were able to be combined into the above figure from our dataset; we attempted to display at least 48 images. It is evident from this sample image that we collected photographs from multiple disease classifications to generate our dataset. Approximately six photos are presented in each disease category.

Distribution of images in the ODIR dataset:

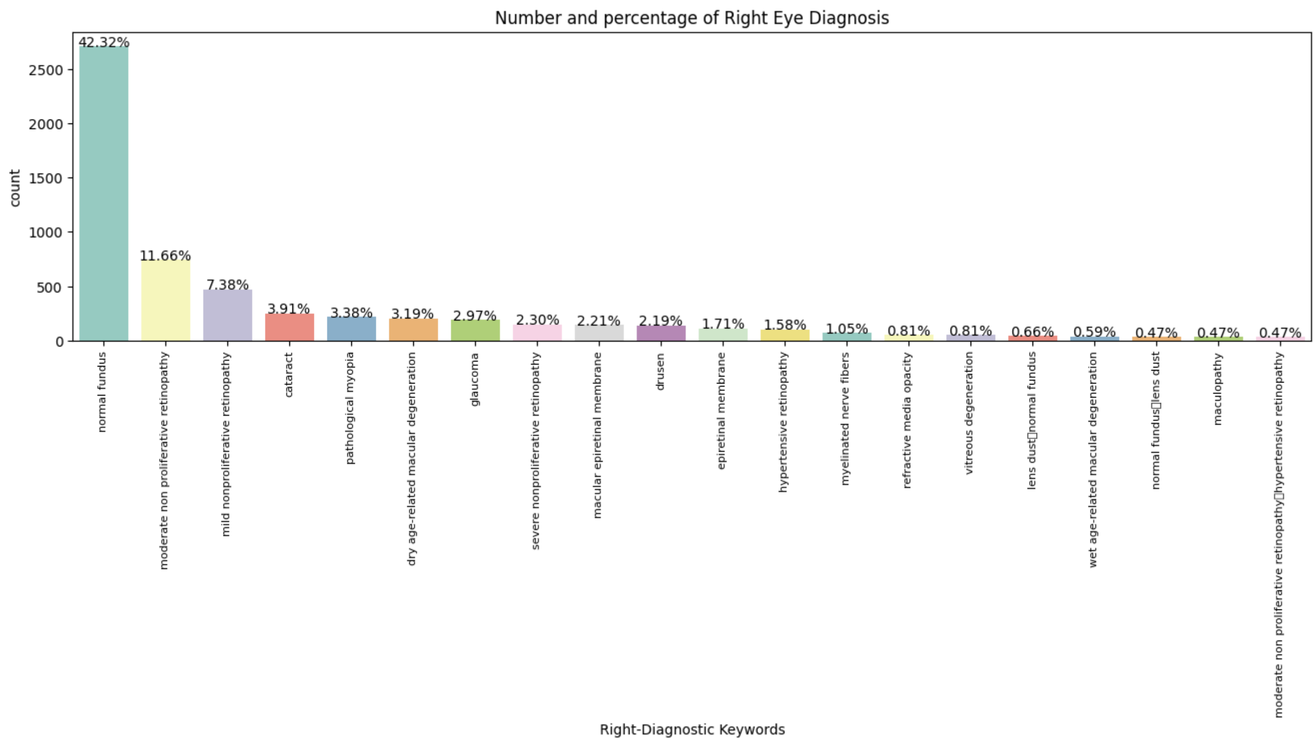
No. of Classes	Labels	Training cases
1	*Normal (N)*	*1135*
2	*Diabetes (D)*	*1131*
3	*Glaucoma (G)*	*207*
4	*Cataract (C)*	*211*
5	*Age - related macular degeneration (A)*	*171*
6	*Hypertension (H)*	*94*
7	*Pathological myopia (M)*	*177*
8	*Other diseases / abnormalities*	*944*

Ophthalmology ODIR Dataset Class Distribution:



The bar chart of the **distribution of ophthalmology ODIR dataset** class distribution according to the number of patients with different types of eye diseases. ***The most common eye disease in the dataset is normal fundus, followed by age-related macular degeneration, cataracts, and glaucoma.*** Other common eye diseases include hypertensive retinopathy, diabetic retinopathy, and drusen. The chart shows that a significant number of patients in the dataset have preventable or treatable eye diseases, such as cataracts and age-related macular degeneration.

Number and Percentage of Right Eye Diagnoses:



The graph is divided into two sections: the **x-axis shows the different types of diagnoses**, and the **y-axis shows the number and percentage of diagnoses for each type**. The ***most common right eye diagnosis is normal fundus****, followed by* ***cataracts, age-related macular degeneration, and glaucoma***. Other common diagnoses include **mild and severe non-proliferative retinopathy, diabetic retinopathy, drusen, and epiretinal membrane**. The graph shows that a significant number of right eye diagnoses are preventable or treatable, such as cataracts and age-related macular degeneration. It is important to have regular eye exams to detect and treat these conditions early.

Most common right eye diagnoses, are:

- **Normal fundus:** This means that the eye is healthy and there are no signs of disease.
- **Cataracts:** This is a clouding of the lens of the eye. Cataracts are common in older adults, but they can also occur in younger adults and children. Cataracts can cause blurred vision, difficulty seeing at night, and increased sensitivity to light.
- **Age-related macular degeneration:** This is a condition that causes the macula, the central part of the retina, to deteriorate. Age-related macular degeneration is the leading cause of vision loss in people over the age of 65. It can cause blurred vision, difficulty seeing fine details, and difficulty reading.
- **Glaucoma:** This is a group of eye diseases that damage the optic nerve and lead to vision loss. Glaucoma is the second leading cause of blindness worldwide. It can cause peripheral vision loss, which can eventually lead to blindness.

➢ Cataracts

Cataracts are caused by changes in the protein that makes up the lens of the eye. As we age, the protein in the lens can break down and clump together, forming a cataract. Cataracts can also be caused by certain medical conditions, such as diabetes, eye trauma, and prolonged use of corticosteroids.

➢ Age-related macular degeneration

Age-related macular degeneration (AMD) is caused by damage to the macula, the central part of the retina. The macula is responsible for sharp central vision, which is needed for activities such as reading, driving, and recognizing faces.

***There are two main types of AMD: dry AMD and wet AMD*. *Dry AMD is the more common form of AMD*.** *It causes the macula to thin and dry out over time. Wet AMD occurs when abnormal blood vessels grow under the macula. These blood vessels can leak blood and fluid, which can damage the macula and lead to rapid vision loss.*

- **Glaucoma**

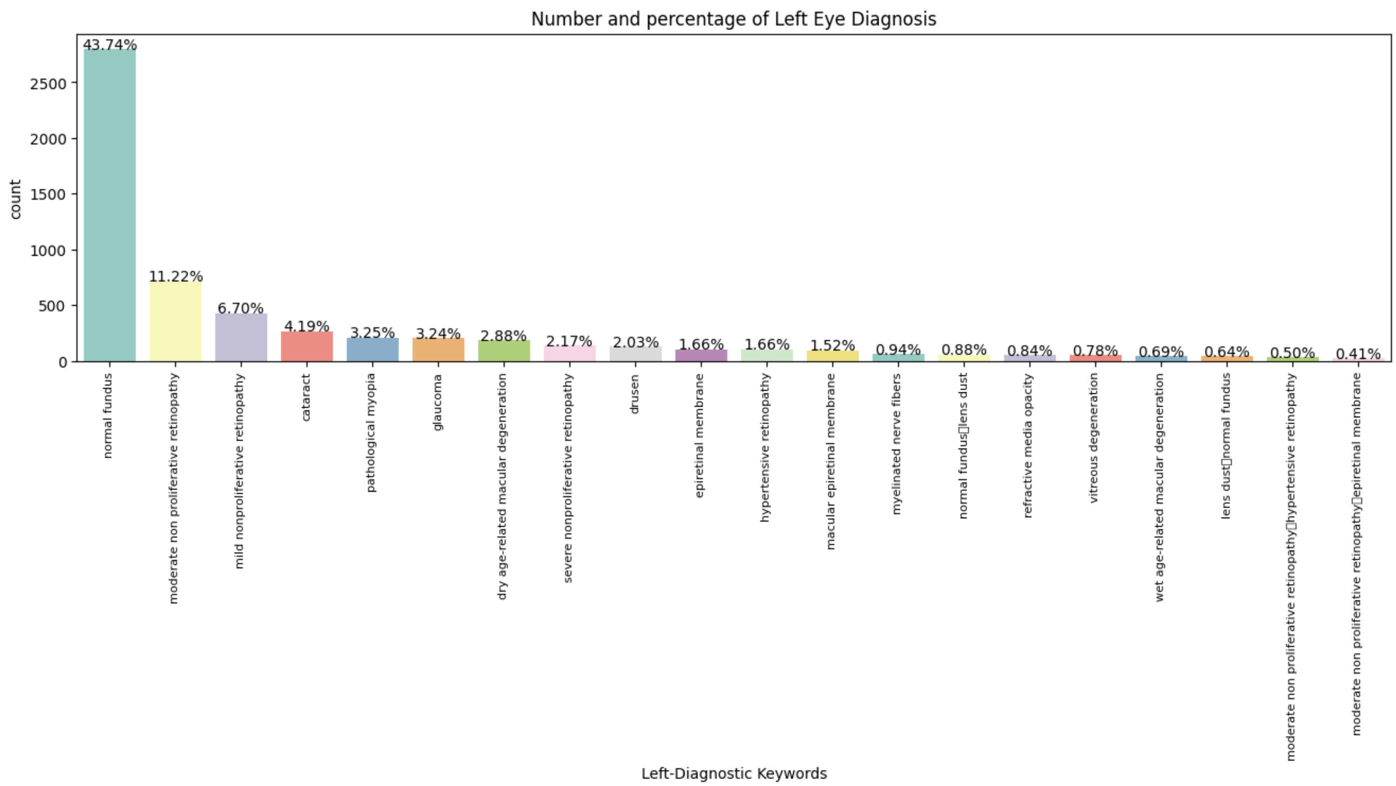
Glaucoma is a group of eye diseases that damage the optic nerve. The optic nerve is responsible for carrying visual signals from the eye to the brain. Glaucoma can cause peripheral vision loss, which can eventually lead to blindness.

There are two main types of glaucoma: primary open-angle glaucoma and angle-closure glaucoma. Primary open-angle glaucoma is the more common form of glaucoma. It is caused by a gradual buildup of pressure in the eye. Angle-closure glaucoma is a less common but more serious form of glaucoma. It is caused by a sudden increase in pressure in the eye.

Other common right eye diagnoses:

- ✓ **Mild and severe non-proliferative retinopathy:** *This is a type of diabetic retinopathy, which is damage to the blood vessels in the retina caused by diabetes. Non-proliferative retinopathy is the most common type of diabetic retinopathy.*
- ✓ **Diabetic retinopathy:** *This is damage to the blood vessels in the retina caused by diabetes. Diabetic retinopathy can lead to vision loss and blindness.*
- ✓ **Drusen:** *These are small yellow deposits that form under the retina. Drusen are common in older adults, but they can also occur in younger adults. Drusen can lead to AMD.*
- ✓ **Epiretinal membrane:** *This is a thin layer of tissue that forms on the surface of the retina. Epiretinal membranes can cause blurred vision and distorted vision.*

Number and Percentage of Left Eye Diagnoses:

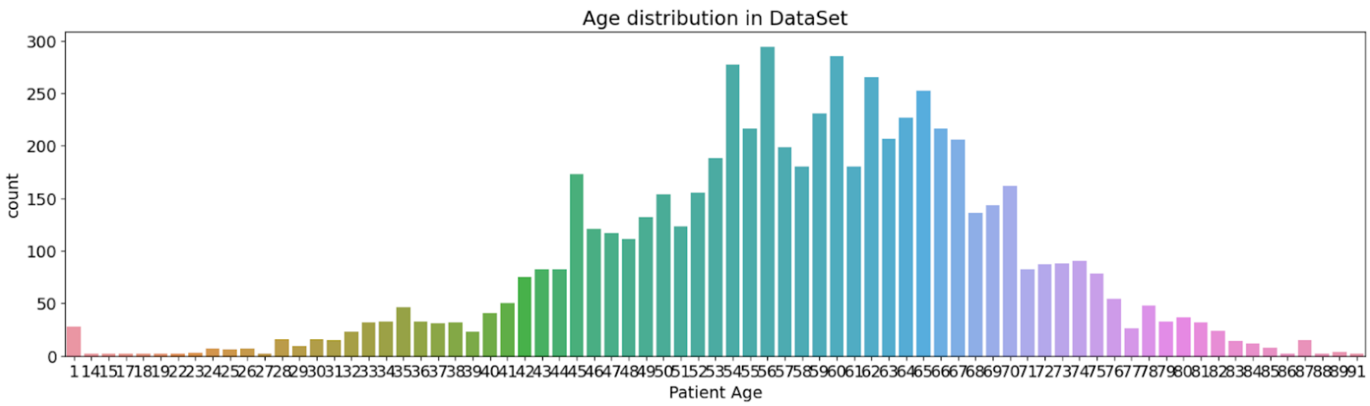


The graph is divided into two sections: **the x-axis shows the different types of diagnoses**, and the **y-axis shows the number and percentage of diagnoses for each type**. The ***most common left eye diagnosis is normal fundus, followed by cataracts, mild non-proliferative retinopathy, and moderate non-proliferative retinopathy***. Other common diagnoses include **age-related macular degeneration, glaucoma, severe non-proliferative retinopathy, drusen, and epiretinal membrane**. The graph shows that a significant number of left eye diagnoses are preventable or treatable, such as cataracts and age-related macular degeneration. It is important to have regular eye exams to detect and treat these conditions early.

Most common left eye diagnoses:

- **Normal fundus:** This means that the eye is healthy and there are no signs of disease.
- **Cataracts:** This is a clouding of the lens of the eye. Cataracts are common in older adults, but they can also occur in younger adults and children.
- **Non-proliferative retinopathy:** This is a type of diabetic retinopathy, which is damage to the blood vessels in the retina caused by diabetes. Non-proliferative retinopathy is the most common type of diabetic retinopathy.
- **Age-related macular degeneration:** This is a condition that causes the macula, the central part of the retina, to deteriorate. Age-related macular degeneration is the leading cause of vision loss in people over the age of 65.
- **Glaucoma:** This is a group of eye diseases that damage the optic nerve and lead to vision loss. Glaucoma is the second leading cause of blindness worldwide.

Age Distribution of People with Eye Diseases:



The histogram shows that the most affected are those in the age group of 45-70. This is likely due to the fact that many eye diseases, such as ***cataracts, glaucoma, and age-related macular degeneration, are more common in older adults***. There are a number of reasons why older adults are more at risk for eye diseases**. One reason is that the eye changes with age**. The lens of the eye can become cloudy, the muscles that control the eye can weaken, and the blood vessels in the eye can become damaged. These changes can make it more difficult for the eye to focus and see clearly.

Another reason why older adults are more at risk for eye diseases is that they are more likely to have other medical conditions that can increase their risk of eye problems. For example, people with diabetes are more likely to develop diabetic retinopathy, a condition that can damage the blood vessels in the retina. People with high blood pressure are more likely to develop hypertensive retinopathy, a condition that can also damage the blood vessels in the retina. ***It is important to have regular eye exams, especially if you are over the age of 45***. Eye exams can detect eye diseases early, when they are most treatable. Early treatment can help to prevent vision loss and blindness.

Data Preprocessing:

The dataset used for experimentation in ODIR, representing 5000 patients' left and right eyes from different types of cameras, was categorized into eight ocular diseases: **diabetic retinopathy, age-related muscular degeneration, glaucoma, cataract, hypertension, pathological myopia,** and **other abnormalities**.

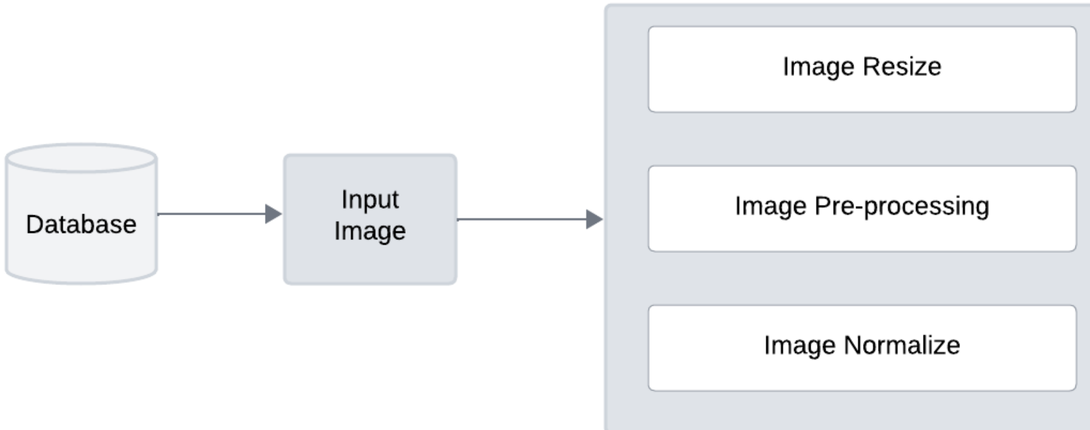
In order to extract features for the dataset we used to recognize a variety of objects from a single image, we needed to load some **necessary libraries, resize images, and initialize the directory**. Before data preprocessing can begin, we imported **OS, TensorFlow, NumPy as np, glob, matplotlib.pyplot as plt, and other libraries**. Once the libraries had been imported, we set up the dataset's directory so that the data can be obtained by the model, and details can be extracted before images can be resized.

For this experiment, we selected a total of **4201 fundus images labelled Normal, 594 fundus images of cataract, 3411 fundus images of diabetes retinopathy,** and **616 fundus images of glaucoma**. The unbalance in the dataset is clear, with the normal and diabetes retinopathy classes representing the majority of the total images. Thus, the image dataset has been prepared **with four categories of images consisting of both disease- and disease-free eyes**. The images collected are in different shapes and sizes. Therefore, we resize the images to **224x224.**

The dataset was shared between three subsets, where **70% was used for training, 20% for validation,** and **10% for tests.** To keep the same ratio for each application, we **stratified a train-test split**. We also needed to face the unbalanced problem that generated bias towards **diabetes retinopathy**, and **normal classes** because they were highly represented, which led to over-predicting these two categories. Sampling strategies solved this issue by balancing the distribution in the dataset. We used the **image augmentation technique** to increase the size of our dataset by **rotating 90 degrees, flipping vertically,** and **filipping horizontally**.

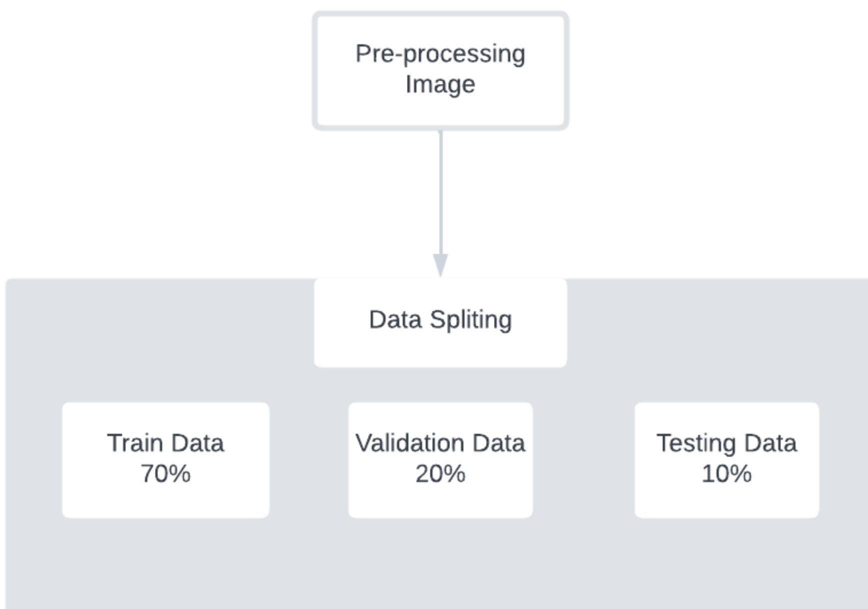
Image Resizing:

To get optimal performance, each type of CNN architecture necessitates a different image size to resize or distort our image from the one-pixel grid to another. Image resizing is applied based on the model architecture. The input ***shape is set as (224,224) in our proposed selected model***.



Splitting Train and Validation Set:

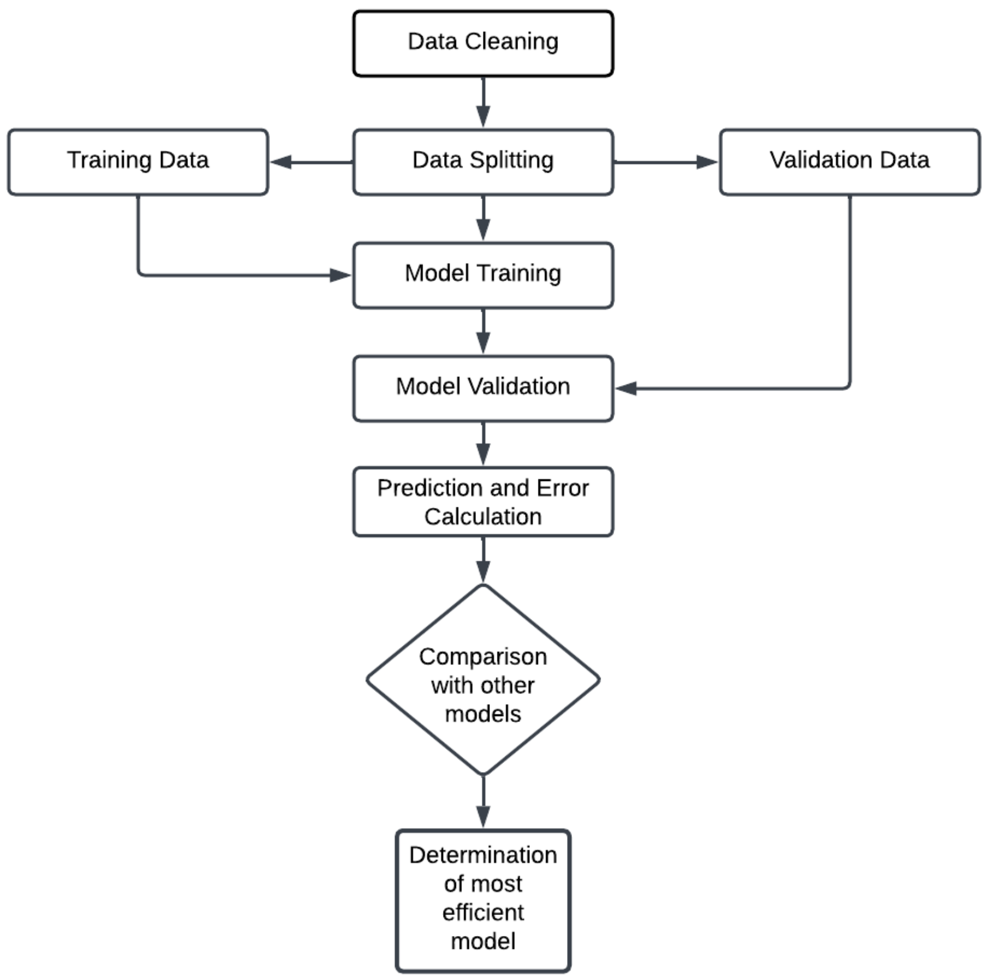
The data set is divided into two parts, a training set and a validation set. In order to determine the accuracy of our CNN models, we will train them using batch size and epochs on the training set, and then evaluate them on the validation set. A comparison of models will take into account precision, recall, f1-score, and accuracy. Our multiple eye diseases dataset included 5,000 fundus images that were **split into three subsets: 70% for training, 20% for validation and 10% for testing**. We tested the accuracy of the model on a subset of these images. Before fitting the model, we shuffled the images so that the model does not memorize anything if the images belonging to same category are fed in a consequent manner.



Data distribution table:

4-classes	Before creating dataset	After creating new dataset	From Left Eye	From Right Eye
Cataract	*594*	*588*	*304*	*290*
Normal	*4201*	*4143*	*2100*	*2101*
Diabetes	*3411*	*3369*	*1673*	*1783*
Glaucoma	*616*	*613*	*332*	*284*
Total =	***8822***	***8713***	***4409***	***4458***

Process of Model Training:



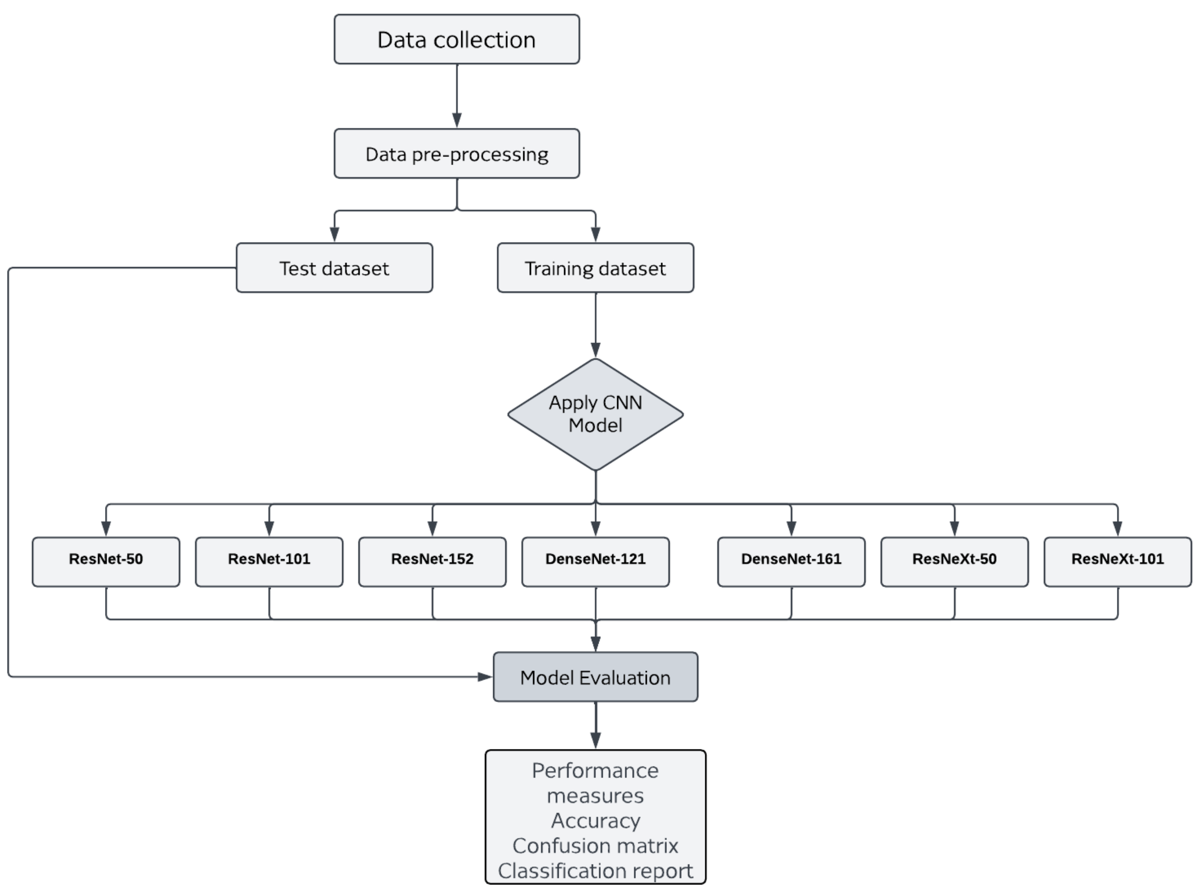
Model Selection:

We present CNN and seven deep learning-based models for targeted ocular disease diagnosis. For this project, we trained cutting-edge classification algorithms such as **ResNet-50, ResNet-101, ResNet-152, DenseNet-121, DenseNet-161, ResNeXt-50,** and **ResNeXt-101** on the ODIR dataset consisting of **5000 fundus images** that belong to **4 different classes**. Each of these classes represents a different ocular disease.

Our CNN model architecture involves **convolutional layers**, **activation functions** (**ReLU** and **softmax**), **max-pooling layers**, **dropout layers**, **hidden fully connected layers**, and the **output layer (for 4-classes)**. The model also uses the **Adam** and **RMSProp optimizers** and a **categorical cross-entropy loss function**. Hyperparameters such as **batch size**, **learning rate**, and **number of epochs** are also used.

Then what are we going to do with that last layer that every particular model has to remove? because the pre-trained model has thousands of layers. We worked on four categories of images. So, we added the **four output layers**. This is called the art of algorithmic transfer learning. In all the models, we use **0.001 as the learning rate** for both the **Adam** and **RMSprop optimizers**. To reduce the error rate, **we used Adam and RMSprop optimizer equations**, which update network weights iteratively in the training dataset by replacing the **stochastic gradient descent method**. **Adam** and the **RMSprop optimizer** play a key role in **minimizing the training error**. After that, the model was run for different sizes of epochs. We also use a **training callback function** called "**early stopping**" that stops the training process if the **validation loss** of the model does not decrease for more than 20 epochs.

To perform our work, all the deep learning models have been trained with **GPU support**. Now, we needed to collect our required images and pre-process them. After that, we applied different deep learning algorithms and analyzed their results.



Model Explanation:

ResNeXt-50:

ResNeXt-50 is a deep learning model that is used to classify images. It is based on the Res-Net architecture, which is a type of neural network that is good at learning long-range dependencies in images. ResNeXt-50 has 50 layers, and each layer is made up of a set of residual blocks. Residual blocks are a type of layer that allows the model to learn more complex relationships in the data. ResNeXt-50 is trained on a large dataset of images, and it can be used to classify a wide variety of objects and scenes. It is particularly well-suited for tasks such as image classification, object detection, and image segmentation. One of the advantages of ResNeXt-50 is that it is very efficient. It can run quickly and accurately, even on mobile devices. This makes it a good choice for applications where performance is important, such as real-time image recognition.

ResNeXt-101:

ResNeXt-101 is a deep learning model that is based on the Res-Net architecture. It is a convolutional neural network (CNN) that has been trained on a large dataset of images to classify them into different categories. ResNeXt-101 is particularly well-suited for image classification tasks, such as identifying objects, scenes, and faces. ResNeXt-101 works by extracting features from images using a series of convolutional layers. These layers learn to identify different patterns in the images. The features extracted by the convolutional layers are then fed into a fully connected layer, which makes a prediction about the category of the image. ResNeXt-101 is a very powerful model, and it has achieved state-of-the-art results on many image classification benchmarks. It is also a relatively efficient model, meaning that it can be trained and used on a variety of devices.

ResNet-50:

ResNet-50 is a deep convolutional neural network (CNN) architecture that consists of 50 layers. It was introduced to address the problem of vanishing gradients in deep neural networks. ResNet-50 uses skip connections, known as residual connections, to enable the flow of gradients directly from earlier layers to later layers. This allows for easier training of very deep networks and helps to alleviate the degradation problem. ResNet-50 has been widely used and has achieved excellent performance on various computer vision tasks, including image classification and object detection.

ResNet-101:

ResNet101 is a type of convolutional neural network (CNN) that is used for image classification. It is a deep learning model that was developed by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in 2015. ResNet101 has 101 layers, which is more than most other CNNs. This allows it to learn more complex features from the images it is trained on. ResNet101 works by breaking an image down into smaller and smaller pieces. It then uses a series of filters to extract features from each piece. The features are then combined to form a representation of the image. ResNet101 is trained on a large dataset of images that have been labeled with the objects they contain. This allows the model to learn to recognize the objects in new images. Once ResNet101 is trained, it can be used to classify new images. To do this, the model takes an image as input and produces a probability distribution over the different classes. The class with the highest probability is the predicted class of the image. ResNet101 is a very powerful CNN and has been shown to achieve state-of-the-art results on a variety of image classification tasks. It is often used in medical imaging applications, such as the diagnosis of diseases.

ResNet-152:

ResNet152 is a deep learning model that is used for image classification. It is a convolutional neural network (CNN), which is a type of neural network that is well-suited for image processing tasks. ResNet152 is made up of 152 layers, which are arranged in a sequential manner. Each layer takes the output of the previous layer as input and produces a new output. The first few layers of ResNet152 extract low-level features from the input image, such as edges and corners. The subsequent layers extract higher-level features, such as shapes and objects. The final layer of ResNet152 produces a prediction of the image class. ResNet152 has been trained on a large dataset of images, and it has been shown to be very accurate at classifying images. It is often used as a benchmark model for image classification tasks.

DenseNet-121:

DenseNet-121 is another CNN architecture that is known for its dense connectivity pattern. In DenseNet, each layer is directly connected to every other layer in a feedforward fashion. This dense connectivity helps in improving feature reuse, reducing the number of parameters, and enhancing gradient flow. DenseNet-121 has 121 layers and has been shown to achieve high accuracy on image classification tasks while requiring fewer parameters compared to other models.

DenseNet-169:

DenseNet-169 is a convolutional neural network (CNN) architecture that belongs to the family of DenseNet models. It was introduced by Huang et al. in 2017 as an extension to the original DenseNet architecture. DenseNet stands for "Densely Connected Convolutional Networks," and it is known for its dense connectivity pattern within the layers. Unlike traditional CNNs where layers are connected in a sequential manner, DenseNet introduces dense connections between layers. This means that each layer in DenseNet receives input not only from the previous layer but also directly from all preceding layers. This dense connectivity facilitates feature reuse, encourages gradient flow, and enables efficient information propagation through the network.

Model Evaluation:

Evaluate Accuracy Matrix: for measuring goodness, we will use some metrics relating to accuracy in determining whether a particular image represents a disease. After all, we have used classification models. So, we will use the most widely used metrics for classification problems.

Training and Test Accuracy: while training our models using the training data, we will find out how much model is learning from that training dataset. The main purpose of training accuracy is to extract the hyperparameters and to check whether our models have over-fitting or underfitting issues.

$$\textit{Accuracy} = \frac{\textbf{True Positive }(\boldsymbol{TP})+\textbf{True Negative }(\boldsymbol{TN})}{\textbf{True Positive }(\boldsymbol{TP})+\textbf{True Negative }(\boldsymbol{TN})+\textbf{False Positive }(\boldsymbol{FP})+\textbf{False Negative }(\boldsymbol{FN})}$$

When we are done with training our models using our training dataset and we have also cross-checked that our models are doing well in the validation dataset, only then can we go for the test accuracy, which is the final accuracy of our models. When mentioning accuracy in our report, we mean test accuracy.

Precision: sometimes accuracy alone is not enough. We just cannot say that our model is very accurate by only looking at the accuracy, because, in this project, we have to classify both diseases and non-disease correctly. In terms of deep learning, we can say that those who have a disease are called ‘positive’ and those who do not have a disease are called ‘negative’. Precision gives a clear view of how many of the positive-meaning disease patients are identified correctly among the entire dataset.

$$\textit{Precision} = \frac{\textit{True Positive}\ (TP)}{\textit{True Positive}\ (TP) + \textit{False Positive}\ (FP)}$$

Recall: sometimes even precision is not enough. For example, if the dataset is highly biased towards one target, recall provides the number of correctly classified true positives, meaning those people who are actually diseased and our model has predicted that they are diseased.

$$\textit{Recall} = \frac{\textit{True Positive}\ (TP)}{\textit{True Positive}\ (TP) + \textit{False Negative}\ (FN)}$$

This recall is the most important metric of our project, because if we have a poor recall, then our model can predict a diseased person incorrectly.

F1-Score: The F1 score is called the harmonic mean of precision and recall. So, if someone claims to give equal priority to precision and recall, then he/she can focus on the F1 score. In this study, our second highest priority after the recall is the F1 score.

$$\textit{F1-score} = \frac{2 \times \textit{Precision} \times \textit{Recall}}{\textit{Precision} + \textit{Recall}}$$

Confusion Matrix: The performance of classification models has been assessed by the confusion matrix for a given dataset. Only if the true values of the test data were known could it be determined. A matrix in which predicted and actual values are separated, as well as the total number of forecasts, has two dimensions. A true value is determined by the observational data, whereas a projected value is determined by the model.

Project Mitigation Strategies:

The following are some specific project mitigation strategies for enhancing ocular disease diagnosis in fundus images with CNN models and deep learning approaches:

- **Data collection and preprocessing:**

✓ Collaborate with eye hospitals and clinics to collect a large and diverse dataset of fundus images.
✓ Use a variety of sources to collect data, such as clinical trials, public datasets, and social media.
✓ Carefully curate the dataset to ensure that it is high quality and free from errors.
✓ Preprocess the data to remove any artifacts and normalize the pixel values.

- **Model selection and training:**

✓ Experiment with different CNN model architectures and hyperparameters to find the best combination for the specific dataset.
✓ Use transfer learning to initialize the model weights with the weights of a pre-trained model. This can help to reduce the amount of data that is needed to train the model and improve its performance.
✓ Train the model on a large and diverse training dataset.
✓ Use a validation dataset to evaluate the performance of the model during training and to prevent overfitting.

- **Model evaluation:**

- ✓ Evaluate the performance of the model on a held-out test dataset to ensure that it is not overfitting to the training data.
- ✓ Evaluate the model on a variety of metrics, such as accuracy, precision, recall, and F1-score.
- ✓ Look at the confusion matrix to identify any potential biases in the model.

- **Model deployment:**

- ✓ Deploy the model to a production environment that is secure and scalable.
- ✓ Monitor the performance of the model in production and retrain it if necessary.
- ✓ Use the model in a way that respects the privacy of patients.

Challenges:

- **Data collection:**

  Collecting a large and diverse dataset of fundus images can be challenging. This is because ocular diseases are relatively rare, and patients may be reluctant to share their medical images.

- **Data preprocessing:**

  Fundus images can be noisy and difficult to interpret. This can make it difficult to develop deep learning models that can accurately classify the images.

- **Model selection and training:**

  Choosing the right model architecture and hyperparameters is important for achieving good performance. This can be a challenging task, especially for researchers who are new to deep learning.

- **Model evaluation:**

  It is important to evaluate the performance of the model on a held-out test dataset to ensure that it is not overfitting to the training data. This can be difficult to do if the dataset is small.

- **Model deployment:**

  Deploying the model to a production environment can be challenging. This is because it is important to ensure that the model is secure and scalable.

- **Ethical considerations:**

  It is important to consider the ethical implications of deploying a deep learning model for ocular disease diagnosis. For example, it is important to ensure that the model is fair and unbiased, and that it does not discriminate against any particular group of people. It is also important to ensure that the model is used in a way that respects the privacy of patients.

Future work:

- **Developing more accurate and robust deep learning models:**
  - ✓ Researchers can continue to develop new deep learning model architectures and train them on larger and more diverse datasets. This will help to improve the accuracy and robustness of deep learning models for ocular disease diagnosis.
- **Using multi-modal data:**
  - ✓ Deep learning models can be trained on a variety of data modalities, such as fundus images, OCT images, and patient medical records. Using a multi-modal approach can improve the performance of the model.
- **Using ensemble learning:**
  - ✓ Ensemble learning is a machine learning technique that combines the predictions of multiple models to produce a more accurate prediction. Ensemble learning can be used to improve the performance of deep learning models for ocular disease diagnosis.
- **Developing deep learning models for specific ocular diseases:**
  - ✓ Researchers can develop deep learning models that are specifically designed to diagnose specific ocular diseases, such as diabetic retinopathy, glaucoma, and age-related macular degeneration. This could lead to the development of more accurate and efficient diagnostic tools for these diseases.
- **Making deep learning models more accessible to clinicians:**
  - ✓ Researchers can develop user-friendly interfaces for deep learning models that make them easy for clinicians to use. This could help to accelerate the adoption of deep learning in clinical practice.

conclusion:

In conclusion, deep learning has the potential to revolutionize the field of ocular disease diagnosis. Deep learning models can be trained to accurately classify fundus images and identify a variety of ocular diseases. This could lead to the development of more efficient and accurate diagnostic tools that can help to improve the lives of patients with ocular diseases.

However, there are still some challenges that need to be addressed before deep learning models can be widely deployed in clinical practice. One challenge is collecting a large and diverse dataset of fundus images. Another challenge is developing models that are robust and accurate, even in the presence of noise and artifacts in the images. Additionally, it is important to ensure that deep learning models are fair and unbiased, and that they are used in a way that respects the privacy of patients.

Despite these challenges, the potential benefits of using deep learning for ocular disease diagnosis are significant. By developing and deploying deep learning models, researchers and developers can help to ensure that patients receive the care they need as early as possible.