Multi-Class Eye Disorder Classification using CNN

MSc Research Project
MSc in Data Analytics

Shahrukh
Student ID: 23135379

School of Computing
National College of Ireland

Supervisor: Dr Chirstian Horn

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Shahrukh
Student ID:	23135379
Programme:	MSc Data Aanlytics
Year:	2024
Module:	MSc Research Project
Supervisor:	Dr Chirstian Horn
Submission Due Date:	12/08/2024
Project Title:	Multi-Class Eye Disorder Classification using CNN
Word Count:	8217
Page Count:	23

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Shahrukh
Date:	16th September 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Multi-Class Eye Disorder Classification using CNN

Shahrukh
23135379

Abstract

The sense of vision in humans is critical, and the eyes are very important in everyday life. Different conditions of eye disorder could cause blindness, like, as cataract disease, the disease AMD also known as Age-Related Mascular Degeneration, Diabetic Retinopathy (DR) disease, pathological myopia, and glaucoma disorder, these all could damage the visual system leading to blindness. Early diagnosis and accurate detection are crucial to prevent the loss of sight. Traditional ways of diagnosing them are time-consuming with high error rates hence there is a need for the development of better diagnostic tools. This research considers the implementation of deep learning specifically CNN that can help enhance accuracy and speed in ocular disease diagnosis using fundus images. The paper seeks to review existing methodologies critically, construct the CNN model, employ data preprocessing and augmentation approaches as well as assess their influence on model performance to make multiple eye disease diagnoses more accurate and faster resulting in improved patient outcomes. After the preprocessing and data augmentation, several experiments are performed to increase the accuracy of disease prediction. The final approach achieved an accuracy of 81% for the multi-class disease classification for ocular disorders.

1 Introduction

The human body consists of many crucial organs that all have their importance for the proper functioning of the body. There are five core senses utilized by human beings. Among them, the sense of sight is paramount, contributing to approximately 80% of our observation of the environment. The eyes, being the organs of vision, are crucial for daily functioning and quality of life. Consequently, protecting the eyes from diseases that could harm vision is of utmost importance. Many diseases could affect the proper functioning of our eyes and can lead to blindness and serious effects on our sight.

According to WHO report on vision 2019 , cataract is one of the key reasons for loss of sight that has affected 100 million people Qiao et al. (2017) . The catract is a minor, opaque section in the lens of the eye that becomes larger over time and may lead to blurred vision if it gets enough large it causes total vision loss. Age Related Mascular Degeneration also known as (AMD) , is the third largest cause and has affected over 8 million people Ganesh et al. (2024). AMD affects the part of the optic nerve that is known as the macula, responsible for central vision activities and it is common in elderly people. Glaucoma disease is the fourth-largest cause of blindness and affects 4 million people Pandey et al. (2020). It is a nerve disease that is associated with damage to the nerve that connects the eyes to the brain. This disease is common in diabetic type 2 patients and causes irreversible loss of sight.

Retinal disease leads to loss of vision if it is not detected at the earlier stages as detection of disease through traditional ways is a time-consuming process due to the complexity of this major part of the eye.The retina is a nerve layer located behind the eye responsible for detecting light. Diabetic Retinopathy, DR is also a well-known cause of blindness that affects over 4 million people it is produced by diabetes in the blood vessels of the retina and it is associated with diabetic patients Behera and Chakravarty (2020). Hypertensive retinography is caused by high blood pressure which causes problems in the blood flow from the vessels in the retina and may lead to permanent sight loss if it is not treated timely. Pathological myopia is a severe condition that develops due to changes in anatomical eye structure and causes irreversible blindness if it is not treated timely. Several other conditions cause issues of eyesight and affect the regular working of the eye.

The early diagnosis of all these diseases is vital to keep the patient safe from blindness. Figure 1 illustrates each disease along with the normal retinal image.

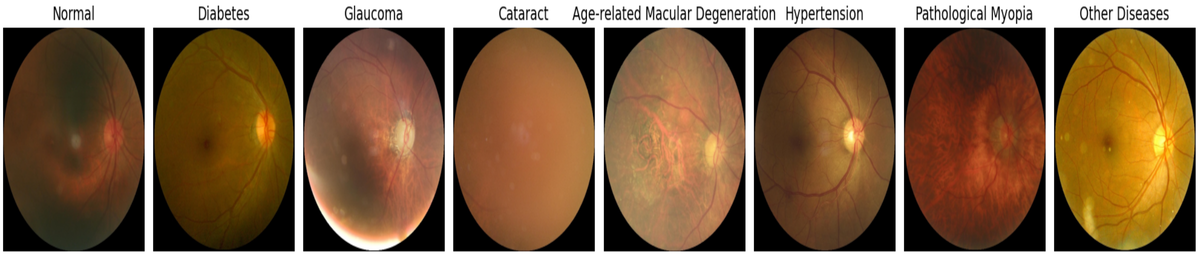

Figure 1: Typical Retinal Images of diseases

1.1 Motivation and Background

The high occurrence and significant impact of these eye diseases emphasize the need for efficient diagnostic tools. In general, the traditional diagnosis techniques are based on the interpretation and analysis of the eye by medical experts which is usually very time-consuming, and there are chances of error. Machine learning has improved the process of diagnosing eye diseases early attempt used Support Vector Machine(SVM), Qiao et al. (2017) used a SVM approach to predict cataract disease along with the genetic algorithm(GA) and was able to classify it efficiently. Later on, Behera and Chakravarty (2020) targeted DR also known as diabetic retinography using SVM to efficiently predict the disease by getting a good value of accuracy.

The implementation of deep learning especially on medical images provides a good solution to such challenges due to their automated, accurate, and efficient diagnostic capability on the images. Due to its fitness to automatically learn and obtain features from raw image data, the Convolutional Neural Network (CNN) has transformed the field of image analysis. The success of CNNs is widespread in applications like ocular disease prediction, from the fundus images or optical tomography scans . Previous research showed that CNN has the capability of diagnosing ophthalmic diseases. A full-scale survey of deep learning techniques for diabetic retinopathy classification by Atwany et al. (2022) demands special attention due to its ability to assess the health state in retinal fundus images and even help preserve a patient's vision. Deep learning-based automated diabetic retinopathy screening has been proven cost-effective by saving manpower resources and facilitating labor costs in low -to middle-income countries. For example, research of Shankar et al. (2020) proposed an advanced automatic approach namely Hyperparameter Modification Inception-v4 (HPTI-v4) to discover and categorize diabetic retinopathy disease from fundus pictures that outperformed other techniques. Another research by Qummar et al. (2019) suggested a composite deep model of five deep convolutional neural network models to enhance the categorization capacities for different stages of diabetic retinopathy and exceed existing approaches based on the Kaggle dataset. These examples emphasize the importance of such improvement in designing and enhancing CNN for ocular disease prediction. The challenge is to improve precision and speed and to effectively diagnose multiple eye diseases using CNN. This research attempts to contribute to the multiclass classification of eye disease based on the real-world dataset of fundus images , Ocular Disease Recognition available on kaggle.

1.2 Research Question

What are the most effective data preprocessing and augmentation techniques for enhancing the performance of the CNN model in the area of Ocular disease classfication?

1.3 Research Objective

1. To perform the critical analysis of the existing methodologies their benefits and their drawbacks in accurately predicting single and multiple diseases in eyes
2. Building and implementing the CNN model on the dataset
3. Implementing and employing the image processing and data augmentation techniques on the dataset
4. To evaluate the results of different data preprocessing and augmentation approaches on the performance of the CNN model.
5. Comparison with the other research in similar domains.

1.4 Outline of the Remainder of the paper

The next section of the paper is the literature review that focuses on previous work done in the domain of ocular disease diagnosis. Section 3 explains the methodology. Section 4 explains the design specification. The fifth part of the paper explains the implementation and evaluation process. The last section covers the conclusion along with future work.

2 Literature Review

In this section prior and most recent research conducted in the area of ocular disease diagnosis is highlighted. This section is divided into three parts in the first part the research work that is related to machine learning and its approaches is presented. In the second part, the research that focused on CNN and Deep Neural Network (DNN) is presented. Furthermore, in the third part, the research work related to transfer learning and pre-trained models is mentioned, and the last part describes the research gaps in the previous research related to disease diagnosis.

2.1 Papers Related to Machine Learning for Eye Disease Detection

SVM is considered a strong ML model that is commonly used for the task of classification. Qiao et al. (2017) proposes a framework for the diagnosis of cataracts. For diagnosis, SVM enhanced with the genetic algorithm for the optimization of feature weights is used. The results indicate achieving a high accuracy of 95.33% for the binary classification task and for the four classes (normal, mild, moderate, and worst) an accuracy of 87.52% is accomplished. The segmentation of fundus images into 17 parts and extracting broad features for example color, textures, and wavelet coefficients increase the capabilities of the model for better accuracy. The strength of this research lies in achieving better classification accuracy and using the genetic algorithm to optimize the feature weights. However, computational complexity and the time required for the genetic algorithm are the main weaknesses of this research study. Future work is focused on decreasing processing demands.

In Carrera et al. (2017) SVM is also used in used for the early detection of diabetic retina. This research used the Messidor database. Detailed preprocessing of retinal images is done to separate the blood vessels, microaneurysms, and the hard exudates, moreover, standard deviations of RGB components, blood vessel densities, and microaneurysm counts are extracted. These features were used to train the SVM model for detecting the presence of diabetic retinopathy and classifying the NPDR which is called non-proliferative diabetic retinography. An accuracy of 85.1 is achieved and sensitivity is 95% in the classification task. Key strengths are the high accuracy of the model's broad feature extraction and robustness against the parameter changes. The weakness of the research study lies in computational complexity.

In Pandey et al. (2020) three methods are presented for the detection of glaucoma which is an eye disease by using the Bin Rushed data set which is publicly available. The methodology contains the detailed preprocessing of the retinal images to separate and extract the Cup to Disk Ratio(CDR) and Radius to Disk Ratio(RDR). These features are used to train various ML models such as SVM, KNN, decision tree, and CNN. Moreover, the VCG-16CNN model was implemented to get the best accuracy of 99.6%. While the best accuracy of SVM, KNN, and the decision tree was 98%. The deep learning model employs the 17 layers including the convolutional and the max-pooling layer to accurately classify glaucoma. The decision tree accuracy was 87%. The research has potential for clinical application as the strength of this research lies in achieving the high accuracy for glaucoma the room for improvement is to further validate it across the diverse dataset.

Behera and Chakravarty (2020) the study addresses the Diabetic Retinopathy (DR) which is one of the major reasons for blindness, using SVM. The methodology used in this research is to preprocess retinal images to obtain the green channel, apply the equalization of the histogram, and enhance the edges. Features are then extracted by using algorithms such as Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF) that capture the regions that are indicative of DR. These features are then compiled into a feature matrix that is labeled as normal or of DR and trained on the SVM classifier. The model attained a sensitivity of 94% and an accuracy of 98% on the test data. The strengths of this research are the accuracy and sensitivity along with the effective feature extraction and preprocessing. However, the SIFT and SURF increased computational complexity.

Another research study Purandare and Noronha (2016) focused on the DR also known as diabetic retinography by utilizing the fundus images. Image processing techniques along with the SVM are used for the early detection of disease. The methodology of the paper involves the preprocessing of the retinal images and segmenting the clinical features like the region of the blood flow, exudates regions, and bifurcation points. Textual features like correlation and homogeneity entropy are extracted. The preprocessing steps involved the resizing of images and employing the adaptive histogram equalization to amend the technical differences of the images. In the segmentation process active contour model is used to remove the optic disc to avoid the problem of false detection and blood vessels are enhanced by using a 2-D Gabor wavelet. Bifurcation points are identified by analyzing the segmented blood vessels. These features are compiled into an SVM classifier which attained an accuracy of 92.55% and a sensitivity of 78%. The aspect of this research is the accuracy and suitability of feature extraction, but moderate sensitivity is still challenging.

In Harshini et al. (2023) multiple diseases related to eyes are detected by the machine learning approach. A combination of pre-processing techniques and ensemble learning can get a high value of accuracy. The methodology involves the preprocessing of the retinal images, by extracting the features through the pixel matrix conversion and ML algos such as SVM, RF, and deep learning models like MobileNet and InceptionV3. A modified voting classifier that is combined with XGBoost, decision tree, and the Gradient Boosting classifier achieved an accuracy of 98%. The image resizing of the image is 227 * 227 pixels. Seven different types of ocular diseases are targeted in this research. The strength of this research is the high accuracy and effective ensemble learning that give significant advancement, but this also increased the computational complexity of the study.

In Xu et al. (2018) a framework is suggested for the detection of eye disease named MKLclm. This framework integrated knowledge of various informatics domains and used learned SVM for the integration. The methodology involves the preprocessing of data from various domains of informatics. The dataset that is used is named SiMES which includes 2258 subjects that have comprehensive data from a different domain. Each base kernel is developed from the augmented features and MLclm optimizes the combination of kernels for the detection of disease. The preprocessing techniques include histogram-based analysis and SIFT. The study got significant performance improvements and high classification accuracies for three diseases named glaucoma, AMD, and PM. The strengths of the study are high performance and effective data fusion. While negative aspects are high computational complexity and the need for validation on different datasets.

In Ganesh et al. (2024) multiple diseases related to eyes are targeted using different datasets. Firstly, data augmentation is performed to balance the dataset for this purpose rotation, shear, and zoom transformations of images are performed which increase the dataset size and ensure better model performance. For efficient predictions of the diseases VGG-19 model is utilized that is followed by the normalization to ensure a consistent pixel value range. The model is then trained with convolutional and dense block layers. Convolution layers are used for the extraction of the features while the dense blocks are used to enhance the learning capacity of the model and to mitigate the gradient problem. the training is performed separately for both eyes. The model performance is calculated using accuracy and loss rates. The VGG-19 model achieves higher accuracy rates better than other pre-trained models. For the Challenge-AMD dataset, 96% accuracy is achieved and for the Challenge-PM dataset 94% accuracy is achieved, and 98% for the Challenge-GON dataset. The strength of this study is high accuracy and efficient data augmentation but due to dense architecture, the computational complexity is a challenge.

2.2 Papers Related to CNN/DNN for Eye Disease Detection

In Khan et al. (2020) multiple diseases in the eyes are detected by using deep CNN. The methodology of the work involves preprocessing images with adaptive histogram balancing and morphological tasks are performed to enhance the condition of pixels in the images. Pretrained models after fine-tuning are used for these tasks they evolve transfer learning. Models that are used are ResNet50, InceptionResNetV2, and the EfficientNet. In this research output from the multiple models is combined to increase the overall performance. The dataset of ODIR is used that contains the images of 5000 patients and covers multiple eye diseases. The enhancement of images showed improved performance with an overall accuracy of 86 percent and sensitivity of 81%. The strength of this study lies in the high accuracy and the large dataset. The challenge involved is the high computational complexity because of the usage of combined models to improve performance.

In Shamia et al. (2022) deep neural network is used to predict three common diseases of the eyes. The methodology involves the preprocessing of images with first grayscale conversion, then resizing of the images, CLASHE, correction of lighting, and Gaussian filtering. The Deep CNN architecture includes convolutional, pooling, and fully connected layers that consist of one hidden layer and 16 input neurons. The publicly available dataset is used for training, the validation of model, and testing. An online platform is developed for the rapid and easy diagnosis of the disease. The model achieved an accuracy of 91% for DR, 90% for the disease known as cataract, and 86% accuracy for the disease known as glaucoma. The strength of this research is developing an online platform for easy diagnosis and the accuracy values for these diseases that makes it capable of use in the health care environment while broader validation should be done if it is implemented in the clinical environment.

In Shaik et al. (2023) the multiple classification of eye disease is done by using deep CNN. The dataset used for the research is a publicly available dataset that contains 5530 fundus images of a total of 8 classes. The preprocessing step involves data augmentation for up-sampling to deal with the issue of class imbalance, Tensorflow is used for the the dataset preparation. The study used DenseNet , InceptionResNet, EfficientNet with transfer learning applied to VGG-19. In the training process of images, Goolge Colab is used with ReLU activation and binary cross-entropy loss functions. The best accuracy model achieved is for the class cataracts which is 91% but the model struggles with the other classes' detection the overall accuracy for all the other classes is 78%. The strength of the study lies in the thorough preprocessing and the effective use of transfer learning. However, there is room for improvement in the multi-disease diagnosis.

In Helen and Gokila (2023) CNN is used for multiple diseases of eye detection on fundus images. The methodology involves the preprocessing of images and resizing them into 100x100 pixels while contrast image quality through an adaptive histogram. The open-sourced data set that is used consists of 383 images of each category of a total of five diseases. For the training purpose, 269 images are used and for the testing purpose, 114 images are tested on the model. The architecture of CNN includes 64 layers and 128 filters with Relu activation, followed by max pooling and dense layers. Adam is used for optimization with a Learning Rate of 0.001, achieving an overall accuracy of 92.3%. out of all the five classes of accuracy, bulging eyes give the best accuracy. The effective preprocessing and high accuracy are the positive points of this research. However, the quantity of images is still challenging because of the small dataset. The accuracy values for the five classes should be tested on the large dataset.

In V et al. (2023) CNN is utilized to diagnose Glaucoma disease from the retinal images. The preprocessing of images is done through grayscale conversion and equalization of the histogram to improve the contrast and enhance visibility. The morphological tasks to remove the noise and to highlight the critical features are performed. The tasks involve isolating the critical region and accurate extraction of features to improve the reliability of the classification of images. The study used ResNet-50, VGGNET-16, and a combination of multiple architectures. CNN architecture begins with several convolutional layers for the detection of edges and textures, and max pooling layers to reduce dimensionality and highlight important features. Further convolutional layers are added for complex patterns. Drop-out layers are added to prevent overfitting by randomly disabling neurons during the training. The model is optimized by Adam optimizer with 0.001 Learning Rate. Additionally, data augmentation approaches, such as rotation, perform flipping and scaling to increase the divergence of training data. On the PSGIMSR dataset 91.13% accuracy is achieved and 98% on the HRF dataset. The strength of this work is high accuracy values, but large datasets should be used for further validation of results.

In Li et al. (2020) the correlation between the left and right-eye fundus images is identified. The study proposed a DCNet that combines the CNN for feature extraction, SCM for getting the pixel-wise correlation, and a classifier for the generation of disease probability scores. By the use of SCM the DCNet model outperformed the base models and got a good classification, particularly when using deeper ResNet architectures. Specifically, the ResNet backbone with SCM has achieved the highest accuracy of 93%, accuracy values get low when SCM is not used with the model. The dataset consists of 5000 patients that are divided into a total of 8 classes of diseases because of the large amount of data robust training ground for the validation is achieved. The results specify the good performance gain even though the model has computational complexity which is one of the drawbacks of this research.

In Prasad et al. (2019) a deep neural network is proposed to classify two diseases DR(Diabetic Retinopathy) and Glaucoma. These two conditions are considered as the leading causes of blindness so that's why the author focused on these diseases. The proposed model is trained on datasets from Kaggle and Medimrg that consist of labeled images for both of these diseases. The model consists of a layers neural network with convolutional, pooling, and dense layers to classify the features from the input images. In the phase of training the images are preprocessed by rescaling and resizing which is followed by a series of convolutional and pooling operations to extract the relevant features. The features that are extracted are processed through the dense layers for the final classification into healthy or disease categories. For the identification of DR and glaucoma, an accuracy of 80% is achieved. Moreover, the integration of the model is done into a human interactive GUI that allows the real-time analysis of disease and this is one of the positive aspects of this research. However, with more preprocessing for 2 classes accuracy could be improved.

In Vayadande et al. (2022) the eye disease classification by using the deep learning approach is performed. Three models are used for the research purpose the first one is CNN and the others are V3 and VGG-19. The disease that the author targeted in this research is the cataract. The dataset used for this purpose contains images of five thousand patients with normal and cataract images. VGG-19 got the best results in this research with an overall accuracy of 95.87% .V3 performed better than the CNN model, and CNN performance is identified as the lowest. Further research directions aim at to use large datasets to increase performance. The positive aspect of this research is the high accuracy of the algorithm for the classification of cataract disease. However pre-trained models require a high-performance machine for execution.

In Islam et al. (2019) the Neural Network is used for the classification of diseases related to eye disorders. CNN is used to detect the multiple diseases that are present in the eye. The study focused on the need to detect eye diseases in their early stages such as glaucoma, cataracts, and myopia disease to prevent blindness. The study used the ocular disease prediction dataset that consists of images that are captured through various commercial cameras so which also increased the complexity of images. Preprocessing steps involved the use of CLAHE and data augmentation along with label corrections of the images. The model of CNN attained an accuracy of 80% an F1 score of 85% with a kappa score of 31%. The good point of this research is the high accuracy value that has been achieved. However, the negative aspect of this research is the low kappa score which indicates limitations in agreement between the actual classification and the predicted class. Overall the study contributes significantly to the field of ophthalmic disease diagnosis.

In Li et al. (2021) the challenge of multiple eye disease detection is done with the help of fundus images and by utilizing the OIA-ODIR dataset. The author introduced the dataset from the data collection stage to the challenges involved in the dataset. The dataset consists of 10000 binocular fundus images of a total of 5000 patients. This dataset mirrors a real clinical scenario where both eyes could detect multiple eye diseases. The author benchmarked several state-of-the-art deep learning algorithms that include ResNet, VGG, and Inception. Increasing the size of the network does not improve the multiple classification task. Advanced feature fusion techniques significantly enhance the accuracy of the model. VGG-16 and the Inception-v4 performed best when the model was evaluated on the testing data. This research gives a solid foundation for future research in the area of multi-class ocular disease detection.

2.3 Papers Related to Transfer Learning for Eye Disease Detection

In Prasher et al. (2023) a transfer learning method is developed for the disease classification. Models that are utilized for this purpose are MobileNetV3 and the other one in EfficientNetB0. The classes that are targeted for this research are Normal, Diabetic, Glaucoma, and Cataract class. The author utilized an open-sourced datset that is available on kaggle (Eye Disease Classification) that contains over 4000 retinal images of these four classes. The study aims to enhance the performance of these four class predictions. The EfficientNetB0 model attained the highest accuracy of 94% and the MobileNetV3 reached an overall accuracy of 73%. The positive aspect of this research is the good performance of EfficientBetB0. However negative aspects include the low performance of MobileNetv3 which could be further improved by adding more pre-processing techniques and better data augmentation.

In Aranha et al. (2023) a transfer learning method for the diagnosis of ocular disease is performed on the less in quality images. The author used CNN and the transfer learning techniques to identify four disorders of the eyes Diabetic, cataract, excavation, and blood vessel disease conditions. The number of images that are utilized for this purpose is 38,727 high in quality images and 13000 lower in quality images. The model achieved 87 percent accuracy for cataracts, 90% for the DR, 87% for the excavation, and 79% for the blood vessels. The positive aspect is the validation of transfer learning on the low-quality images that could be accessible where constrained on the setting. However, blood vessel accuracy needs further improvements. The study contributes significantly to the automated diagnostic of eye disease. Additionally, the model is adaptive for low-quality images so this is another good point of this research.

In Gour and Khanna (2021) a CNN-based transfer learning approach is proposed that addresses the multi-disease classification on the eye retina image dataset. The study used fundus images to detect 8 different kinds of diseases. In this research two models are proposed, one using two input CNNs that deal with both the right and left fundus images separately, and the other one is concatenated CNN architecture that combines images before classification. The image size used for the research is 256x256 and the data augmentation is performed for balancing of data among different classes. The positive aspect of this research is the high performance of VGG-16 with SGD optimizer which achieved the best accuracy of 84%. However, the negative aspect of this research is the unbalanced dataset that does less performance for the less presented classes.

In Wang et al. (2020) EfficientNet is used for the various different eye disease classifications using the fundus images. The study used a combined model that integrates EfficientNet for feature extraction and a custom neural network for different class classifications. The dataset used for this purpose consists of 8 different diseases of the retina. The ensemble model combines the output probabilities of different models to achieve the final results. The resizing of images they have done in this research is 229x229 pixels the data augmentation techniques like random rotation and performing histogram equalization for the enhancement of image contrast. Results on the 40 images gave an accuracy of 89 percent. The positive aspect of this research is to handle the multi-label classification and achieve good results with the limited data.

In Singh et al. (2023) a deep learning technique is used to detect the cataract. The study uses an open-source dataset Ocular Disease Recognition avialble on kaggle. The preprocessing steps involved converting the image size into 224x224 pixels and applying data augmentation on the dataset. The VGG16 model was fine-tuned with the LR of 0.01 and evaluated at various epochs the highest accuracy achieved was on epoch 20. Overall, 96% accuracy was achieved for the disease cataract. The positive aspect of this research is the high accuracy value while the negative aspect is the dataset size used for training the model.

2.4 Breif Comparison of Previous Research

Dataset	**Binary/Multi-Class**	**Model**	**Accuracy**	**Reference**
Messidor	Binary Classification	SVM	85%	Carrera et al. (2017)
Bin Rushed	Binary	SVM, Decision tree, KNN, VGG 16	KNN = 98% VGG= 99%	Pandey et al. (2020)

Dataset	Binary/Multi-Class	Model	Accuracy	Reference
Private	Binary	SVM	92%	Behera and Chakravarty (2020)
SiMES	Multi Classification	Standard SVM Multi Kernel Learning	Glaucoma 93% AMD 83% Pathological Myopia 94%	Xu et al. (2018)
Challenge-AMD,Challenge-PM Challenge-GON Dataset	Multi-Class Classification	VGG-19	94% 94% 86%	Ganesh et al. (2024)
ODIR	Multi-Class Classification	ResNet 50 InceptionResNetV2	86%	Khan et al. (2020)
ODIR	Multi-Class	VGG-19	78%	Shamia et al. (2022)
PSGIMSR , HRF	Binary classification	ResNet-50 VGGNET-16	91% 78%	V et al. (2023)
ODIR	Multi-Class	DCNet	89%	Li et al. (2020)
Messidor	Binary	DNN	80%	Prasad et al. (2019)
Eye Disease Classification	Multi-Disease	EfficientNetB0	94%	Prasher et al. (2023)

Table 1: Summary of Classification Approaches

2.5 Research Gap

Even though a lot of research can predict ocular disease by machine learning, deep learning, and using transfer learning approaches the computation complexity and the use of large datasets remain challenging tasks. Most of the research work that is getting a high value of accuracy deals with a particular disease and a small and nondiverse dataset. Effective pre-processing is crucial, but it often affects the computational power, this research aims the use of CNN model that could be able to predict the disease deal with the challenge of computational power, and increase the accuracy value of the prediction while using fewer resources of the machine.

3 Methodology

The overall research followed the KDD methodology that describes the steps from the data collection to the final evaluation of this research. KDD has 6 steps to be followed all the steps and tasks performed to achieve those steps are mentioned in detail from the data collection stage to the final evaluation stage.

3.1 Data Collection

The first phase of the research was to collect data that contained a good number of images so that the model could effectively train. The image dataset is gathered from Kaggle name of the dataset is Ouclar Disease Recognition. The link of the dataset is Ocular Disease Recognition (ODIR-5K). It contains images of multiple eye disorders. That contains images of 8 different eye disorders consisting of a total of 6392 images. Images are not in equal distribution and all the images are of high-resolution. The total number of images along with the distribution according to the disease type is presented in Figure 2.

3.2 Exploratory Data Analysis

As the data set was not balanced the first task was to find the number of images of each disease. After the analysis of the count of each disease, it is found that the dataset is highly imbalanced it contains 2873 images of Normal retina, 1608 images of the retina that has diabetes, 708 images of other diseases or abnormalities, 293 images of cataract disease, 284 images of Glaucoma, 266 images of AMD disease, 232 are related to PM disease and 128 of hypertension. The dataset also contains verbal diagnosis which covers multiple diagnoses and severity of disease. Figure 2 illustrates the number of images graphically.

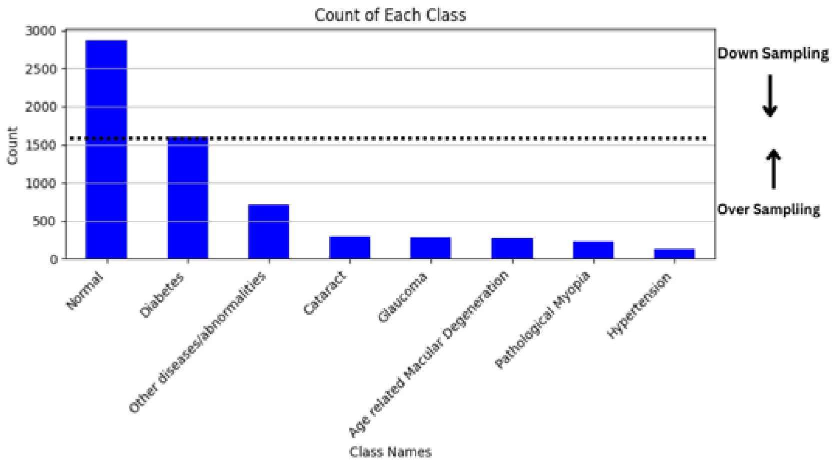

Figure 2: Imbalanced Dataset

All the images were of medium resolution of 512x512 pixels. The dataset contains other columns such as patient gender, the age of the patients, the diagnosis keywords, the label column, the next step that is performed in EDA is to get the information of the gender column how many male patients are present, and how many female patients are present in the dataset for this purpose a bar chart is created that is used to identify gender of patients in Figure 3.

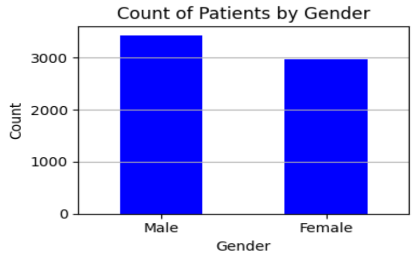

Figure 3: Gender Ratio (Moderately Balanced)

The next step is to figure out missing values, there is no missing value in the data frame. Images files are cross validated with the the dataframe. All the images that are mentioned in the data frame were in the actual directory of images

3.3 Data Prepossessing

Data preprocessing is an important task to put the data in shape that could help in obtaining better results. As the data was highly imbalanced, to obtain the data in a balanced form several preprocessing steps were performed. The class H has been eliminated because of low number of images. Firstly, classes of cataract, glaucoma disease, AMD, and pathological myopia disease are targeted for the preprocessing. A horizontal flip is given to these classes in such a way that each image of these classes is selected and then the horizontal flip to that image and then save it in a separate directory. After performing this step, a vertical flip is performed on the dataset. The same classes that are used for the horizontal flip are used for the vertical flip so that the number of images in the specified classes increases 4 times in the data set. Figure 4 presents the results of the vertical flip in which an image is presented before performing the flip and the same image after performing the vertical flip. In the next step to increase the number of images in the low representation disease classes rotation is performed. In the rotation process, other abnormalities class is added with other classes for the rotation. The value for rotation is 10 degrees, the strategy of rotation is to take the one image from the original images rotate it and save it in the directory where all previous augmentation is stored. Figure 4 presents data augmentation

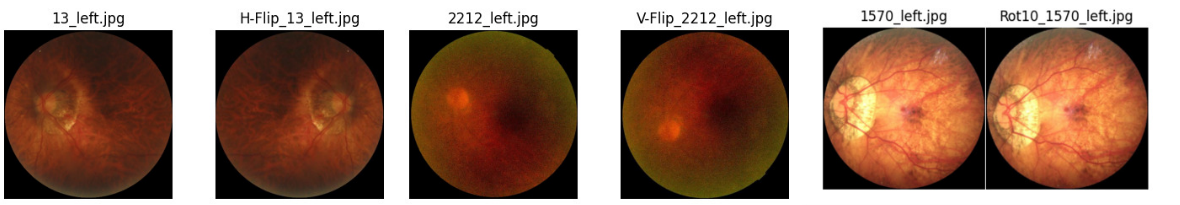

Figure 4: 3 Augmentation Results (Horizontal and Vertical Flips, Rotation)

By performing augmentation on classes, the total number of images in classes increased as shown Figure 5.

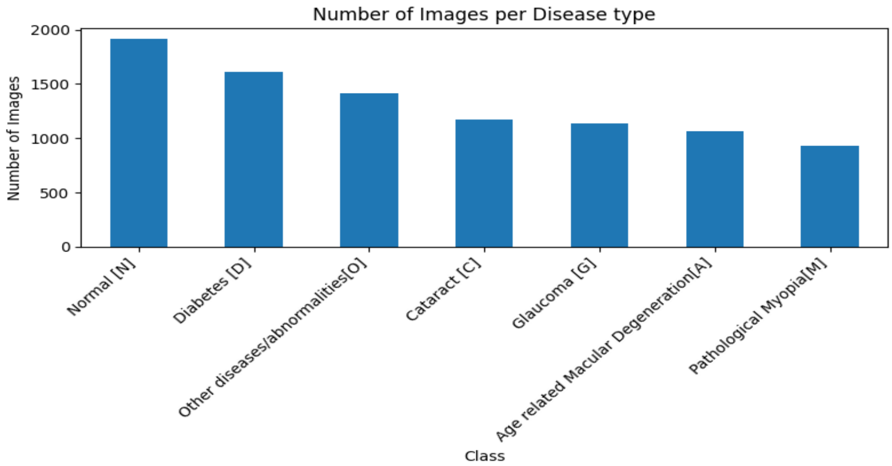

Figure 5: Image count after augmentation

The number of images has now increased and the rest of the imbalanced classes could be managed by downsampling the classes with high number of images For preprocessing in the label column as it was present in the string format it is converted into list format. Pre-processing is also performed on the keywords column for testing the model on the keyword column as well. As the disease conditions are mentioned in the diagnostic keywords column the specific conditions are checked in the keyword column, for instance, normal, cataract, etc. These conditions are checked both in the right and left diagnostics keywords if the disease is present, it will write 1 in the new column and 0 if the particular disease is not present then the images are fetched based on these columns' results.

3.4 Data Transformation

Data transformation in this research is to convert the image size to 150x150 pixels and compare the performance of the model on this image size. Then image size is increased to 224x224 pixels and the model performance is on it. The label column in the dataset was present in the string format so it is converted into the list format after it converting the label to one-hot encoded vectors by using a label encoder. Moreover, normalizing the data by dividing the images by 255 to get the image pixels in 0-1 ratio. The training and validation data is generated by using ImageDataGenerator.

3.5 Modeling

The next step is to a CNN model, the model is designed to perform feature extractions by using the convolutional layers and then classification tasks by using dense layers. In the convolutional layer, 32 filters are used with kernel size 3x3. In the second layer, 64 filters are used with kernel size 3x3 then 128 filters are used . Max pooling layers are used to reduce the longitudinal dimensions. Flatten layers to convert 2D features into 1D vector. Dense layers are used for classification with dropout 0.5 along with SoftMax activation function for multi-classification tasks. Figure 6 presents overview of the model

Layer (type)	Output Shape
conv2d (Conv2D)	(None, 222, 222, 32)
max_pooling2d (MaxPooling2D)	(None, 111, 111, 32)
conv2d_1 (Conv2D)	(None, 109, 109, 64)
max_pooling2d_1 (MaxPooling2D)	(None, 54, 54, 64)
conv2d_2 (Conv2D)	(None, 52, 52, 128)
max_pooling2d_2 (MaxPooling2D)	(None, 26, 26, 128)
flatten (Flatten)	(None, 86528)
dense (Dense)	(None, 128)
dropout (Dropout)	(None, 128)
dense_1 (Dense)	(None, 7)

Figure 6: Model Summary

3.6 Evaluation

The evaluation of the model is based on the accuracy of the model, specificity value, sensitivity, and loss. Accuracy is out of all the predictions made how many inputs the model is predicted correct. Accuracy is weighted across all diseases. Sensitivity is the rate of true positive values; it is defined as the ability of the model to correctly detect patients who are suffering from a particular disease. Specificity is the opposite of it to detect patients who don't have a disease. And the loss measures the difference between the actual and the values that are predicted. Sensitivity and specificity are reported per disease

4 Design Specification

The CNN model is supposed to classify fundus images into different categories by using a convolutional neural network for feature extraction and then dense layers for classification. Different strategies are adapted at different levels to improve the results of the models, such as testing the models using a label column and testing the same model using a keywords column. Figure 7 presents my design model. It consists of three parts.

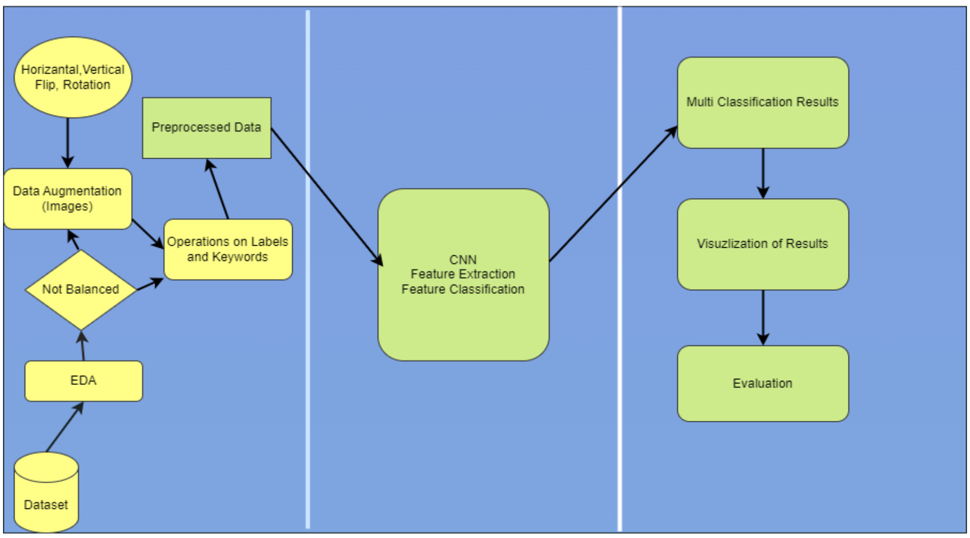

Figure 7: Research Design

The model of design consists of three parts, the first part describes the activities that are performed on the dataset. We will start by collecting the dataset and performing EDA on it. The EDA results show that data is not balanced, and 2 classes are dominating as compared to the other classes so the decision node is attached for this purpose if the data is not balanced it will move into the stage of performing data augmentation on the images. In the data augmentation stage, three different operations are performed on the images.Horizantal and vertical flip and image rotation of images

After performing this step the low classs images are in balanced form. While the class has more images like the Normal class. The down-sampling is performed to make the dataset balanced for the model. The next step is to perform the preprocessing on the column that we are going to predict. The label column and the keyword column are used for the prediction. On the label column, the string format is converted into list format, and on the keywords, disease conditions are checked in the diagnostic keyword column and stored the information of matching in another column, then images are extracted based on the disease type mentioned in the column.

5 Implementation and Evaluation

This section focuses on the implementation of CNN and the evaluation of the model by using different samples of ocular disease dataset classes for the classification of disease. Different strategies for increasing the dataset size are adapted to reach the final solution along with different columns for prediction. The model has convolutional layers supported by a max-pooling layer to find and down-sample local prominent features within images, which in turn explains characteristic elements like edges and textures. These layers are then followed by a flattening layer and dense layers, with ReLU activation to handle complex patterns. To prevent overfitting a dropout layer with a 50% rate has been integrated. The last layer uses the softmax activation function to give output in terms of class probabilities. It is compiled with the optimizer known as Adam and the loss function that is used is categorical cross-entropy 1 and batch size of 32 and model is trained over 30 epochs.

5.1 Environmental Setup

The environmental setup that is used for training the model is, it is trained over a system that is intel core m3, 7th generation CPU along with 8GB RAM. The platform that is used for this research is Anaconda Jupyter Notebook with a version of Python installed 3.11.4. The Keras version installed on the notebook is 3.4.1 model is built using Keras API using TensorFlow.

5.2 Implementation and Evaluation of 200 images of each Disease

The process is started with 200 images of each disease. H class has less number of images than 200 so all the images of H class are selected. An image size of 150x150 pixels is used in the first run.140 images are used for training and 40 for training/validation for Hypertension[H] 26 images are used, The CNN model architecture contains three convolutional layers with 3x3 kernels and ReLU activation, each of the convolutional layers is followed by a 2x2 max pooling layer. The dense layer is included in the network, with an output layer with 8 units (softmax). The optimizer that is used is Adam and the loss function that is used is categorical cross-entropy, the model is run for 30 epochs. This setting resulted in an accuracy of 40.5%. A confusion matrix, presented in Figure 8, in is created to evaluate the performance of each disease

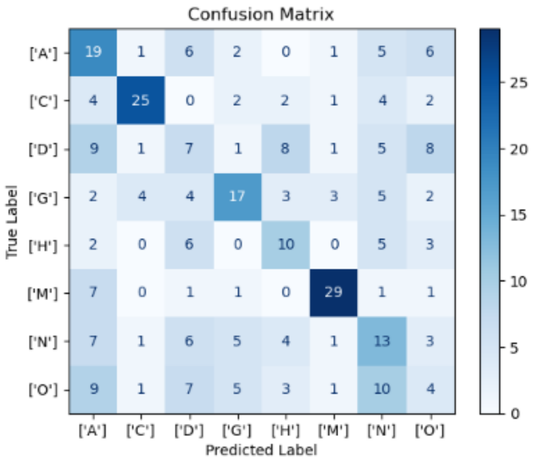

Figure 8: Confusion Matrix

5.3 Implementation and Evaluation without H class (200 images for each disease)

The dataset is then evaluated without the Hypertension class because it contains less number of images the total number of images in the whole dataset of this class is 128 images. Image size of 150x150 and labels column along with images are used corresponding to these labels. Images and labels are stored in the list and using images the mentioned size images are fetched from the directory. Images are then converted into NumPy array and then normalized by dividing it by 255. Labels are then converted into one-hot encoded format. The same CNN model is then run into the sample dataset of 200 images of each disease without hypertension class. The value of accuracy reached 41%. The confusion matrix in Figure 9 presents the performance of each disease class

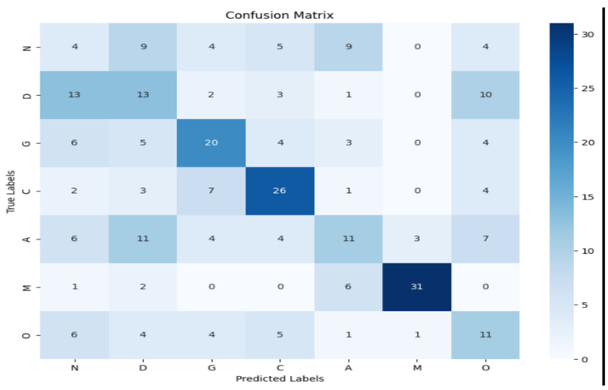

Figure 9: Confusion Matrix

5.4 Implementation and Evaluation without H and O class (200 images of each class)

The dataset is then evaluated on 200 images of each class but without considering H which is the hypertension class and O class which is the other disease class of the dataset. The O class is the collection of epiretinal membrane, drusen, and lens dust that is why it seems to create confusion during the learning process.The model is first trained on 30 epochs and then the epochs are increased to 100 to evaluate the performance of the model on the dataset. 45% accuracy is achieved on the testing dataset. Confusion matrix is shown in Figure 10

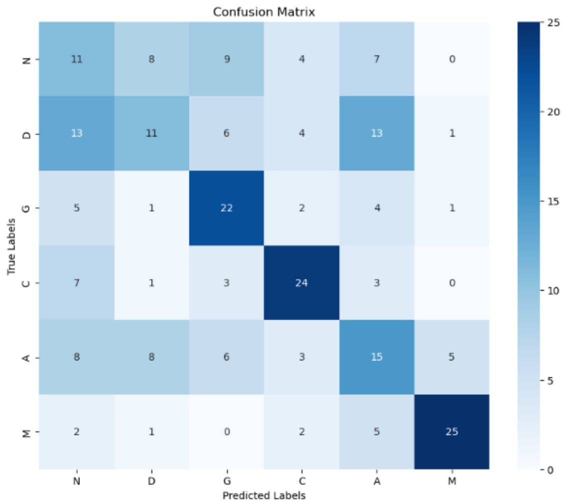

Figure 10: Confusion Matrix

5.5 Implementation and Evaluation after Vertical and Horizontal flip

The images of disease class, pathological myopia [M], Glaucoma disease [G], disease of age-related macular degeneration labeled as [A], and Cataract disease [C] are first flipped horizontally in such a way that all the images are taken give them a Horizontal flip so that pictures of these classes become double. The same step is repeated for the Vertical flip on these classes. Previously the number of images was M = 232, C=293, A=266, and G= 284. Their number increased to 696, 879,798, and 852 after these two flips, for Normal and Diabetes classes down sampling is done to reduce the number of images. For the Normal sample Fraction = 1/3 is used for taking a sample, which will take one-third of normal class images. Fraction =0.5 is used for the D class to take half of the sample of Diabetic disease class images. The total number of images loaded for this model is Normal class 958, Cataract class 879, Glaucoma disease class 852, Diabetes class 804, A disease class 798, and M disease class 696. Image size 224x224 is used along with the 30 epochs and batch size of 32. After splitting the images into training 60, validation, and test 20,20 image pixel values are normalized by dividing them by 255. Labels are converted from string representation to list and then labels are first encoded and then converted to one-hot encoded format. The model achieved an accuracy of 63% on test data with an average recall, precision, and f1 score of 63%. The graph of model accuracy and loss is represented in Figure 11 and the result of class classification by the use of a confusion matrix is in Figure 12.

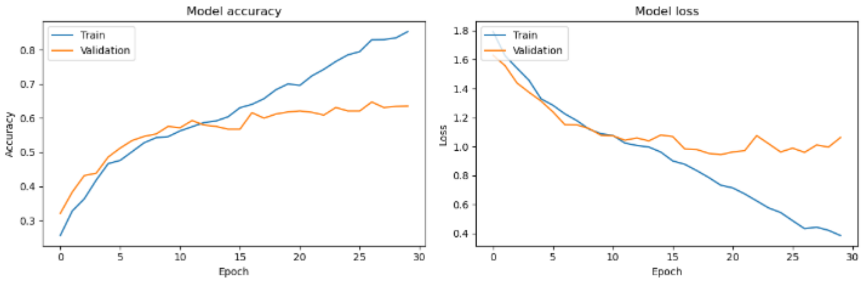

Figure 11: Model accuracy and Model loss

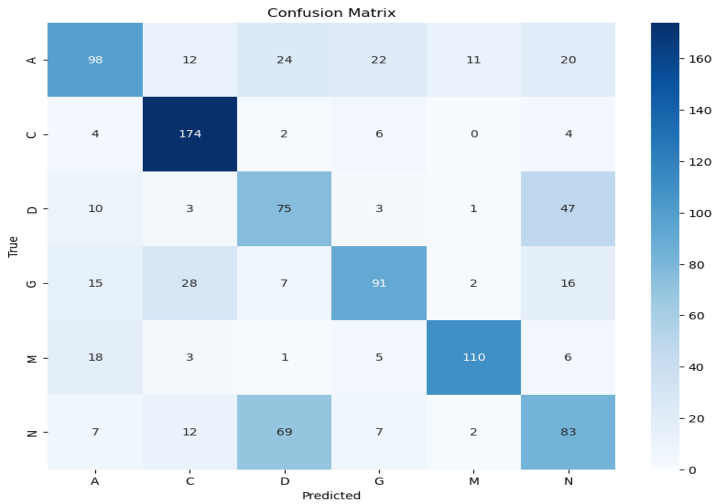

Figure 12: Confusion Matrix

5.6 The implementation and Evaluation after adding Rotation to images

The same number of classes is then evaluated by doing the rotation operation on the images on the other disease class [O], pathological myopia [M], Glaucoma disease [G], disease of age-related macular degeneration labeled as [A], and Cataract disease [C]. For this operation, original images of these classes are taken and a rotation of 10 degrees is given to them. The number of each class image after performing this rotation operation. Other disease=1416, Cataract disease =1172, Glaucoma =1136, Age-related MD =1064, Myopia=928 . The number of Normal class images selected for this model is Fraction=$\frac{2}{3}$ (two-thirds of images from the original dataset) and all images of the Diabetic class. The model is then trained with an image size of 150x150 pixels, for 30 epochs. The optimizer that is used is Adam and the loss function is categorical cross entropy. The accuracy achieved through this run is 59% with a weighted average of precision, recall, and f1 score of 0.60. The D class contains 30% of cases of mild non-proliferate retinography which is probably responsible for the misclassification as Normal.Figure 13 presents Confusion Matrix

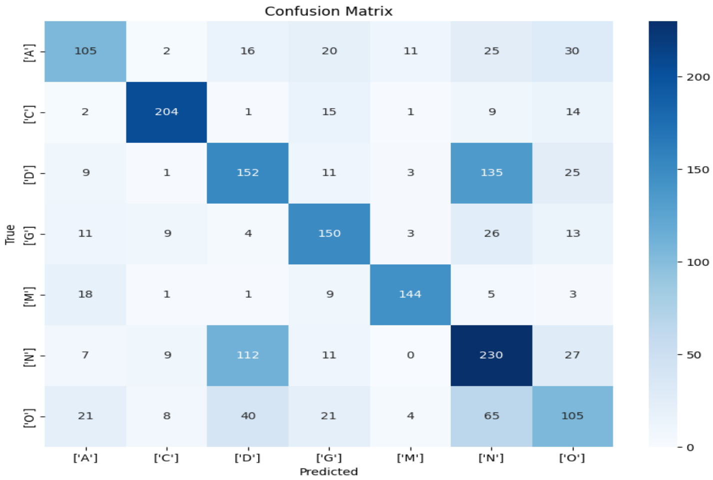

Figure 13: Confusion Matrix

5.7 Implementation and Evaluation by using Disease Diagnosis Keywords

So far, all the models are run by using the labels column, now the diagnosis keywords are considered for the execution of the model to classify disease. All seven classes are considered in the keyword column for processing. For this purpose all the disease are searched in the keyword column if the disease is present in a certain row of the keyword the information is stored as 0 or 1, 0 is used for the certain disease is not present in the keyword and 1 is used when the disease is present in the certain row . This process is repeated for all the diseases and then images are extracted based on the presence on this criteria. This process is used both for the left and right fundus images because the keywords column contain the diagnosis for the right and left fundus. This process is adapted to deal with the problem that may exist in the label column which is why keyword column is considered. The image are loaded and then resized into 224x224 pixels . The dataset is then split into training validation and testing sets then CNN model is defined using multiple convolutional by the max polling layers. Softmax activation function to output probabilities of seven classes . The optimizer that is used is Adam and the loss function that is used is categorical cross-entropy . Training is conducted on 20 epochs and the performance of the model is evaluated on the validation and the testing sets. The model achieved an accuracy of 81% with the weightage average of precision-recall and f1 score of 81%. The metrics indicate the model performed well in classifying the images into the diseases. The graphs of accuracy and loss concerning epochs are shown in Figure 14 . The confusion matrix of the classification of disease is shown in Figure 15

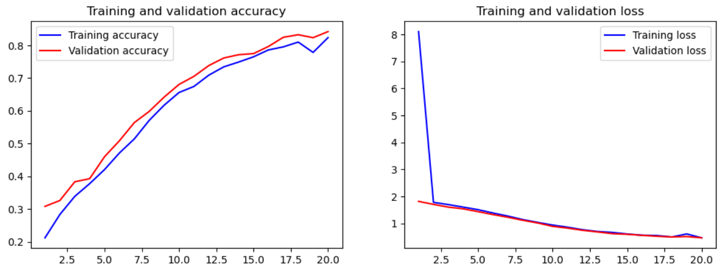

Figure 14: Training and validation accuracy , loss

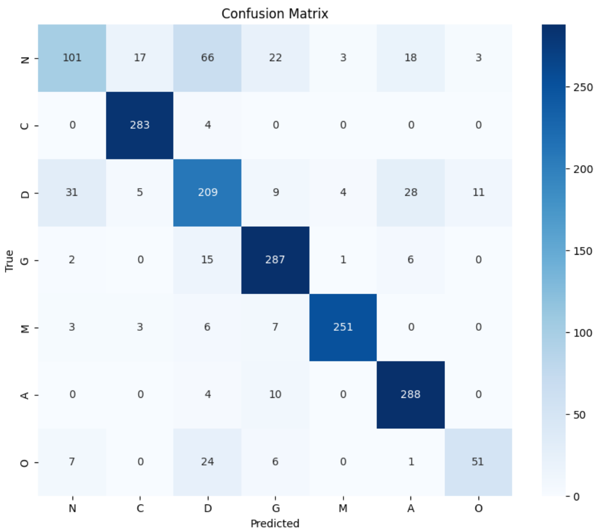

Figure 15: Confusion Matrix

6 Conclusion and Future Work

The focus of this research is to classify eye disease into multiple categories by the use of CNN. To answer the research question and to achieve the objectives of the research CNN model is applied to different variations of the dataset, and uses different columns for the predictions. The model is able to achieve the best accuracy of 81 percent when the keywords column is used for the prediction and 59 percent when the labels column is used for the prediction. These accuracies could be achieved if the data set increased by applying horizontal flip, vertical flip, and ten-degree rotation operation to the diseases that have fewer images and to the diseases that have more images like normal and diabetic disease down sampling is performed on them. In future work, it is planned to do processing on the other column and their impact in classifying the multiple diseases accurately. The dataset has a few columns that contain important information such as the patient gender and patient age column. There could be a possibility that there are few diseases that target a specific age group, like diabetic disease is more prominent in elderly people as compared to young people. Moreover, same like there could be a few diseases that target a specific gender, either male are female or they could be more prominent in the specific gender as compared to the other gender, the direction for the future is to take these two columns and define that how much they contribute to efficiently predict the disorder of the eye.

7 Acknowledgment

Many thanks to Dr. Christian Horn for the valuable supervision of this research project, helping me with developing my academic writing working through problems set, and helping with better formulation of this research project. His wealth of knowledge and wisdom has played a key role in the success of this work. I must also express my warm gratitude toward my family for their continual support and motivation during this project. It provided me with such motivation and resilience in their firm faith.

References

Aranha, G. D., Fernandes, R. A. and Morales, P. H. (2023). Deep transfer learning strategy to diagnose eye-related conditions and diseases: An approach based on low-quality fundus images, *IEEE Access* **11**: 37403–37411.

Atwany, M. Z., Sahyoun, A. H. and Yaqub, M. (2022). Deep learning techniques for diabetic retinopathy classification: A survey, *IEEE Access* **10**: 28642–28655.

Behera, M. K. and Chakravarty, S. (2020). Diabetic retinopathy image classification using support vector machine, *2020 international conference on computer science, engineering and applications (ICCSEA)*, IEEE, pp. 1–4.

Carrera, E. V., González, A. and Carrera, R. (2017). Automated detection of diabetic retinopathy using svm, *2017 IEEE XXIV international conference on electronics, electrical engineering and computing (INTERCON)*, IEEE, pp. 1–4.

Ganesh, K., Kishore, P. and Julian, A. (2024). Ocular disease recognition using machine learning, *Artificial Intelligence, Blockchain, Computing and Security Volume 2*, CRC Press, pp. 504–508.

Gour, N. and Khanna, P. (2021). Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network, *Biomedical signal processing and control* **66**: 102329.

Harshini, G., Saradhy, B. D., Varma, K. S. S. P. K. and Vadladi, V. K. (2023). Machine learning approach for various eye diseases using modified voting classifier model, *2023 International Conference on Inventive Computation Technologies (ICICT)*, IEEE, pp. 88–95.

Helen, D. and Gokila, S. (2023). Eyenet: An eye disease detection system using convolutional neural network, *2023 2nd International Conference on Edge Computing and Applications (ICECAA)*, IEEE, pp. 839–842.

Islam, M. T., Imran, S. A., Arefeen, A., Hasan, M. and Shahnaz, C. (2019). Source and camera independent ophthalmic disease recognition from fundus image using neural network, *2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON)*, IEEE, pp. 59–63.

Khan, I. A., Asaduzzaman, F. and Fattah, S. A. (2020). An automatic ocular disease detection scheme from enhanced fundus images based on ensembling deep cnn networks, *2020 11th International Conference on Electrical and Computer Engineering (ICECE)*, IEEE, pp. 491–494.

Li, C., Ye, J., He, J., Wang, S., Qiao, Y. and Gu, L. (2020). Dense correlation network for automated multi-label ocular disease detection with paired color fundus photographs, *2020 IEEE 17th international symposium on biomedical imaging (ISBI)*, IEEE, pp. 1–4.

Li, N., Li, T., Hu, C., Wang, K. and Kang, H. (2021). A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection, *Benchmarking, Measuring, and Optimizing: Third BenchCouncil International Symposium, Bench 2020, Virtual Event, November 15–16, 2020, Revised Selected Papers 3*, Springer, pp. 177–193.

Pandey, A., Patre, P. and Minj, J. (2020). Detection of glaucoma disease using image processing, soft computing and deep learning approaches, *2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)*, IEEE, pp. 1–7.

Prasad, K., Sajith, P., Neema, M., Madhu, L. and Priya, P. (2019). Multiple eye disease detection using deep neural network, *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)*, IEEE, pp. 2148–2153.

Prasher, S., Nelson, L. and Gomathi, S. (2023). Automated eye disease classification using mobilenetv3 and efficientnetb0 models using transfer learning, *2023 World Conference on Communication & Computing (WCONF)*, IEEE, pp. 1–5.

Purandare, M. and Noronha, K. (2016). Hybrid system for automatic classification of diabetic retinopathy using fundus images, *2016 online International conference on green engineering and technologies (IC-GET)*, IEEE, pp. 1–5.

Qiao, Z., Zhang, Q., Dong, Y. and Yang, J.-J. (2017). Application of svm based on genetic algorithm in classification of cataract fundus images, *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, IEEE, pp. 1–5.

Qummar, S., Khan, F. G., Shah, S., Khan, A., Shamshirband, S., Rehman, Z. U., Khan, I. A. and Jadoon, W. (2019). A deep learning ensemble approach for diabetic retinopathy detection, *Ieee Access* **7**: 150530–150539.

Shaik, K. M., Anupama, C., Paluru, S., Pedada, S. C. and Attuluri, B. K. (2023). Application of deep cnn networks in ocular disease detection, *2023 3rd International Conference on Smart Data Intelligence (ICSMDI)*, IEEE, pp. 372–377.

Shamia, D., Prince, S. and Bini, D. (2022). An online platform for early eye disease detection using deep convolutional neural networks, *2022 6th International Conference on Devices, Circuits and Systems (ICDCS)*, IEEE, pp. 388–392.

Shankar, K., Zhang, Y., Liu, Y., Wu, L. and Chen, C.-H. (2020). Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification, *IEEE access* **8**: 118164–118173.

Singh, G., Guleria, K. and Sharma, S. (2023). A pre-trained vgg16 model for cataract eye disease prediction, *2023 3rd International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)*, IEEE, pp. 1–6.

V, P., M, N., U, S. M., V, T. T., CH, M. and Y, D. (2023). Glaucoma detection using convolution neural networks, *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, pp. 368–372.

Vayadande, K., Ingale, V., Verma, V., Yeole, A., Zawar, S. and Jamadar, Z. (2022). Ocular disease recognition using deep learning, *2022 International Conference on Signal and Information Processing (IConSIP)*, IEEE, pp. 1–7.

Wang, J., Yang, L., Huo, Z., He, W. and Luo, J. (2020). Multi-label classification of fundus images with efficientnet, *IEEE access* **8**: 212499–212508.

Xu, Y., Duan, L., Fu, H., Zhang, Z., Zhao, W., You, T., Wong, T. Y. and Liu, J. (2018). Ocular disease detection from multiple informatics domains, *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, IEEE, pp. 43–47.