# Detection and Revelation of Multiple Ocular Diseases using Transfer Learning Techniques

MSc Research Project
Data Analytics

Uppiliappan Vijayaraghavan
Student ID: 20198442

School of Computing
National College of Ireland

Supervisor: Dr.Jorge Basilio

**National College of Ireland**
**Project Submission Sheet**
**School of Computing**

| | |
|---|---|
| **Student Name:** | Uppiliappan Vijayaraghavan |
| **Student ID:** | 20198442 |
| **Programme:** | Data Analytics |
| **Year:** | 2021-2022 |
| **Module:** | MSc Research Project |
| **Supervisor:** | Dr.Jorge Basilio |
| **Submission Due Date:** | 19/09/2022 |
| **Project Title:** | Detection and Revelation of Multiple Ocular Diseases using Transfer Learning Techniques |
| **Word Count:** | 9120 |
| **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:** | Uppiliappan Vijayaraghavan |
| **Date:** | 16th September 2022 |

## 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): | |

# Detection and Revelation of Multiple Ocular Diseases using Transfer Learning Techniques

Uppiliappan Vijayaraghavan
20198442

**Abstract**

Loss of vision is one of the most vigorous defect for surviving in the modern world and leading a peaceful life. The reason causing the defect in vision is a major problem in current generation since the use of unlimited electronic smart devices and gadgets which provides the light to create damage in the vision of the eyes. The detecting ability of examining the type of eye disease impacted, needs to be treated at the very initial stage so that the loss of vision can be prevented by providing proper treatment or surgery at timely manner with stages to improve the vision. This paper provides how the multi-classification of eye diseases such as Glaucoma, Diabetes, Pathological Myopia, Cataract, Age related Macular Degeneration, Hypertension and other eye related diseases are detected with the use of deep learning algorithms. The transfer learning algorithms of types : InceptionNetV3, VGG-16, VGG-19, MobileNet, RestNet, AlexNet are trained and executed in order to get the best results in detection. The InceptionNetV3 model provides the best results in terms of accuracy with 92.1% when compared to other models under consideration. The accuracies of other respective models are stated - VGG-16: 88.3%, VGG-19: 88%, RestNet50: 86.9%, MobileNet: 90.7%, AlexNet: 86.9%. Hence the automatic detection of ocular diseases and providing what kind of classification it belongs to, is accurately recognised by InceptionNetV3.

## 1 Introduction

There are 5 sense organs in Human body out of which eyes are considered as the most vital sense organ. The eyes provides vision to view or watch an object present in front to be visible. The loss in that sensing organ of vision causes a great damage to the person as the part of vision is missed for the individual as the visual perception is having issues which may affect the sensory vision of identifying object, colour. The analysis of visual perception works on the basis of electrical signals that is passed to the brain what is being visualized as an object or person with the view of normal analogy. The modern monkey statistics currently states that 77% of the current world population has any of the disability in viewing perceptions on the eyes which is due to the use modern electronic gadgets which we use on the routine day to day life from the range of young kids to older people. The vision loss can always be prevented if the treatment of eyes is done on the very early stage of having any defective issues or any issues in the possibility of viewing through eyes.

The use of computer vision has grown to a very great extent with the use of imaging techniques of Optical Coherence Tomography. The use of X-rays and ultrasound images is used for the analyzing the damage or detection in the particular region of an any image of organ analysis to be done. The eye diseases are of multiple types which classified as multi-labelled with the following classification such as Myopia, Diabetes, Hypertension, Cataract, Age related macular degeneration and other abnormalities depending on focus region around the retina region of the eyes. For the detection and treating this kind of diseases modern solutions can be provided only the basis of modern diagnosis methods which can be done only by the use deep learning methods in the field of ophthalmology. The main aim of this research is to analyze and provide the assistance in the diagnosis of eyes in the field of ophthalmology where the current lists of patients is very high due to the regular usage of modern electronic gadgets by everyone which raised to a exponential growth with the pandemic situation of pandemic of global virus in 2021.

The error in the vision which can be causing the problems in the eye is the refractive index which leads to near sightedness, far sightedness and also the distortion in distance vision of viewing. The important aspect of getting a new deep learning technology in detection of eye related diagnosis is said to be proved in this paper of what's new and to what extent it can help in the diagnosis with the transfer learning algorithms which is not completely justified in other papers to the maximum. As per the futuristic goal objective of how life can be in 2050, it is said that nearly 90-93% of people will have eye defect of any kinds which requires modern solution of having a loss. Hence the development of these kind of medical diagnosis is much more mandatory and it helps the ophthalmologists very easily in proceeding the stages of treatment to done and reduces the major surgeries while diagnosing at the earliest.

## 1.1 Motivation

The disturbance in the ability of visual perception is one of the common disease that 1 in 3 people in the current modern world has issues with. Hence the topic provides the information on how that can be solved with the use of modern technology which is used in the diagnosis of detection of any issue that leads to this problem. The use of electronic gadgets and its regular has increased a lot which involves the major time of using it. The light glazed with the use of those devices cause damage to the eyes and also affects the vision. The vision distortion is said to be treated with regular eye checks and test to be conducted for the vision to be clearly analyzed. The day of most human beings start with their phone in the hand as the first important task to check and the young kids are also more feasible to be attracted with the usage of these gadgets which in turn affects the eyes. The usage of laptops in the field of study and work plays a very important role which leads and affects the vision of the eyes. As the regular eye checks and tests are mandatory. The eye diseases are also related to age relational in most causes with the modern lifestyle necessities and habits in getting to know if a person is of above 40 years, the cataract and age related macular degeneration is common in this age. The hypertension is common in the age group of 25 and above which leads a great motivation to adverse the impact of diagnosis in the eye diseases. The main motivation and guide of the research is provided from the paper of (Guergueb and Akhloufi, 2021) where the detection of different eye diseases is done with the use of CNN methods using the EfficientNet and DenseNet algorithms. But the paper wasn't able to provide the complete information in the classification of multi-labelled diseases.

## 1.2 Value

The most important value of this paper is to provide proper assistance in the field of ophthalmology which provides more development in diagnosis and detection of eye diseases. The identification and detection of multi-labelled classification of eye diseases provides more information in finding the issues at the early stage which would be more complicated at the later stage. The use of multiple deep learning algorithms in order to identify the best results and finding the better accuracy and getting the good performance with the use transfer learning algorithms.

## 1.3 Research Question

***"How accurately the multi classification of ocular diseases can be detected using transfer learning algorithms?"***

## 1.4 Research Objective

This section provides the road-map of how this research is done in a section wise manner as listed below in the Figure 1.

| S.No | Objective | Tasks |
|---|---|---|
| 1 | Critical Review about research on previous papers and deep learning techniques to be used | Analysing the related work section and providing references of the work done and inspirations taken from |
| 2 | Data-preprocessing | Cropping the images. Resizing, Augmentation techniques |
| 3 | Data - Visualisation | Building the charts based on data and types of categorical diseases |
| 4 | Using transfer learning techniques building the best models using the algorithms | GoogleNet, VGG-16, VGG-19, RestNet, AlexNet, MobileNet |
| 5 | Building the GoogleNet model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 6 | Building the VGG-16 model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 7 | Building the VGG-19 model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 8 | Building the RestNet model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 9 | Building the MobileNet model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 10 | Building the AlexNet model and analysing the performance | Metrics : Accuracy, Precision, Recall, AUC value |
| 11 | Evaluation and finding the best model | Comparing all the results from the respective models |

Figure 1: Objectives of the Research

# 2 Literature Review

The development of machine learning and deep learning technologies in the modern era has been widely expanded in Image Classification and Object detection. The involvement of these methodologies in all the domains leads to modern development and helps in improving the trends accordingly. But in the field of medicine, it plays a vital role since, as it involves the life of living populations which affects the eco-system and provides modern treatment which is needed for the world population to get rid and accommodate to the usage of modern technologies. The eye vision loss is one of the mandatory disease to be treated by detecting what kind of ocular disease has caused the blurring or loss of vision. This section provides the critical overview of approaches and techniques that has been used in the previous researches which is related in finding the eye diseases.

## 2.1 Research on CNN models

The paper Wang et al. (2020) provides the research done on the eye fundus images with the ability of learning the features which is more powerful along with Computer Aided Diagnosis (CAD) which provides the doctors while screening or in performing the reference values with the ability of proper processing and providing the image analysis of the fundus images. The retinal fundus images are used in which the CNN (Convolutional Neural Networks) ensemble model is used to detect the fundus diseases which is of multiple type. The initial part involves the feature extraction using the EffiicentNet algorithm and the second is the classification of multi-label which is done by the neural network of custom classification. The images has been resized to 299x299 pixels with the images set to aspect ratio 1:1. The enhancement techniques of data rotation of 45,90 degrees and retaining the original images with main features. The use of ensemble strategy is the main highlight of this paper which provides how the overall model for classification can be improved. The colour complexion of the images are in grey and coloured states the information that there is a complexity in terms of designing done on this paper.

The research of Glaucoma detection is explained using the CNN on the paper of Chen et al. (2015) which is single type of disease detection. The model explains how the CNN architecture is designed with six learned layers along with 2 fully connected layers followed by 4 convolutional layers. The strategies of dropout and augmentation techniques are used in order to get the perfect results in terms of detection. The structural image cues are studied for glaucoma which provides the images related to optical disc. The deep CNN is used for the detecting the features of the images. The response-normalisation happens from the first to second layer. The overlapping and pooling layers functions are described which recap the feature statistics over the image region. 2 datasets are used ORIGO and SCES with the re-construction based methods and pixels of the image during augmentation are resized to 256x256. The result provides an AUC value of 0.83 and 0.87 in the two databases used.

Pratt et al. (2016) provides the information of research done on the diagnosis for the detecting Diabetic Retinopathy using the Convolutional Neural Networks. The severity is classified for the disease from the digital fundus images. The network is provided with batch normalisation in all the layers of the CNN. The pixels are resized to 512x512. The paper is addressed with the problem of overfitting which doesn't provide much better involvement in diagnosis even though weighted class weights are relative to the image amounts in each individual class. To reduce the initial time of training the

Gaussian initialisation was done to the network. The kernel for max pooling has the stride of the size 3x3 and 2x2. The classification is split into 5 categories based on the severity of the diseases and the final model achieved accuracy of 75%.

Wang and Wang (2019) states the research done on OCT datasets with respect to Diabetic Macular Edema (DME) and Age related Macular Degeneration (AMD). The re-usable features are proposed for the usage of effective features in the netork with respect to dataset size and adaptability enhancement with respect to different datasets. The adaptation with of different dataset to build-up the difference is approached on this paper. The CNN model of CliqueNet type of method is used. The CliqueNet is not a pre-trained model in comparison with other transfer learning techniques. The paper also provides information on FastNIMeans and BilateralFilter methods for the purpose of denoising which is done by the noise interfered. The pixel filling was performed carried out. The environmental setup is done based on the fine tuning of the network where the CNN has the 4 convolutional blocks and 12 convolutional kernels in each block. The features are extracted are in each block while execution of the experiment. The parameters in the CliqueNet are much optimised for the both the datasets used along with classification used with technique of SVM (Support Vector Machine). The optimiser of type SGD with the learning rate of 0.0001 and dropout of 0.5 value is given. The use RestNet and DenseNet allows the pre-training steps for the execution of the CliqueNet. The model provides the accuracy of 98% which is not the ability to get the research in a proper in handling 2 datasets and merging the difference between them.

Ram and Reyes-Aldasoro (2020) focus on identification of 3 eye related diseases such as Myopia, Age related macular degeneration and cataract. The dataset used in this research is ODIR-2019. The Convolutional Neural Network does the feature extraction part while the network feature mapping is changed. The paper aims to provide linking connection between the layers which are fully connected and number of classes. But the overlapping of datas going up is the main complexity which is not solved in this paper in which the neural networks struggles even if the number of layers if increased to n, the problem is incomplete without any solution. The oversampling technique is used for balancing the data. The size resolution of the images are taken as 1500x1500 while the images which can't be in those pixels are eliminated which shows the data used would be quite small with the given amount of dataset. The fibrometric filter is used for the data pre-processing which is used as its a type of hessian. The dataset involves multiple diseases while the research is done on Myopia, Cataract and Age related macular degeneration which would be a limitation to the research. The accuracy of the CNN model is 82%.

## 2.2 Research on Other Neural Network models

The glaucoma diagnosis and its detection is done using the Multi-branch Neural Networks (MB-NN) is proposed on the paper of Chai et al. (2018). The domain knowledge is easily analysed and learnt by the use of MB-NN models. 2 approaches are used in this research. Both are related to the segmentation techniques which is related to the cup size of the optic disc region and the setup is done between 0.3-0.5. The R-CNN model is used for detecting the region of optic disc which is more faster along with cup area, optic disc area and peripapillary atrophy (PPA) area. The CNN model checks only the retinal nerve fiber layer defects and the CNN of each block has 5 layers and SGD type of optimiser is used on this technique. The use of R-CNN provides the faster processing technique for the diagnosis of glaucoma as the result is produced from the MB-NN. The feature extraction of optic disc area is explained in a better way on this research. The image pixels are resized to 192x192 and the model provides an accuracy of 90%.

The paper Peng et al. (2019) provides the use of DeepSeeNet deep learning model classification of patients who suffer from Age related eye diseases (AREDS). The uniqueness of this paper when compared to other papers is that it has made a method of finding the severity score based on the AREDs type of detection from the disease detection while other papers has used only the raw image dataset. The bilateral classification of colour fundus photographs is done by the DeepSeeNet model. The pigmentary abnormalities and size of the drusen area in the region is learnt by the algorithm. The accuracy performed after the classification is of 67%. The use of assigning a severity score as per the stage of detection is more attractive feature provided on this paper. The use of this techniques in the detection of multi-classification of diseases could give good value to the research.

## 2.3 Research on Transfer learning models

The research of Guergueb and Akhloufi (2021) provides more information with respect to the actual research to be done on this specific topic and type of techniques. This paper provides information how the screening of fundus images helps in ophthalmology which is cost cutting, effective and efficient approach. The objection of only single disease detection was done on many papers, but this provides the information of having a multi-classification stage of detection but the paper is not justified in providing the legit information of the research. The ODIR-5k dataset is used for the research along with 6392 images as data with segregation of categories of ocular diseases. The data augmentation is done via the Mixup technique where the raw inputs is converted into random convex combination and encoding are done in one label. The methods of DenseNet and EffficentNet models with various versions are designed and executed. Global Average pooling layer (GAP) is used for avoiding the over-fit of the model and normalising the output is done by the soft-max layer for obtaining the final prediction. The learning rate annealing uses warm restarts using the Stochastic Gradient Descent. The images resizing are done at 512x512 pixels. The conclusion of this paper doesn't justify the classification of the multi-labelled diseases. The accuracy provides on results is of 87%.

He et al. (2021) states the research carried out on DenseNet Correlation Network (DCNet). The DCNet is the backbone of feature extraction for the CNN model which provides the spatial correlation in support of feature correlation and for the classification of score generation the classifier is used. The ODIR-5k dataset is used in this research. Spatial correlation function receives the 2 sets of features from the CNN model and provides the output of 2 interrelated features sets by receipting the correlation between them. During the feature extraction process since there is no fusion, the RestNet technique is designed to truncate the fully developed CNN model. The random cropping of image to 448x448 pixels is done from the earlier range of 512x512 pixels. Four RestNet models of type 18, 34, 50 and 101 are trained. The results are evaluated in comparison with the use of Spatial correlation and without the spatial correlation. The AUC score of 78.5% is achieved by using this technique.

The research paper of Gour and Khanna (2021) delivers the content of research done on the classification of ocular diseases using the deep learning algorithms. The main motive of this paper is providing the use of optimisers such as SGD and Adam rather than the actual detection. The paper of research doesn't involve any augmentation techniques as they have proceeded with the raw set of data. The Convolutional Neural network model us built along with few transfer learning models such as VGG-16, RestNet, MobileNet, InceptionV3. The paper was lacking the pattern of pre-processing techniques and concatenation of images is done. The images from the original dataset has been reduced in terms of count from the ODIR-2019 dataset. The model provides high accuracy which shows the clear under-fitting issues which is not addressed properly on this paper. Also the pixels size for the images is not mentioned in the paper.

The paper Khan et al. (2020) provides the research done on normal and diseased cases classification of the retinal fundus imaging using ensemble approach of deep learning architectures by the use of transfer learning approach. The adaptive histogram equalisation technique along with morphological operations by the use of enhancement technique which provides better classification of normal and diseased image separation. For getting the best result of classification, the predictions acquired from the transfer learning models such as RestNet50, InceptionNetV2, EfficientNetB2 and EfficientNetB0 are fused. The public datasets of DRIVE, ORIGA and RIM ONE are used to get the pressure from the Intra-ocular nerve. The enhancement algorithm is used for the acquisition procedure to counter the loss of the quality which can be inherent. The RestNet and InceptionNetV2 networks are combined to get the final decision which is made from the use of ensemble technique. The pattern is using the ensemble network is one of best highlights of this paper which yields an accuracy of 86% when combined the use network with respect to different approaches.

Research of Fu et al. (2018) provides the diagnosis of glaucoma using the segments from the main structure through screening methods. The visual features are hidden in most experiments but in this paper, the deep learning technique is practiced to provide additional relevant information of the images. The segments are done from the eye region with 4 deep streams on different modules and streams such as disc polar transformation stream, local disc region stream, global image stream and segmentation guided network. The research is done on SCES and new SINDI datasets. The 2 datasets involves whether there is presence of glaucoma or not and whether there is a loss of vision or not. The better polarised balance is handled by morphing the pathological regions and enhancement of the detection regions. The initial stream is done with the use of RestNet and has 5 down sampling blocks which is CNN based network. The utilisation of sigmoid layer is done via the use of classifier with layer 1x1. The input image has been resized to 224x224 pixels. The positive of this research is the setup encoding and decoding techniques which helps in getting the element wise rectified linear non-linearity (ReLU). The model provides the accuracy of 62% but the run time of program takes 5 hours for the algorithm to run.

Govindaiah et al. (2018) illustrates the screening of Age related macular degeneration (AMD) using the Deep Convolutional Neural Networks. The models of VGG-16 and RestNet approach of transfer learning methodologies is experimented. The initial step of classification is done in 2 classes based on the clinical significance. One with No or early AMD and the other with Intermediate or Advanced AMD. The AREDS dataset is used in this research. The 4 categories is split based on the size of infection and drusen extent in each eyes. The blurred and poor quality images are removed and cropping is done around the center region of the macula with a threshold value where the image is re-scaled. To create the gray-scale images, the green channel intensities are taken into consideration. The RestNet model is build for the pre-training purpose and the 16 layers architecture of type VGG-16 is used. The batch normalisation is added to enhance the generalisation at the last connected layers. The images are re-scaled to 224x224 pixels. The paper states that pre-processing of fundus images might remove the glare of the lens but the techniques are not implemented. The VGG-16 network provides the accuracy of 83% with respect to the categories of infection.

He et al. (2020) proposes the detection on early diagnosis of the ocular diseases using the transfer learning methodologies. Because of minor symptoms in the image it is hard to predict the type of diseases affected. Hence the use of AUB-Net is classify the data of the patients automatically depending on the categories of the diseases. The combination of Feature extraction module (FEM), classification module and feature fusion module (FFM) is the composition of the AUB-Net. From the patient photographs, two feature vectors extracts the bilateral fundus images which is done by FEM. For the preparation of representations of the features, the FFM proceeds with 2 levels of fusion and feature weighting and CFM handles the classification part. The features from the two eyes are treated equally using the concatenation and conventional summation. The RestNet models are built for the execution and training the data. For training the images are re-sized to 448x448 pixels and SGD optimiser is used along With the classification ability, the RestNet-101 provides the accuracy of 82.9%.

The automatic glaucoma diagnosis using the transfer learning and its comparative analysis is executed on the paper of Elakkiya and Saraniya (2019). The public source of datasets is used on this research from of Origia, Drishti-GS1. The CNN handles the classification tasks of the neural network. The feature extraction of the input images is analysed and given proper information of it can be done which provides the best reference to this paper of getting to know what to be done. The Recurrent neural networks (RNN) are designed to work for the sequence problems of predictions.The models of GoogleNet, RestNet, AlexNet, VGGNet involving the convolutional layers were designed. Different types of optimisers are used such as Adam, Adagram, RMSprop and Momentum in this research which helps in getting to utilise these optimisers in the multi-classifications to get the better accuracy in terms of results. The GoogleNet model outperformed with the best results on this paper 91%.

## 2.4 Research on Pre-processing models

The paper of Islam et al. (2019) provides the information of detection of ocular diseases of specific types such as Cataract, Glaucoma and Myopia using the CNN methodology. The tests for the detection of these diseases usually have manual tests such as acuity visual test, retinal examination and ocular tonometry. The pre-processing steps were performed with the use of contrast adjustment using the CLAHE (Contrast-limited adaptive histogram equalization) methods which is used for adjusting the contrast of the image by modifying the distribution of the intensity in the histogram. The visibility of the details in the local image is increases which is the main advantage of this technique in the field of medicine. The complexity of building the network is reduced by proper labelling techniques which enables the prediction of the model. After the labelling process, the single labelled and multi-labelled images are segregated and prone to dataa augmentation process. The network modelling is done by keeping the parameters as low as possible and it provides the result with respect value of 80.5%.

Shankar et al. (2020) provides the information of pre-processing techniques along with feature extraction done in the process of detection of diabetic retinopathy (DR) with the use of deep learning. CLAHE (Contrast-limited adaptive histogram equalization) methods is used along with the segmentation of the histogram model. The HPTI-v4 is run for getting the main features from the images which is segmented which provides a significant way of result. The Bayesian optimisation technique is introduced for tuning of the parameter along with MLP - Multi layer perceptron classifier is applied for the classification of DR images. The generation of confusion matrix at each stage from the source of original images, segmented images and classified images is displayed.

## 2.5 Research on Data Augmentation

The research paper Mikołajczyk and Grochowski (2018) provides the information of data augmentation techniques with respect to colour perspectives, histogram based methods, rotating, zooming and cropping. The image transfer style has been introduced, to get the new images of high quality which is perpetual and the created images are pre-trained using the deep learning methodologies with the help of neural networks. The image classification augmentation techniques are performed on medical datasets involving his-pathological images, skin melanomas diagnosis and breast magnetic imaging resonance. The popular methods of increasing the contrast, here it is done by 20%, histogram equalisation, White balance and sharpening of the images are described in the traditional transformation techniques. The Generative Adversial Networks (GAN) which is used for un-supervised new images generation using the strategy of min-max is shown. The texture transfer with respect to visual attributes of the raw image, lines, strokes, shading and a technique of image quilting where the small patches are stitched in the new images.

Araújo et al. (2020) illustrates the augmentation process done in the research of detecting Proliferative diabetic retinopathy (PDR) . The stage of extreme diabetic retinopathy (DR) and the final stage is known as Proliferative diabetic retinopathy. The factors involving in the original datasets didn't show the proliferative stage of DR. The heuristic technique of data augmentation using the structures of Neo-vessels synthesis is done. The augmented dataset is tested using the deep learning neural network algorithms. The image is studied by giving location parts to the Neo-vessels for getting the insights where the algorithm is handled by the Cartesian space which is from the parent tree.

The image enhancement and augmentation techniques is the highlight of the paper Karkuzhali et al. (2020). The glaucoma diagnosis with the use 2D fundus images using the super pixel screening of optical discs and optical cups is done on this paper. The Gabor filter is used for the feature space extraction where the clustering is done by the Convolutional Neural networks. The images are patched at the size of 256x256 pixels. The contrast improvement and morphological transformation are done along with the top-hat transform which takes all the information even from the small patch of information from the images. The feature extraction proves to be uniformly distributed but the paper has the drawback of having the malignant micro-calcifications on the images which appear to be small.

## 2.6 Summary of Literature Review

The review of the related works done with respect to the topics, dataset and techniques to be used has provided a clear description how this research can be formulated and prove to be different kind of research in terms of novelty. The research papers on ocular diseases proves that multiple classification of ocular detection can be done but the papers wasn't providing enough evidence how the actual research can be executed. This plotted the way of having transfer learning algorithms to be implemented for multiple classification of ocular diseases detection such as Pathological Myopia, Cataract, Glaucoma, Diabetes, Cataract, Age related macular degeneration and other eye related diseases. The use of algorithms such as InceptionNetV3, MobileNet, RestNet, VGG-16, VGG-19 and AlexNet can be done and was never employed in any research of having 6 transfer learning techniques applied on a single research with the use of optimisers and augmentation techniques. The research also provides clear information of support and it addresses the requirement of this new technology in the field of ophthalmology.

# 3 Methodology

The research is done on the methodology of Cross-Industry Standard Process for Data Mining (CRISP-DM) as it involves in determining the performance of deep learning algorithms and deals with different phases of project which is interlinked to each phase and creating a relationship between the tasks performed. The phases of CRISP-DM methods Paik and Ataka (2019) are explained below as follows from the Figure 2.

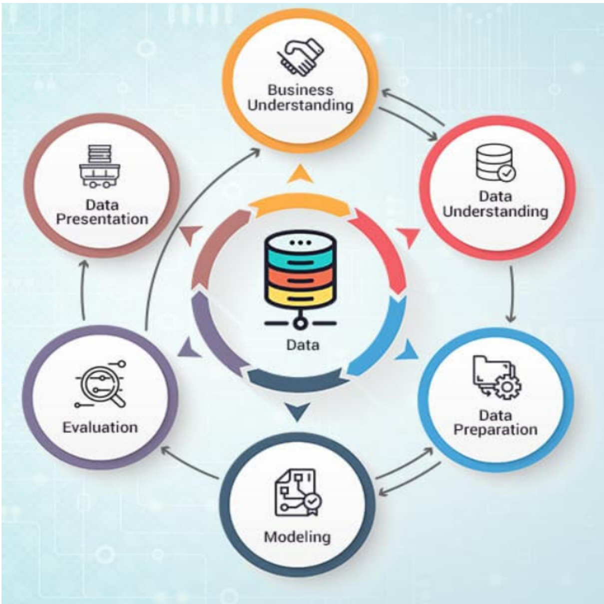


Figure 2: CRISP-DM Methodology

## 3.1 Business Understanding

The Business understanding defines the phase of grasping the technique of what to be done in the case of identification of the problem and describing it and ability to solve the issue by the use of techniques. The current scenarios is analysed for which the insights and goals to be executed is defined. In the medical industry, the radiologists and doctors via the scanned images of X-rays and photographs, the type of disease detection is done. But the use of deep learning technology, the human error can be completely taken off with the avoidance of manual mistakes. The 6 specific deep learning techniques has been applied with the evaluation of their performance is analysed by accuracy, recall, AUC value and precision. But accuracy is the main source of metrics along with the model fit to be considered. The main value of business is that, by detecting the type of ocular disease at the early stage will have a impact which would be completely positive for the patients as it can be replenished and avoiding the increase in severity, surgeries and loss of vision.

## 3.2 Data Understanding

The data understanding of how the data source is produced from or taken. The understanding how the data can be utilised and what all problems may arise with the data and formats of data while the implementation of the research starts with the dataset. The checking possibility getting the relevant answers to the questions that we have is possibly achievable. The dataset consists of 5000 patients history of data with 8000 images of dual eyes having any ailments of diseases in eye along with normal eyes. The dataset is available for public use and it is taken from the public source: kaggle. The category of the images in dataset is displayed on the Figure 3 below:

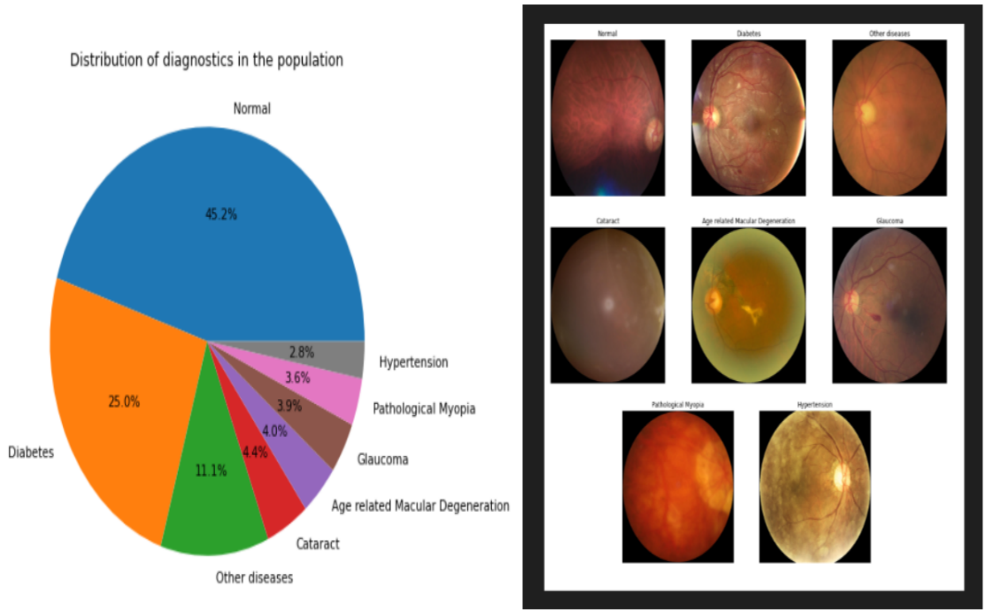


Figure 3: Distribution of Diagnostics and Image classification

## 3.3 Data Preparation

The data preparation involves the phases of selecting, constructing, cleaning, developing, integrate and making all the datas in same format. The data images are in the format of JPG. Each image was having different pixels but the process of having a standard resizing and steps involving normalisation and re-scaling has also been done for the data to be used for execution of the research. The steps of Data preparation is done using the Tensorflow package and Keras library in python 3.7.

### 3.3.1 Data Cleaning

The data cleaning is done to correct the messy and inadequate informative data to be aligned as per the actual data. The data cleaning is done based on the optical disk which is not visible in the fundus images along with the lens in the eye region which is having issues with. The hot encoding techniques is done for the representation of having categorical data which is very much needed for the prediction and detecting of the ocular diseases by the algorithms used. The hot encoding techniques is done based on the diseases representation in the dataset with the value of having represented with values 1 in the case of any diseases which represented below in Figure 4.

Create Hot Encoding for each disease using pandas series

```
[ ] #                              [N, D, G, C, A, H, M, O]
    dico_ds = {
        'normal'      : pd.Series([1, 0, 0, 0, 0, 0, 0, 0]),
        'retinopathy' : pd.Series([0, 1, 0, 0, 0, 0, 0, 0]),
        'glaucoma'    : pd.Series([0, 0, 1, 0, 0, 0, 0, 0]),
        'cataract'    : pd.Series([0, 0, 0, 1, 0, 0, 0, 0]),
        'age'         : pd.Series([0, 0, 0, 0, 1, 0, 0, 0]),
        'hypertensive': pd.Series([0, 0, 0, 0, 0, 1, 0, 0]),
        'myopia'       : pd.Series([0, 0, 0, 0, 0, 0, 1, 0])
    }
```

Figure 4: Hot encoding

### 3.3.2 Image Cropping

The images from the dataset is cropped as per the retina region alignment which yields the proper detection method. The image is cropped based on the cropped regions of retina to analyse the type of ocular diseases as everything is detected based on the retinal region, optic disc curves around the eye. The ImageCrop class function is used for cropping of the image along with the removal of black pixels in the image dataset. The cropped images are saved in the destination folder.

### 3.3.3 Image Resizing and scaling

The class function ImageResizer used as the function allows to align the image dataset with specific set of rules in resizing and mirroring of an image. The function is analysed by image quality, width, height and aspect ratio. The images are resized using the opencv function to 128x128 pixels because it allows the program to run in an efficient way and the allocation of RAM utilisation is much faster in terms of predicting the result. The scaling of images involves conversion of the images in the standard format based on the category of diseases which would be represented as 0s and 1s as stated in the Hot encoding techniques.

### 3.3.4 Data Augmentation

The function DataAugmentationStrategy class is being used for the process of data augmentation. The image is created and saved as original vector from the sample. The package cv2 is being used for reading the datas and generating new images. The techniques of having weight based models are build with different phases used such as contrast, saturation, gamma, hue, scaling based on the original images to create new images with respect to the regions of the retina where the actual input dataset is provided with complete information around the eye region. The Figure 5 shows the process of Augmentation.

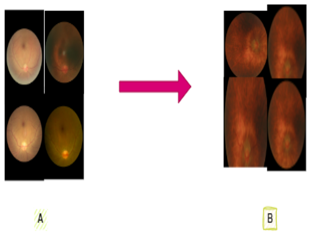


Figure 5: Augmentation Images A:Normal and B:Augmented

## 3.4 Data Labelling using Numpy

The multi-label classification problem is always addressed with use of Numpy. The images of multiple classification always needs to organised properly depending on the categories which creates a separate category for each individual class in the dataset. Since the datas of images by NumPy is stored in n-dimensional arrays, this will be easy for adding the augmented images along with original image dataset and merging as the input file for the research. The functions of NumpyDataGenerator is used for generating the numpy records for vector distribution in terms of arrays. The class GroundTruthFiles is used along with populate_vectors function with respect to the category of diseases. for saving the numpy labels, the absl package importing the app function and many other packages are imported which are time, cv2, csv, os and glob. The data files in the CSV format are converted to .py file format.

## 3.5 Modelling

This section provides information on selection of relevant models in deep learning methodologies for the classification of detecting the type of ocular diseases. The test scenarios is produced to get model's quality and validating it accordingly. The 6 models are described as follows:

### 3.5.1 VGGNet : VGG-16 and VGG-19

The VGGNet model is usually designed to run large scale of datas which involves images and is considered as the subset of Imagenet and considered as one of the best technique in the building blocks of CNN background. The memory of the algorithm is the main highlights which pre-learns the features of the generics in the images and runs it in a efficient way. The regions of retinas, colours and optic disc locations are read to a maximum extent in this type of method. The VGGNet used in this research is of two type : VGG-16 and VGG-19 where the difference is only in the layers of processing as shown in Figure 6. The VGG-16 holds 13 convolutional layers along with 1 softmax classifier followed by 2 layers which are fully connected. In the other hand the VGG-19, has 16 convolutional layers along with 1 softmax classifier and 2 fully connected layers. The pre-trained model when used provides good object detection and it helps in weight optimisation in which the representation of the output data would be more valuable.

### 3.5.2 InceptionV3 and AlexNet

The InceptionV3 algorithm shown in Figure 7 solves the computational efficiency of other transfer learning models. The economical cost is incurred and the number of parameters

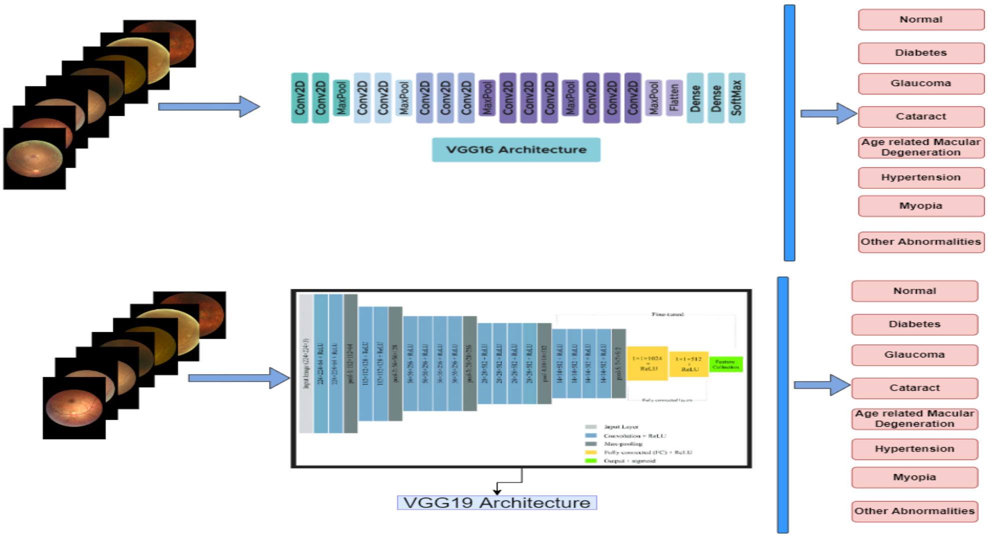


Figure 6: VGG-16 and VGG-19 Architectures

is reduced for good efficiency in InceptionV3. The method has much more techniques in optimising the network. The dimension reduction, convolutions and regulations are factorized. The Bigger convolutions are replaced by the smaller convolutions. The method also has the advantage of using asymmetric convolutions. The auxiliary classifier which is a small CNN based classifier is inserted between the layers. The network has dropout layers, average pooling layers, convolution concats and fully connected layers. The largest value is calculated from the max pooling layer within a feature map.

The AlexNet algorithms solves the classification problems of multiple categories of images very easily as it involves 5 fully linked layers and made of 3 convolution layers shown in Figure 7. The ImageNet runs as the subset of the model. The initial 2 convolutional layers for which overlapping max pooling layer is present in each of the convolutional layers. The ReLU as the activation function accelerates the algorithm to perform the best and training speed limit is nearly 6 times better than other transfer learning algorithms. The AlexNet involves multi-GPU unit.

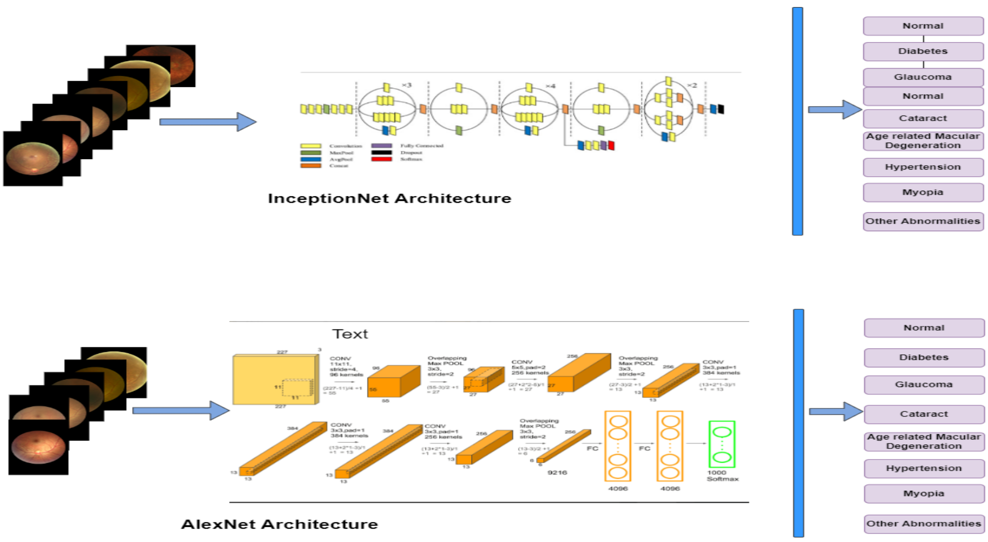


Figure 7: InceptionNetV3 and AlexNet Architectures

### 3.5.3 RestNet50 and MobileNet

The RestNet50 architecture shown in Figure 8, states that the convolution neural network has 50 layers which is a residual neural by stacking residual blocks which is numerous on top of each other forming a network. The problem of vanishing gradients is solved by the use of residual networks. The saturation of accuracy gets triggered by increasing the number of layers after saturation which is the only drawback of this method which is called as problem of vanishing gradient. The Batch normalisation is done at the last connected layers.

The MobileNet architecture shown in Figure 8 is a type of convolutional neural network which are specifically designed for the mobile applications. The number of parameters used on this method is comparatively less when compared to other transfer learning algorithms which is the positive and highlight of this model. With having the limited resources MobileNet provides good accuracy in a effective manner. Single convolution is executed in each channel by the convolutional filter of type depth-wise. It has 2 hyper-parameters which are Resolution Multiplier and Width Multiplier.

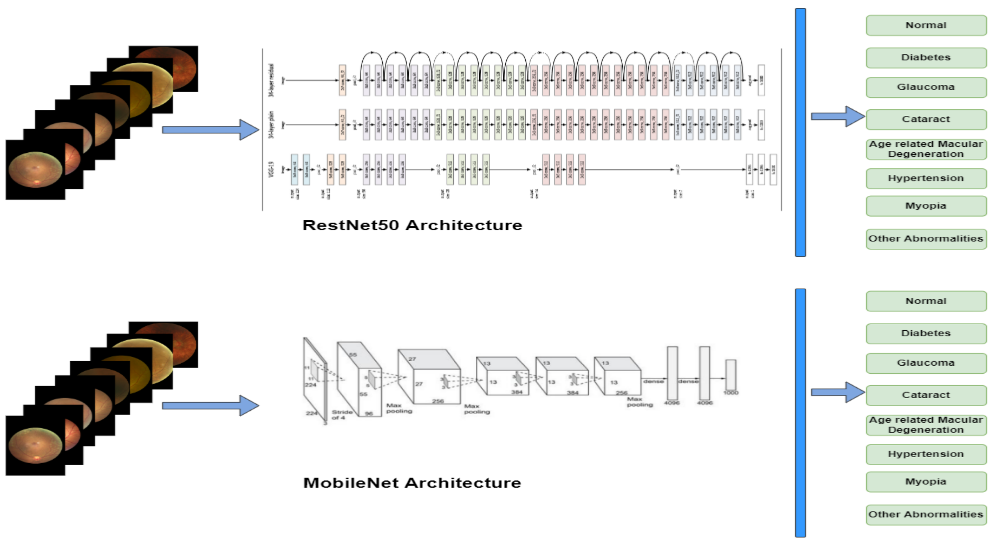


Figure 8: RestNet50 and MobileNet Architectures

# 4 Design Specification

The Design Specification section provides the information of how the research process is designed which is described in Figure 9. This provides the information of framework, architecture, methodologies and techniques used. The data is extracted from the public source kaggle and the pre-processing steps are done in getting the data for research which is explained in the section of Data preparation. The design involves 2 phases of representation where getting the data, pre-processing and augmentation is done in the phase 1. The model classification along with feature extraction techniques is done in the phase 2. The detection of ocular diseases is carried out by 6 different algorithms such as InceptionV3, AlexNet, RestNet50, MobileNet, VGG-16 and VGG-19. The performance metrics are evaluated with accuracy, AUC value and precision.

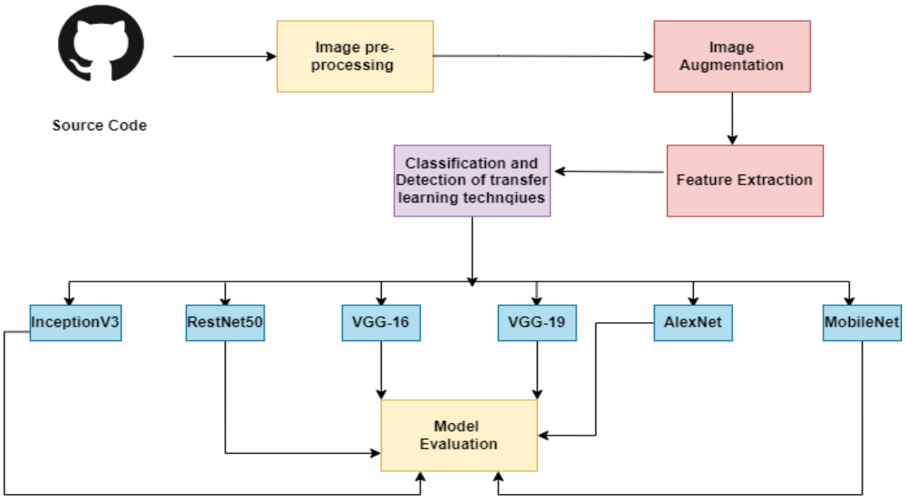


Figure 9: Design Specification

# 5 Implementation

The section of implementation describes how the research has been implemented with help of tools, languages used in computing and data transformations used in this research.

## 5.1 Environmental Setup

The deep learning model execution requires a huge setup of processors with high Graphical processing Unit (GPU) and RAM specifications of computing resources. So, the use Integrated Development Environment(IDE) of Google-Collab is used for performing the code execution using the python language. The use of Google-Collab provides smooth execution of codes involving deep learning techniques since its organised with high GPU based via cloud. The NVIDIA GPU of 27.5GB of RAM of available RAM is utilised through the Google Collab Pro subscription. The Tensorflow package is used along with Keras library for the implementation and building the desired models using Python language.

## 5.2 Data Handling

The data handling method illustrates how the pre-processing techniques are done. The dataset is stored in Google Drive as the software Google collab is used for running the programming. The dataset consists of 8000 images The dataset is being into two parts as per the methodology which is training and testing with the ratio 80% and 20%.The use of Hot encoding technique for data cleaning is described in the section 3.3.1 Data cleaning. The Imagecrop function is used for cropping the images. The ImageResizer along with cv2 package is used for resizing the images. The data augmentation process is done by the use of DataAugmentationStrategy function which are all described in the section 3.3 Data preparation. The parameter tuning and data is normalised. The feature extraction of the images is done before designing the model of transfer learning algorithm used. The research provides the strategy of how a good input data can be utilised with steps of data cleaning using segregation of data accordingly for the transfer learning models.

## 5.3 Transfer Learning

The characteristics of transfer learning is, all models in the methodology are pre-trained models and it has the ability to tackle all kinds of problems which leads to solving the problems created by its own method. The pre-trained models were created using the image fundus datas and the labelling is done using the NumPy. The entire input data is provided to the network and the output provides the label in which the category of the detected diseases is displayed and predicted. The models that are under consideration for this research involves : VGG-16, VGG-19, InceptionNetV3, RestNet50, AlexNet, MobileNet. All the models which is described proves to be more computative and significant that can be used for the future works of development in the technological trends such as Mobile apps etc. In each model the layers are extracted in which the existing network gets new layers getting added and providing new predictions from the old features.

# 6 Evaluation

This section provides the comprehensive description what the performance analysis which is evaluated from the research depending on the experiments of models executed. The metrics of recall, precision, accuracy and AUC-value for each other is analysed .Huang and Ling (2005). The confusion matrix for each model is generated individually with the formula involving true and false - positives and negatives respectively. The classification of taking only positive samples and neglecting the negative samples is known as recall. The accuracy provides the performance of classification in all cases. For each, the calculation of performance metrics is designed using the keras library of the neural network.

## 6.1 InceptionNetV3

The model of InceptionNetV3 is imported in Google Collab from the keras library. The Dense, GlobalAveragePooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.01 with model compiling using binary cross entropy function . 2 Dense layers are used with one layer of 1024 and activation function as "ReLu" and the other layer for prediction is used with "sigmoid" function. The model results are provided in the Figure 10 with results. The model produced the accuracy of 92.1%.

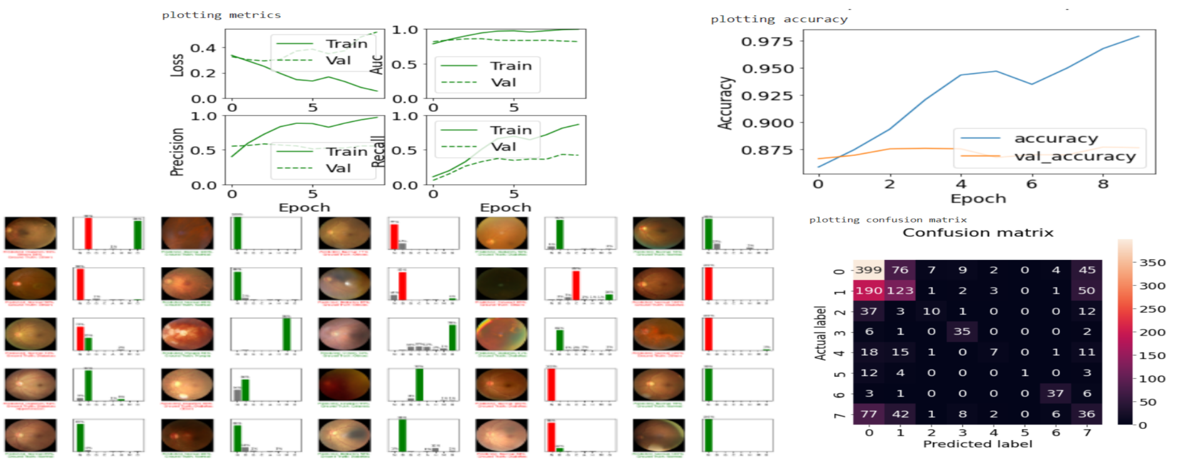


Figure 10: Results of InceptionNetV3

## 6.2 VGG-16

The model of VGG-16 is imported in Google Collab from the keras library. The Dense, GlobalAveragePooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.01 with model compiling using binary cross entropy function . 2 Dense layers are used with one layer of 1024 and activation function as "ReLu" and the other layer for prediction is used with "sigmoid" function. The model results are provided in the Figure 11 with results. The model produced the accuracy of 88.3%.

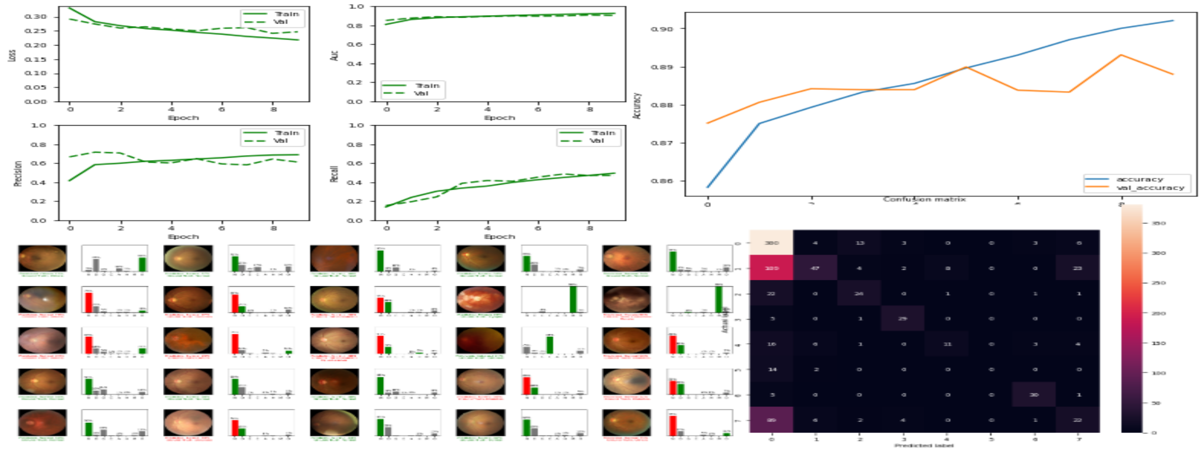


Figure 11: Results of VGG-16

## 6.3 VGG-19

The model of VGG-19 is imported in Google Collab from the keras library. The Dense, GlobalAveragePooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.01 with model compiling using binary cross entropy function . 2 Dense layers are used with one layer of 1024 and activation function as "ReLu" and the other layer for prediction is used with "sigmoid" function. The model results are provided in the Figure 12 with results. The model produced the accuracy of 88%.

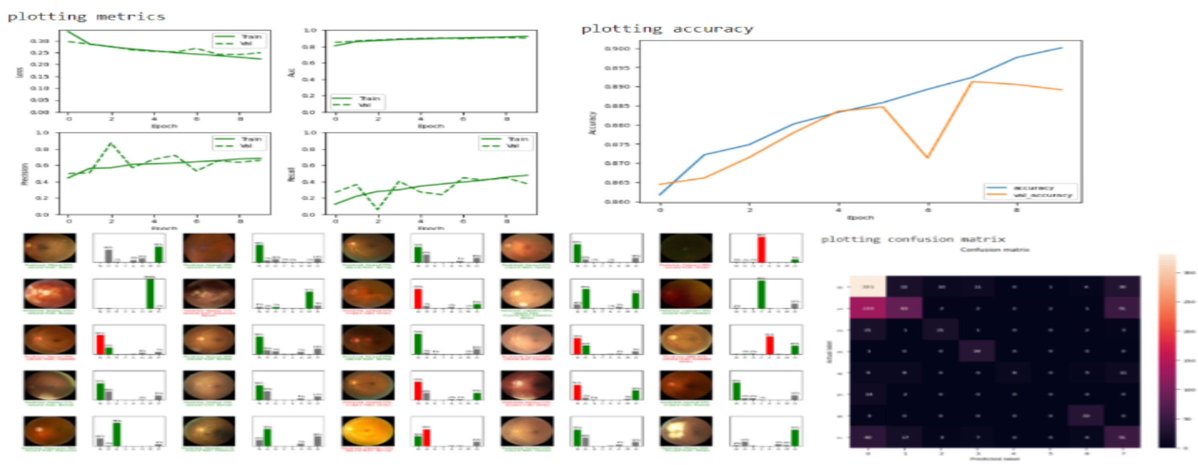


Figure 12: Results of VGG-19

## 6.4 RestNet50

The model of RestNet50 is imported in Google Collab from the keras library. The Dense, GlobalAveragePooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.0001 with model compiling using binary cross entropy function . 2 Dense layers are used with one layer of 1024 and activation function as "ReLu" and the other layer for prediction is used with "sigmoid" function. The model results are provided in the Figure 13 with results. The model produced the accuracy of 86.9%.

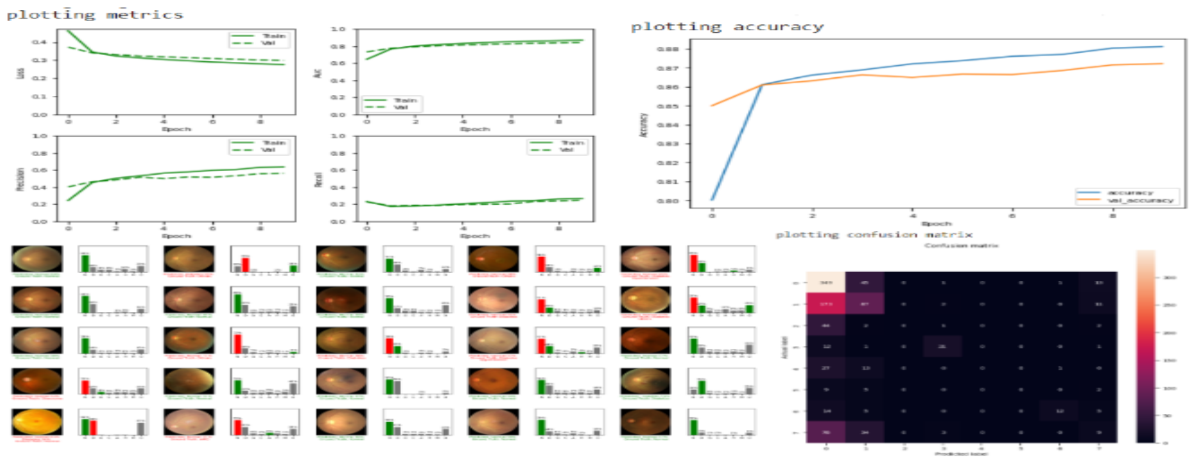


Figure 13: Results of RestNet50

## 6.5 MobileNet

The model of MobileNet is imported in Google Collab from the keras library. The Dense, GlobalAveragePooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.01 with model compiling using binary cross entropy function . 2 Dense layers are used with one layer of 1024 and activation function as "ReLu" and the other layer for prediction is used with "softmax" function. The model results are provided in the Figure 14 with results. The model produced the accuracy of 90.7%.

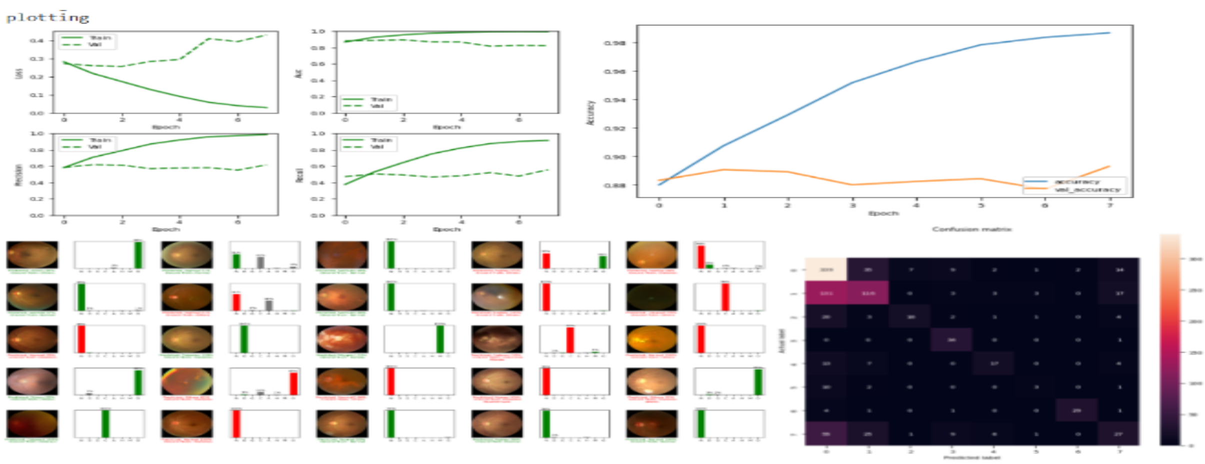


Figure 14: Results of MobileNet

## 6.6 AlexNet

The model of AlexNet is imported in Google Collab from the keras library. The Dense, Activation, Dropout, Flatten, Conv2D, MaxPooling2D layers are imported along with the use of SGD optimiser with the learning rate of 0.01 with model compiling using binary cross entropy function . 5 convolution layers are created along with 2 Dense layers with layers count of 4096 and 1000 are used with one layer of 1024 and activation function as "ReLu" followed by flatten layer and Dropout layers and the output layer for prediction is used with "sigmoid" function. The model results are provided in the Figure 15 with results. The model produced the accuracy of 86.9%.

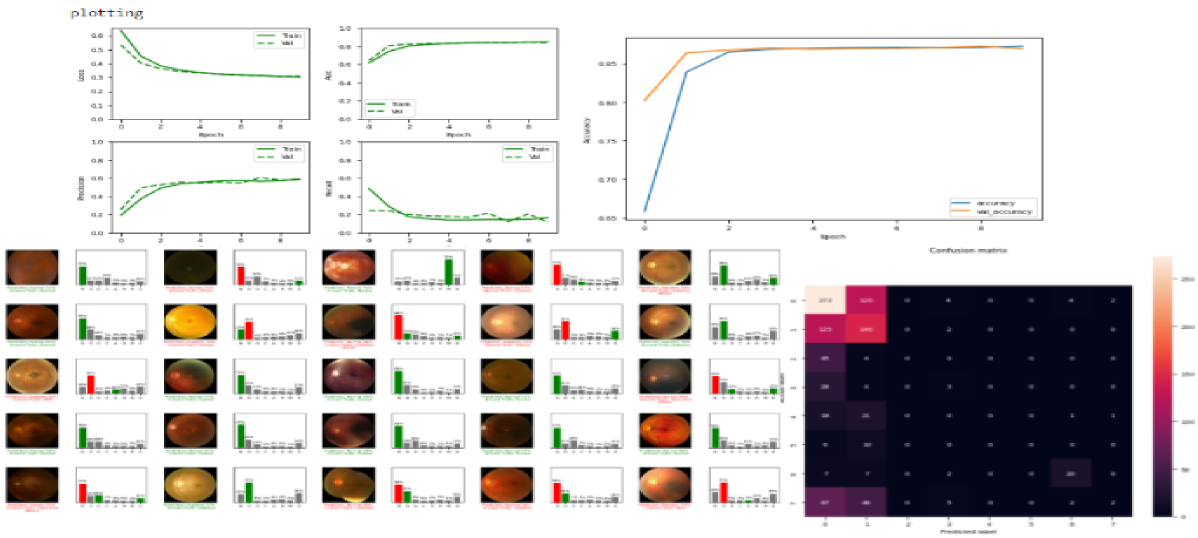


Figure 15: Results of AlexNet

# 7 Discussion

The goal of the research is to examine the multiple classification of ocular diseases such as Glaucoma, Diabetes, Pathological Myopia, Cataract, Age related Macular Degeneration, Hypertension and other eye related diseases accurately by predicting the type of ocular diseases and detecting it so that it will be useful in the medical field of ophthalmology. The performance of all the models provide an overview how each method is designed and is evaluated using the metrics of accuracy and the prediction of the disease by category. The accuracy metrics states that, for all classes the classification is examined precisely.

From the Figure 16 Model comparisons, we can see that that InceptionNetV3 produces the accuracy of 92.1% while the other models such as VGG-16 produces 88.3%, VGG-19 produces 88%, RestNet50 provides 86.5%, MobileNet provides 90.7% followed by AlexNet which produced 86.9% when combined. The examination is also made from the detection of ocular diseases prediction which model has provided best prediction. On comparison with the metrics of evaluation using prediction of ocular diseases charts, accuracy, AUC-value, the InceptionNetV3 model has made the best impact compared to other models.

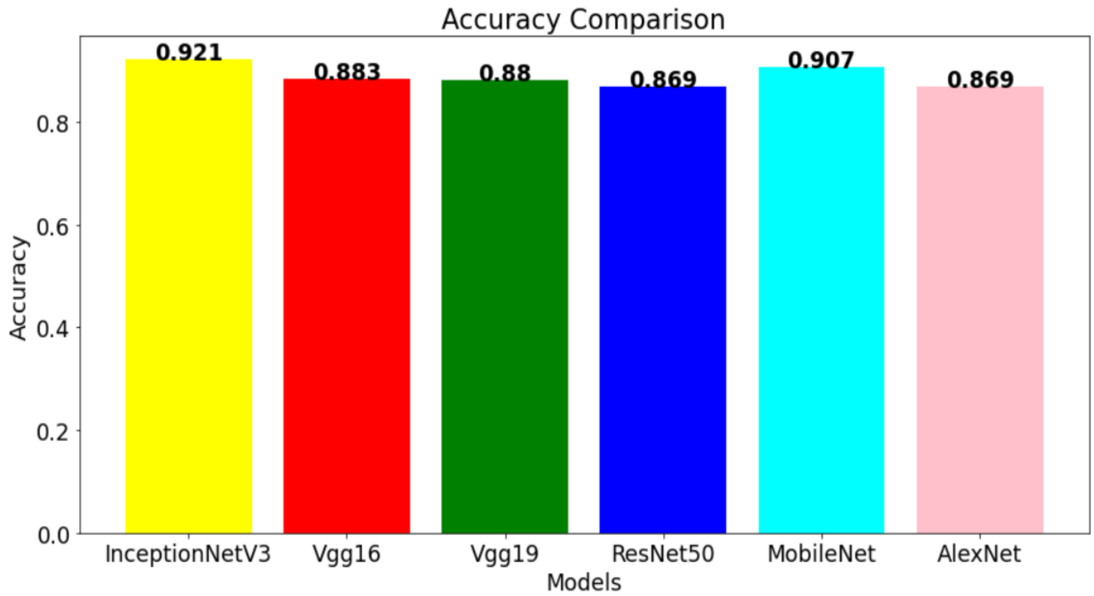


Figure 16: Model Comparisons

# 8 Conclusion and Future Work

From the research done on detection and revelation of multiple ocular diseases using the transfer learning techniques by building the models of InceptionNetV3, VGG-16, VGG-19, MobileNet, RestNet, AlexNet, it is evident that InceptionNetV3 model provided the best result compared to other models. The MobileNet model has also provided the second best in terms of evaluation. The objective of this research, mentioned in section 1.4, states that InceptionNetV3 method of deep learning algorithm can be used for the practical implementation of building the end-to-end design which is more capable of automatic detection of ocular disease when the scanned images of eye is given as input. The future work would be creation of interactive system and since MobileNet model has provided good accuracy on this research, which can also be improved more tweaking of data is done in terms specific techniques. Also the MobileNet model has an advantage of designing a mobile application which could be more handy and will be an added feature in terms of technology and modern trends in which the interactive system can be in much advanced smart automation.

# 9 Acknowledgment

The proposed research has been completed successfully is only with the help of my **Prof. Jorge Basilio.** I would be like to acknowledge my gratitude and thank him for shaping this research,by providing valuable insights and helping with proper feedbacks and guidance. It was a great experience to work under him.

# References


Araújo, T., Aresta, G., Mendonça, L., Penas, S., Maia, C., Carneiro, , Mendonça, A. M. and Campilho, A. (2020). Data augmentation for improving proliferative diabetic retinopathy detection in eye fundus images, *IEEE Access* **8**: 182462–182474.

Chai, Y., Liu, H. and Xu, J. (2018). Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models, *Knowledge-Based Systems* **161**: 147–156.

Chen, X., Xu, Y., Kee Wong, D. W., Wong, T. Y. and Liu, J. (2015). Glaucoma detection based on deep convolutional neural network, *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 715–718.

Elakkiya, B. and Saraniya, O. (2019). A comparative analysis of pretrained and transfer-learning model for automatic diagnosis of glaucoma, *2019 11th International Conference on Advanced Computing (ICoAC)*, pp. 167–172.

Fu, H., Cheng, J., Xu, Y., Zhang, C., Wong, D. W. K., Liu, J. and Cao, X. (2018). Disc-aware ensemble network for glaucoma screening from fundus image, *IEEE Transactions on Medical Imaging* **37**(11): 2493–2501.

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

Govindaiah, A., Hussain, M. A., Smith, R. T. and Bhuiyan, A. (2018). Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images, *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*, pp. 1525–1528.

Guergueb, T. and Akhloufi, M. A. (2021). Ocular diseases detection using recent deep learning techniques, *2021 43rd Annual International Conference of the IEEE Engineering in Medicine  Biology Society (EMBC)*, pp. 3336–3339.

He, J., Li, C., Ye, J., Qiao, Y. and Gu, L. (2021). Multi-label ocular disease classification with a dense correlation deep neural network, *Biomedical Signal Processing and Control* **63**: 102167.

He, J., Li, C., Ye, J., Wang, S., Qiao, Y. and Gu, L. (2020). Classification of ocular diseases employing attention-based unilateral and bilateral feature weighting and fusion, *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pp. 1258–1261.

Huang, J. and Ling, C. (2005). Using auc and accuracy in evaluating learning algorithms, *IEEE Transactions on Knowledge and Data Engineering* **17**(3): 299–310.

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)*, pp. 59–63.

Karkuzhali, S., Mishra, A., Ajay, M. S. and Wilson Prakash, S. (2020). Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models, *2020 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–9.

Khan, I. A., Sajeeb, A. 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)*, pp. 491–494.

Mikołajczyk, A. and Grochowski, M. (2018). Data augmentation for improving deep learning in image classification problem, *2018 International Interdisciplinary PhD Workshop (IIPhDW)*, pp. 117–122.

Paik, I. and Ataka, R. (2019). Adaptable deep learning generation by automatic service composition, *2019 IEEE International Conference on Web Services (ICWS)*, pp. 437–439.

Peng, Y., Dharssi, S., Chen, Q., Keenan, T. D., Agrón, E., Wong, W. T., Chew, E. Y. and Lu, Z. (2019). Deepseenet: A deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs, *Ophthalmology* **126**(4): 565–575.

Pratt, H., Coenen, F., Broadbent, D. M., Harding, S. P. and Zheng, Y. (2016). Convolutional neural networks for diabetic retinopathy, *Procedia Computer Science* **90**: 200–205.

Ram, A. and Reyes-Aldasoro, C. C. (2020). The relationship between fully connected layers and number of classes for the analysis of retinal images.

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.

Wang, D. and Wang, L. (2019). On oct image classification via deep learning, *IEEE Photonics Journal* **11**(5): 1–14.

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.