Knowledge Distillation of the ResNet50 Model for Ocular Diseases Analysis

MSc Research Project
Data Analytics

Saurabh Thakare
Student ID: x20196415

School of Computing
National College of Ireland

Supervisor: Dr. Jorge Basilio

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Saurabh Thakare
Student ID:	x20196415
Programme:	Msc. Data Analytics
Year:	2021/2022
Module:	Msc. Research Project
Supervisor:	Dr. Jorge Basilio
Submission Due Date:	15/08/2022
Project Title:	Knowledge Distillation of the ResNet50 Model for Ocular Diseases Analysis
Word Count:	8276
Page Count:	21

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:	Saurabh Thakare
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):	

Knowledge Distillation of the ResNet50 Model for Ocular Diseases Analysis

Saurabh Babarao Thakare
x20196415

Abstract

Massive amounts of annotated training data contribute to deep convolutional neural network development. It is usually difficult and expensive to obtain data annotations or classification in practice. Deep classification models are prone to overfitting when training with scarce amounts of medical images. Modern deployment demands require interpretability in cases involving computational speed, memory, and complexity of algorithms where performance alone is not enough to satisfy practical needs. In order for deep neural networks to progress further, it is necessary to understand their complex models, how much computation speed and memory are consumed by models, and how models will be portable with different devices. In order to resolve this issue, this paper used the Knowledge distillation of the ResNet50 method using the ocular diseases dataset and compared and visualized the results. The study thoroughly analyzed the result of the model with knowledge distillation and without knowledge distillation of the ResNet50 model in various temperature conditions. This paper is able to create a small student CNN12 model using the knowledge distillation of the teacher ResNet50 model. At t=50, t=70, t=90, and t=100, the accuracy of CNN12 is achieved is 51.4%, 60.9%, 63.7%, and 76.5%, respectively and this accuracy compared with teacher ResNet50 model.

Keywords Knowledge Distillation (KD), ResNet50, CNN12, Ocular Diseases, data augmentation.

1 Introduction

Ocular diseases are one of the main health issues in the present century. It can be noted that a large number of people are suffering from ocular diseases. These include children, young adults, middle-aged persons, and elderly people all around the globe. These diseases are mainly brought about by the changes in the climate of the earth, as well as by the improper routines and foul practices of the people. Excessive reading or excessive use of smartphones is also responsible for ocular diseases among people. It can be noted that the fundus images regarding a given dataset can be split into 4 different classifications and the classification consists of cataract, glaucoma, retina diseases, and normal eyes. This study revolves around the classification of ocular diseases depending on the fundus images, by using the knowledge distillation of the ResNet50 model. This study provides a clear understanding of the concept of ocular diseases, the ResNet50 model, and the knowledge distillation process. Ocular diseases are those diseases that are related to problems in the eye. However, some ocular diseases can be prevented or the chances of getting affected by these ocular diseases can be eliminated by following some good practices. The first and foremost thing which can be done for eliminating the chances of getting affected by ocular diseases is by eating plenty of green vegetables. According to (Spinozzi et al.; 2021), eating fish that are oily can also help in preventing the chances of getting affected by ocular diseases. Moreover, it can also be noted that eating fruits that are rich in Vitamin C, can also prove beneficial in eliminating the chances of getting affected by ocular diseases, like the ones that are mentioned above. Another way by which ocular diseases, like cataracts, can be prevented is by wearing sunglasses. Wearing sunglasses can aid in preventing harmful UV rays from affecting the eyes. Based on the words of (Delmas et al.; 2021), it can be observed that smoking is also responsible for many ocular diseases. This is because smoking can affect the optic nerve and lead to cataracts. Therefore, smoking can be avoided to eliminate the chances of getting affected by ocular diseases.

The various problems which are related to the eyes are mainly considered to be ocular diseases. According to (Park et al.; 2019), ocular diseases often do not have any symptoms. Therefore, these diseases are very hard to detect in their early stages, until they reach an advanced stage. It has also been found that age-related macular degeneration which is included in ocular diseases, can lead to blindness that cannot be treated. It implies that this ocular disease can lead to permanent blindness. Immune dysregulation can be thought to be one of the main reasons for this ocular disease. Medical professionals and medical resources have always been in short supply, which has caused many problems in recent years. These problems have been alleviated to a certain extent by applying machine learning to the medical field, but their shortcomings persist, such as their inability to deploy models on lightweight equipment, and their difficulty in sharing. Knowledge distillation is one of the many methods proposed by researchers. The medical field has widely used knowledge distillation as a technology for compressing and accelerating models. According to many researchers, knowledge distillation can also be used to compress and improve models' performance that is complex and huge. The detection of ocular diseases with the help of knowledge distillation of the ResNet50 model is discussed in this study along with the ways by which the complexity of the model can be reduced. This research is utilizing the knowledge distillation method to reduce the computation processing power, memory, and complexity of the model and also helped to develop a less complex CNN model from the ResNet50 model with good model performance.

1.1 Research Background and Motivation

Ocular diseases are of various types. These can be cataracts, macular degeneration, glaucoma, lazy eye, crossed eyes, ocular hypertension, and so on. The symptoms of ocular diseases are many. Double vision and sudden pains in the eye are considered to be the main symptoms of ocular diseases. Based on the findings of (Ajith; 2020), it can be noted that the other symptoms of ocular diseases swelling in the eyes, redness of the eyes, and severe intensity to light. Changes in the color of the Iris and white patches on the pupil of the eye are other symptoms of ocular diseases. Itching of the eyes and the heavy discharge from the eyes are also sometimes found to be the symptoms of ocular diseases. Ocular diseases mainly affect children and aged people. This is because, children and the young generations are mostly involved in studying for long hours, or using smartphones for long durations. However, in the case of the elderly population, ocular diseases occur mainly due to the degradation of the tissues and nerves in the eyes.

It has also been found that age-related macular degeneration which is included in ocular diseases, can lead to blindness that cannot be treated. It implies that this ocular disease can lead to permanent blindness. Immune dysregulation can be thought to be one of the main reasons for this ocular disease. Glaucoma is another ocular disease that is generally brought about by damage to the optic nerve (Ulhaq et al.; 2020). Many people have high blood pressure in the eye, which leads to damage to the optic nerve, resulting in glaucoma. Cataract, which is another ocular disease, is one of the main reasons, in the current century, for blindness among elderly people. Keratoconus is another type of ocular disease that results in the modification of the shape of the cornea. The original shape of the cornea changes in this disease. Uveitis is another ocular disease in which the inner parts of the eye swell. Colour blindness is also considered to be an ocular disease, where an individual fails to see and identify the proper colors of the objects around. Based on the words of (Guarnieri et al.; 2020), all ocular diseases cannot be prevented.

A combination of Support Vector Machines (SVM) and Convolution Neural Networks (CNN) has been proposed by researchers to overcome the challenge of illness detection (Pahuja et al.; 2022). Deep learning and neural networks were supposed to extend and improve machine learning techniques.(Li et al.; 2020) and (Yang et al.; 2016) found that DC networks were superior for classification, as were aggregations of various deep learning models. Prior to the development of multiple-class eye disease classification methods, researchers used methods that used high computing memory and power. In this paper, the knowledge distillation method is used to reduce the complexity of the model and minimize the computing power and memory usage, so that the classification algorithm for ocular diseases will work with less processing power and memory devices such as mobiles.

1.2 Research Question

How does this research use the knowledge distillation method to develop a simple CNN12 model by reducing the complexity of the large ResNet50 model using the multi-labeled ocular diseases images?

1.3 Research Objectives

1. To evaluate the understanding of knowledge distillation of the ResNet50 method using the ocular diseases images.

2. Ocular image data preprocessing and balancing using the data augmentation method.

3. Implementation of CNN small student model from ResNet50 teacher model using knowledge distillation on a dataset of ocular diseases images.

4. Comparison of teacher model, student model, and distiller model at different temperatures(t) values and accuracy metrics.

5. Comparison of teacher model, student model, and distiller model on the basis of a balanced and unbalanced dataset of ocular diseases images with previous studies.

1.4 Plan of Paper

The 1st section of the research paper outlines the introduction and motivation. The 2nd section of the research paper outlines the literature review. An explanation of the methodology for a research paper can be found in the 3rd section. A description of the design specification can be found in the 4th section. The 5th section explains the implementation and evaluation part of the research. The comparison of results is discussed in the 6th section. The final section discusses the conclusion and future work.

2 Literature Review

This section highlights earlier research on ocular diseases and knowledge distillation. Multiple subsections try to compensate for this section. The prior research is explained in each subsection using a different methodology.

2.1 Ocular Disease Classification Based on Fundus Images

One of the most difficult as well as difficult challenges for an ophthalmologist is too quickly diagnosing and screening retinal disease based on fundus images. Different ocular diseases are capable of causing irreversible as well as permanent damage to the vision of the patients and in severe circumstances, it may potentially cause blindness. Moreover, an automated retinal disorder requires a diagnostics device as well as a strong model that has been properly trained on the various illness related to ocular, therefore it can detect an illness from the various color images of the fundus. The fundus images help in identifying distinct classes related to ocular disease which include diabetic macular edema, no pathology, dry-aged related macular degeneration, and many more. The fundus images are known to be effective models for the purpose of an effective specific and particular classification of the various segments related to the ocular disease as well as retinal vessels (Wang et al.; 2020). The diseases related to ocular are being primarily diagnosed by the help of using color fundus photography or CFP and these respective techniques are being utilized for the purpose of recording various interior surfaces of the human eye, therefore various types of ocular disease are being detected effectively (Nazir et al.; 2021).

Deep learning and computer vision can help automatically detect various types of ocular disease after offering excellent medical retinal fundus images. Fundus images are being captured with the help of various cameras such as Zeiss, Kowa, Canon, and many more resulting in varied resolutions of images (Milea et al.; 2020).). Detecting cataract, as well as myopia, is a considerably simpler process as a result these respective images differ significantly from one another and the standard fundus. Pathological Myopia is a type of progressive and severe nearsightedness which are being characterized by various changes in the fundus of the eye. The various structure of the fundus images which help in the process of examining is retinal parenchyma, retinal vessels, papilla or OD, macula, and many more. Identification of ocular disease one at a time based on the various fundus images is quite limited as the fundus images may reveal many types of conditions. Through the help of various data sets, identification of ocular disease has become easier and various types of ocular disease have been classified without any problems and challenges. It is a difficult task of identifying diabetics early by employing retinopathy images. The blood vessels visible in fundus images have been used in several disease diagnosis techniques. Through the help of various fundus images or photography, various ocular disease has been identified with help of using the appropriate cameras (Sarki et al.; 2020).

2.2 Deep Neural Network Based Classification Algorithms for Ocular Disease Classification

It is a difficult task of identifying diabetics early by employing retinopathy images. The blood vessels visible in fundus images have been used in several disease diagnosis techniques. Several old methods fall well short in recognizing Hard Executes (HE) prevalent in retinopathy and images which are used to evaluate the diabetic disease's intensity (Armstrong and Lorch; 2020). To get around this challenge, the proposed research project utilizes deep networks to extract the features. Using convolutional neural networks (CNN) can generate the focused support plan and helpfully channel the ideas about considerable initiations that can be channelized in this prospect. In the images for mild DR, the microscopic aneurysms could be evident at the beginning of the change from a healthy to a sick situation. The degree of the diabetes illness could be classified using the confusion matrix recognition data. By utilizing the suggested CNN framework, the HE found in a right eye's blood vessel enabled the early detection of the diabetic state.

The intervention to address could also be used to determine if a person has diabetes (Islam et al.; 2019). This particular context provides evidence that the proposed framework has higher accuracy than other conventional detection algorithms. It is a state-related to diabetics that have irreparably damaged human vision. The prognosis is most effective while diabetic retinopathy is still in its initial stages. Diabetes regularly appears as an appearance in the retinal fundus. Diabetic retinopathy could lead to total blindness (DR) (Akinyelu and Blignaut; 2020). Although DR is a progressive process, medical professionals advise that people with diabetes should get their health tested at least twice a year for any signs of sickness. The optical scientist will evaluate the color image of the fundus in diagnosing, which would be the main factor in recognition. This laborious and time-consuming monitoring process has the potential to produce a serious error.

Due to a lack of medical care facilities in some locations, a large number of DR patients do not obtain early treatment, lose their best therapeutic chances, and experience irreversible vision loss (Saxena et al.; 2020). If the DR is promptly diagnosed and managed, the procedure can be successfully controlled and prolonged, linked to premature disorders. The clinician's expertise determines the impact of manual interpreting. When individuals do not contact their doctors, misunderstanding is commonly occurring. Developments in health will be greatly affected by existing ICT developments. In integrative machine learning, pictures can be recognized and interpreted in addition to data representations that underpin the measures, specifically in the context of artificial intelligence (Kumar et al.; 2020). Remarkable observations have been made in the healthcare sector. The discipline of deep learning has a multilayered, data-analysis-friendly design. As sonography becomes more complicated, computers nowadays are capable of classifying medical pictures by applying a variety of methods including examination, verification, detection, and categorization using a deep learning approach. It needs vast experience with fewer prospects which can be channelized by connecting with the projected support plan with constructive directions (Tong et al.; 2020).

The entire challenge could be effectively solved by the process of an automatic approach for grading diabetic retinopathy. Several methods have been taken into consideration in this prospect, where the standard and deeper learning seems to produce efficient and potential outcomes. Conventional techniques including morphological operations, random forest segregation, and fluorescence procedures have shown to strengthen the effects of deep learning algorithms (Pervaiz et al.; 2021). For the diabetic retinopathy analysis, it is highly effective to follow a VGG16 method for precise classification and this prospect connects with the use of action plans and caters to focused processes so that different classifiers get channelized and accessed (Gour and Khanna; 2021).

2.3 Classification of Ocular Diseases Using ResNet Model

A specific example of the CNN network, the ResNet model is mostly utilized for feature extraction and classification. The model of ResNet34 has been used for the purpose of identifying and detecting various types of disease related to ocular. The ResNet34 model provides accuracy in the various results without providing any false results of detection. Among the various ocular diseases, cataract is one of the prevalent diseases. An early diagnosis of cataracts can help in the process of reducing the rates of cataracts in the overall world. The ResNet model helps in the process of early detection of various ocular diseases (Jiang et al.; 2019). The model ResNet does not require heavy as well as expensive equipment based on the images of Fundus. The ResNet model not only aims to achieve accuracy in the prediction but also aims to ease some of the issues and problems which are present in the current method of analyzing and treatment of the various types of ocular disease.

ResNet helps in surpassing the human accuracy of image recognition (Haraburda and Dabała; 2022). The model of ResNet34, as well as ResNet50, are being used for the purpose of proper detection of various ocular diseases. The deep features of ResNet50 along with random forest help in the process of classifying as well as detecting the grading of retinopathy as well as diabetes. A high level of features is determined by the average pooling layer from the trained Resnet50. The ResNet framework structure accelerates the learning of various extremely deep networks of neural and helps in increasing the model performance on the extensive training data. The present model of ResNet34 and ResNet50 helps in identifying various types of diseases such as cataracts, glaucoma, and other retinal diseases. Most of the proposed architecture uses the deep features of the ResNet50 in the combination with a random classifier of the forest for the purpose of achieving the desired results and outcomes regarding the various ocular disease (Ou et al.; 2022).

2.4 Classification of Ocular Diseases Using Knowledge Distillation

When ocular diseases are not treated timely, they can cause irreversible vision loss. Colour fundus photography is one of many imaging techniques used to detect ocular diseases. Because early stages of the ocular disease have few visible symptoms, it is difficult to accurately diagnose them, and it is very difficult to classify them automatically. An imaging-based automatic ocular disease classification model can be improved with the help of a knowledge distillation-based method presented by (He et al.; 2021). The optimization process consists of sequential optimization of two deep neural networks. Using color fundus images and clinical features provided by radiologists, a teacher network is trained. In order to learn that information from only images, a student network distills the knowledge of the teacher model. The proposed method by researchers, based on extensive experiments, can improve the diagnosis of retinal diseases based on fundus images without the need to use clinical features by largely recovering the performance of the teacher model.

A limitation of the collection of images due to their rarity prevents machine learning models from recognizing those diseases. To overcome this problem (Mai et al.; 2021), formulate the problem caused by the lack of labeled data as a learning task between a student and a teacher based on knowledge distillation and discriminative feature space. Depending on the macular structural changes, the imaging device, and the angle, optical coherence tomography (OCT) images differ greatly. A preprocessing pipeline is utilized first to align images in order to alleviate such problems. A horizontal state can be approximated for better feature representation when comparing tissue images taken from different angles. The effectiveness of the proposed approach has been demonstrated in extensive experiments on our dataset.

Since deep learning methods were successful in detecting eye disease on fundus images, they became popular for detecting ocular disease. Medical experts with not enough time to devote to this task are required to manually label medical images for deep learning models. This requires a large number of labeled images for the models to learn. Due to these limitations, few images are annotated. Through supervised fine-tuning and knowledge distillation with a small set of labeled images,(Arrieta Ramos; 2022) proposed a semi-supervised methodology for detecting diabetic retinopathy using unlabeled images and labeled ones. With only 2% EyePACS-Kaggle train labeled images, this method achieved 89% AUC on the Messidor-2 dataset.

2.5 Research Gap

Even though deep neural networks like VGG16, ResNet34, ResNet50 and transfer learning methods are effective and produce interesting results for a variety of ocular diseases datasets but it is not feasible to deploy them on edge devices, such as embedded sensors, and smartphones. The use of a large, pretrained network (teacher) to train a smaller (student) network has been tried in order to compress these networks. A popular method is called knowledge distillation. In order to make the CNN12 (student) model from the ResNet50 (teacher) model deployable on limited memory and low-processor devices while increasing model interpretability, this study applied the knowledge distillation method to ocular diseases images.

3 Research Methodology

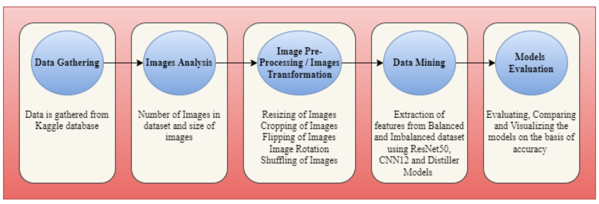

Figure 1: Knowledge Discovery Database Process

Knowledge Discovery in Databases (KDD) is used in this paper to achieve specified research objectives. The KDD process used in this paper is the same as the normal KDD process, with a few modifications (Ahmad et al.; 2022). The KDD process covered in this research includes data gathering, image data analysis, image pre-processing or transformation, data mining, and model evaluation. Above figure 1 showing the KDD process.

3.1 Data Gathering

A dataset of ocular diseases containing fundus images has been gathered from Kaggle. The dataset includes four different ocular disease types: normal, glaucoma, cataracts, and other retinal diseases. The dataset has 601 images in total, which are unequally distributed.

3.2 Images Analysis

The dataset contained four classes labeled as '1_normal', '2_cataract', '2_glaucoma', and '3_retina_diseases'. There are 300 images in the normal class, 101 images in the glaucoma class, 100 images in the cataract class, and 100 images in the class for other retinal diseases. The unequal distribution of images of ocular diseases is shown in the figure 2 below.

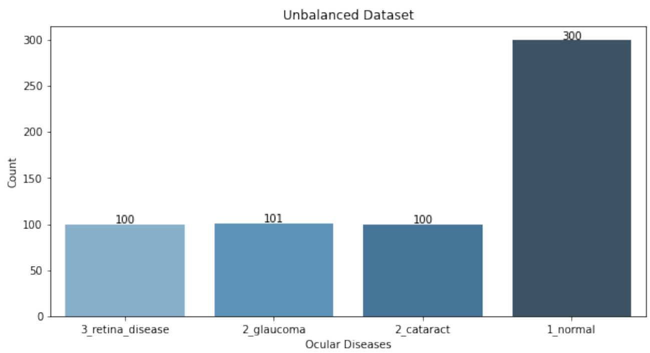

Figure 2: Number of Images in Unbalanced Dataset

3.3 Image Pre-processing and Transformation

The data must be processed in a way that is acceptable to the model, in a format that can be processed, and in a way that is more accurate for the model. A label is given to an image by the folder in which it is stored. In order to label all the images, the folder index was used. Observation reveals that the image's maximum and minimum sizes are (3504 x 2336) and (250 x 188) respectively. The majority of images have high-resolution dimensions. For all images to be the same size, the images have been resized into (256 x 256) sizes. The accuracy of the model is hampered by the black background in the fundus images. The cropping technique is used to remove the images' black backgrounds. The figure 3 displays the resized and cropped images of the original images of ocular diseases.

The imbalanced images of ocular diseases are visible in the figure 2. The accuracy of the model is impacted by the data imbalance, which also causes the model to overfit. The data augmentation method has been used for image balance to solve this issue. The

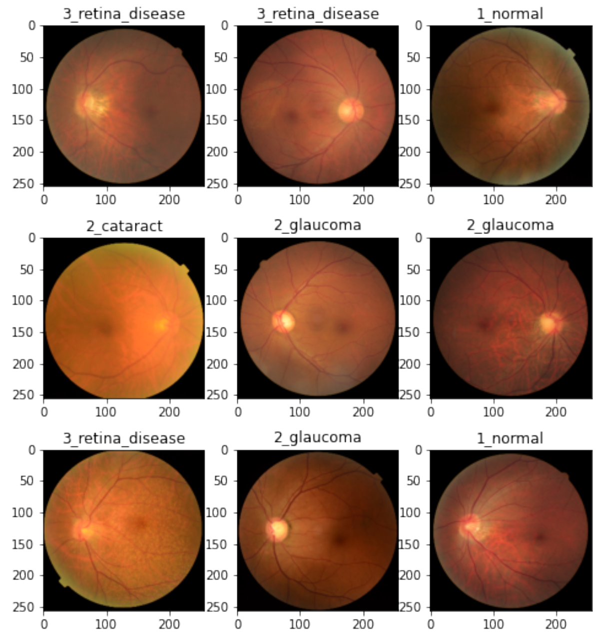

Figure 3: Resized and Cropped Images

images are rotated 180 degrees clockwise and flipped vertically to generate the augmented images. The balanced dataset with 1203 images is shown in the figure 4. In the balanced dataset, there are 300 images in the normal class, 303 images in the glaucoma class, 300 images in the cataract class, and 300 images in the class for other retinal diseases.

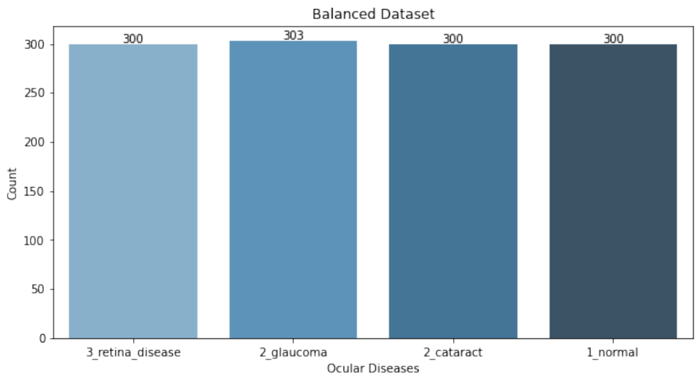

Figure 4: Number of Images in Balanced Dataset

The data is then split into training and testing samples with an 80:20 ratio, after which the images are randomly shuffled. The training dataset contains 80% of the shuffled images, whereas the testing dataset only contains 20%.

3.4 Data Mining

In the data mining stage, features related to ocular disease images are extracted. The feature extraction procedure includes the use of the ResNet50, CNN12, and distiller models. The teacher model is the ResNet50 model. For transfer learning, the ResNet50 model used the pre-trained weight of ImageNet images. The CNN12 model, which does not base on transfer learning, is utilized as a student model. ResNet50 and CNN12 models are used to generate the distiller model. Section 5 provides a thorough discussion of each model and its methodology.

3.5 Models Evaluation

This research's main objective is to derive a less complex model which utilized less computation memory and low power. Therefore, a comparison of teacher, student, and distiller models on the basis of accuracy, parameters, and layers is the best evaluation matrices. Therefore, in this research model evaluation has been done by comparing the training and testing accuracy of the ResNet50 model, CNN12 model, and distiller models. Models are also evaluated on the basis of parameters, and layers used in the model.

4 Design Specification and Architecture

The design specification of the model architecture is covered in this section of the article. This section is broken up into two sections: the model architecture with a brief description of each component used in architecture, and the model compilation specification.

4.1 Model Architecture

As it can be seen from figure 5, the model architecture is divided into 3 stages. Each class of ocular disease images is organized into a separate folder. The image data and labels would be stored in separate tensors in TensorFlow in order to import the data. The architecture is used as a data augmentation block to flip images vertically, and rotate images 180 degrees clockwise. An imbalanced dataset was transformed into a balanced dataset with the aid of the augmentation method. Images of retinal, glaucoma, cataract, and normal are included in a balanced dataset, which has 300, 303, 300, and 300 images, respectively. .

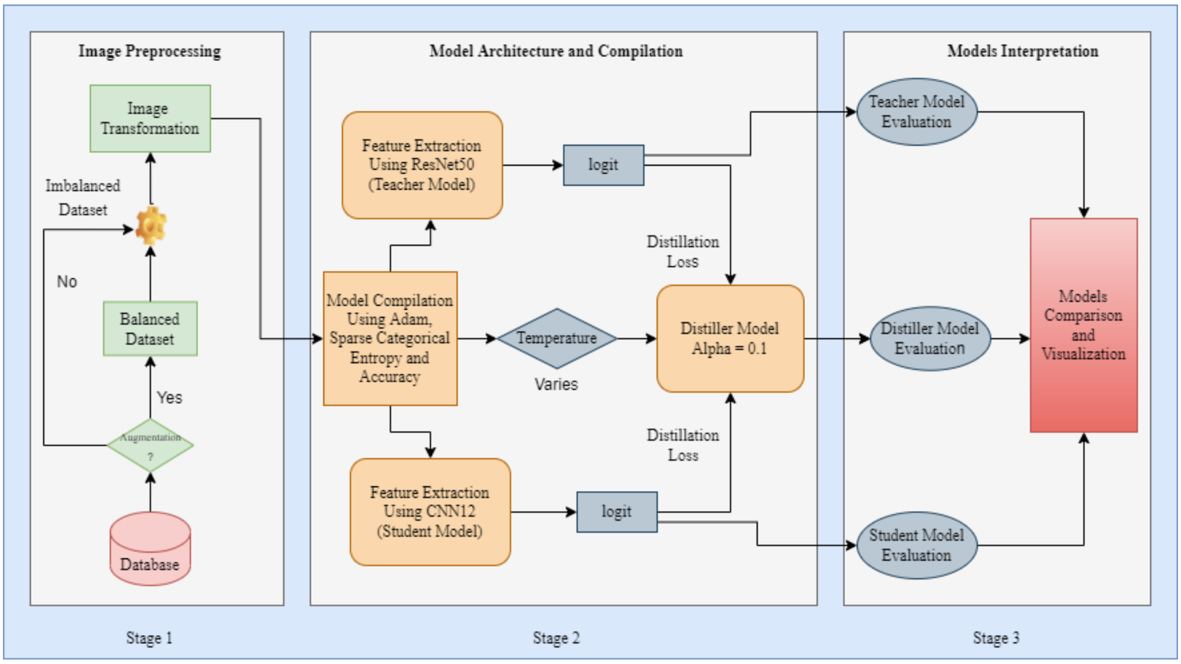

Figure 5: Model Architecture

Images are imported using the TensorFlow library from the balanced dataset after balancing. The first stage of architecture included image resizing into 256 x 256 sizes, cropping to remove black backgrounds, and image shuffling. Moving towards the second stage, the ocular dataset is divided into the train and test samples where the train part contains 80% of the dataset and the test part contains 20%. In order to analyze the ocular images, the train samples were separately trained in the ResNet50 (teacher) model, CNN12 (student) model, and distiller model. In the second stage, the feature extraction part of images has been implemented using various batch sizes and epoch sizes. The distiller model is trained with different values of temperature (t) like t=50, t=70, t=90, and t=100 values to achieve accuracy nearly to the teacher model. The evaluation of each class's images of ocular disease is done in the third stage by utilizing accuracy metrics. At this phase, the ResNet50, CNN12, and distiller models' accuracy and loss measures at various temperature levels have been implemented. The last stage also includes the comparison of the balanced and unbalanced dataset of ocular diseases.

4.2 Model Compilation Specification

Three models: ResNet50 (teacher), CNN12 (student), and distiller are being used in this paper. Each model was designed using a special combination of loss, metrics, and optimizer parameters. The ResNet50 model was compiled with Adam optimizer, Sparse Categorical Cross Entropy loss with using logit values and accuracy metrics. The CNN12 model used the same compiler parameter as ResNet50, whereas the distiller model was compiled with Adam optimizer, Sparse Categorical Cross Entropy student loss, accuracy metrics, KLDivergence distiller loss, alpha=0.1, and different temperature values. While temperature is used to soften the student loss function, alpha is utilized to equate the student loss function to weight and 1-alpha to the distillation loss function. The larger the temperature gives the softer distribution. Higher temperature values indicate a greater transfer of knowledge from the teacher to the student model (Wang and Yoon; 2021). Because this study uses a response-based knowledge distillation method, where the weight of nodes should be constant and the balance of teacher and student loss is based solely on temperature values, therefore the alpha value in this paper is constant and is set at 0.1.

5 Implementation and Evaluation

Throughout this section, the paper discusses how ResNet50, CNN12, and distiller models were implemented and evaluated. The hardware and software tools used to build and run the model are also described in this section.

5.1 Environmental Setup

In order to train the models, this paper used a system equipped with an Intel Core i5 10th Generation processor and an NVIDIA GeForce GTX 1650 Ti GPU. In addition to its 2.50 GHz processor, the system has 8 GB of RAM. The programming is done using Python version 3.8.5 and TensorFlow 2.7.0 on a Jupyter Notebook. TensorFlow's Keras API is used to create the models. The 'thesiskernel' is a new virtual environment created to increase the portability of this research.

5.2 Implementation of Teacher Model (ResNet50)

The ResNet50 model is used to extract features based on transfer learning. Using the pre-trained weights from ImageNet, the model is trained. A model for analyzing ocular diseases was fed 256x256x3 size images of ocular diseases. ResNet50 output is reshaped and passed through a dense layer containing 512 output nodes with Rectified Linear Activation Function (ReLU) followed by an output layer with four classes and SoftMax activation function. The ResNet50 model is created with a total of 90,699,140 parameters whereas 90,646,020 are trainable and 53,120 are non-trainable parameters.

5.3 Evaluation of ResNet50 Model using Unbalanced Dataset

ResNet50 model trained using an unbalanced dataset of ocular diseases which included 601 total images. The model is trained with a batch size of 32 and an epoch size of 25. Each epoch is trained with 15 batch sizes and validated with 4 batch sizes. The average trained accuracy achieved by the model is 68.9% whereas test accuracy is 69.1%. It can be concluded from figure 6, that the model is accurate when the gap between train and validation accuracy continues to decrease with good learning rates for both train and validation losses.

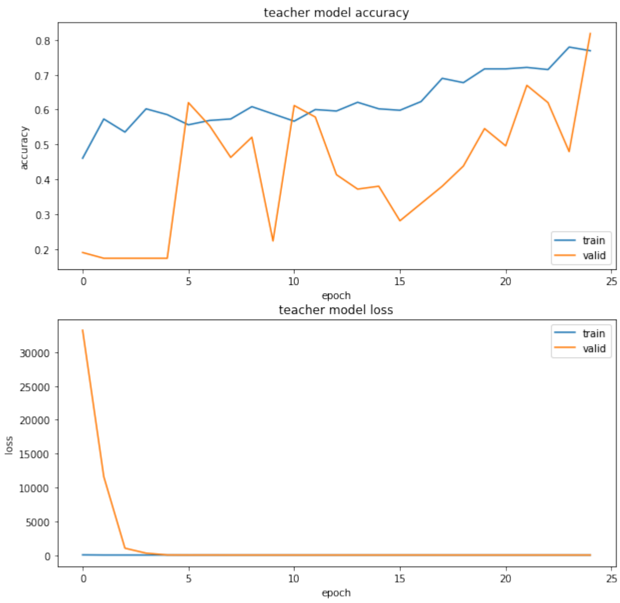

Figure 6: Teacher Model Accuracy and Loss on Unbalanced Dataset

5.4 Evaluation of ResNet50 Model using Balanced Dataset

A balanced dataset of ocular diseases with 1203 total images was used to train the ResNet50 model. A batch size of 32 and an epoch size of 25 are used to train the model. 30 batch sizes are used for training and 8 batch sizes for validation for each epoch. The model average trained accuracy is 82.1%, while its test accuracy is 77.7%. The figure 7 shows that the model is accurate when the difference between train and validation accuracy keeps reducing and the balanced learning rate is more as compared to an unbalanced learning rate.

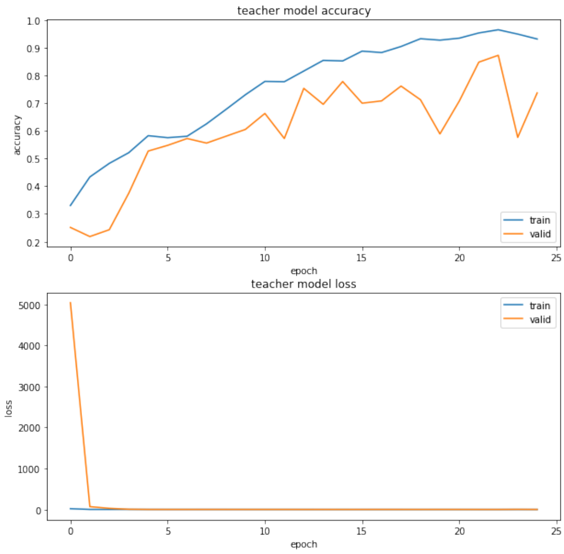

Figure 7: Teacher Model Accuracy and Loss on Balanced Dataset

5.5 Implementation of Student Model (CNN12)

The CNN12 is created using the combination of 2 Conv2D layers with the filter of 32 sizes, 2 Conv2D layers with the filter of 64 sizes, 2 Conv2D layers with the filter of 128 sizes, 3 MaxPool layers with a pooling size of 2x2 and 3 Dropout layer with deactivation rate of the node is 0.2. The size of the kernel used in each Cov2D layer is 3x3. The feature extraction has been done using the CNN12 model. A model was fed 256x256x3 size images of ocular diseases as an input. CNN12 output is reshaped and passed through a dense layer containing 512 output nodes with Rectified Linear Activation Function (ReLU) followed by a dropout layer with a deactivation rate of the node is 0.5. The output from the dropout layer is passed through the output layer with four classes and the SoftMax activation function. The CNN12 model is compiled and trained with a total of 59,271,972 parameters.

5.6 Evaluation of CNN12 Model using Unbalanced Dataset

In order to train CNN12, an unbalanced dataset of 601 images of ocular diseases is used. A batch size of 32 and an epoch size of 20 are used to train the model. A total of 15 batch sizes are used for training and four batch sizes are used for validation for each epoch. As a result, the trained accuracy achieved by the model is 50.4%, while the test accuracy is 49.5%. As can be seen in the figure 8, the model is less accurate compared to the ResNet50 model because both train and validation losses are not decreasing.

5.7 Evaluation of CNN12 Model using Balanced Dataset

The CNN12 model was trained using a balanced dataset of 1203 images of ocular diseases. Models are trained using batch sizes of 32 and epoch sizes of 20. A total of 30 batch sizes are used for training and 8 batch sizes are used for validation for each epoch. Model average trained accuracy is 56.7%, while test accuracy is 52.6%. The trained accuracy of the balanced dataset model grew by over 6%, and test accuracy by almost 3%, as shown in the figure 9, making it more accurate than the unbalanced dataset model. The CNN12

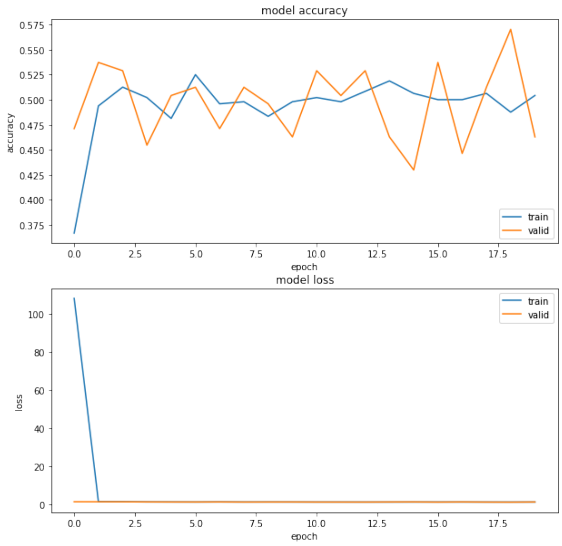

Figure 8: Student Model Accuracy and Loss on Unbalanced Dataset

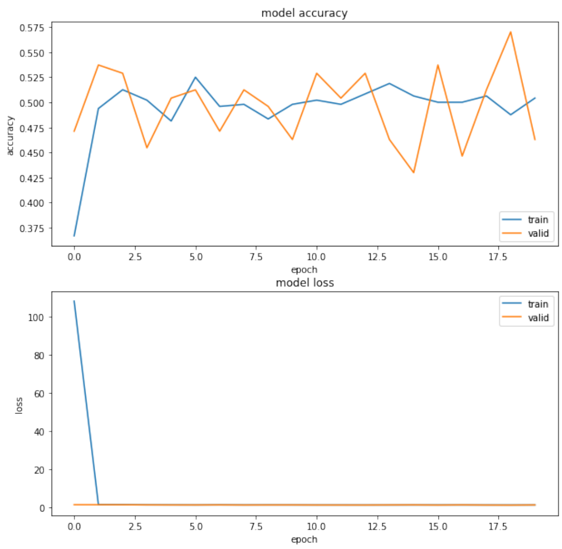

Figure 9: Student Model Accuracy and Loss on Balanced Dataset

model's prediction power, however, is incredibly low when compared to the ResNet50 model.

5.8 Implementation of Student-Distiller Model

The student-distiller model also called as distiller model is made by balancing the loss of the ResNet50 (teacher) model and CNN12 (student) model. The transfer of knowledge from the teacher model to the student model completely depends on the teacher model's final output in this paper, which uses response-based knowledge distillation. In this concept, the student is expected to imitate the predictions of the teacher. In order to accomplish this, a loss function is used, which is the distillation loss. This method evaluates the difference between the logit of ResNet50 and CNN12 models. Using this method, the student model (CNN12) will produce similar accuracy to the teacher model (ResNet50) as this loss is minimized over time, while the temperature parameter is used to soften the accuracy between ResNet50 and CNN12 model. he overall loss of the student-distiller model is defined in this paper using the below expression:

loss = alpha x CNN12 loss + (1- alpha) x distillation loss

Where alpha=0.1, is utilized to equate the student loss function to weight and 1-alpha to the distillation loss function. The distillation loss is calculated using SoftMax values of the ResNet50 and CNN12 models at various temperature values. This paper utilized the KL (Kullback–Leibler) divergence method to match the difference between the ResNet50 and CNN12 model's distillation loss at different temperature values (Kim et al.; 2021).

5.9 Evaluation of Distiller Model using Unbalanced Dataset

In order to train the distiller model, an unbalanced dataset of 601 images of ocular diseases is used. A batch size of 32 and an epoch size of 20 are used to train the model. A total of 15 batch sizes are used for training and 4 batch sizes are used for validation for each epoch. The model has been compiled at 4 different temperature (t) values that are t=50, t=70, t=90, and t=100. The average training accuracy at t=50 is 49.7%, t=70 is 56.1%, t=90 is 68.3%, and t=100 is 97.9%. Whereas test accuracy at t=50 is 47.1%, t=70 is 56.8%, t=90 is 71.9%, and t=100 is 96.6%. As can be seen in figure 10, increasing the value of temperature increases the accuracy of the distiller model. At temperature t=90, the accuracy of the distiller model and ResNet50 model with an unbalanced dataset is similar.

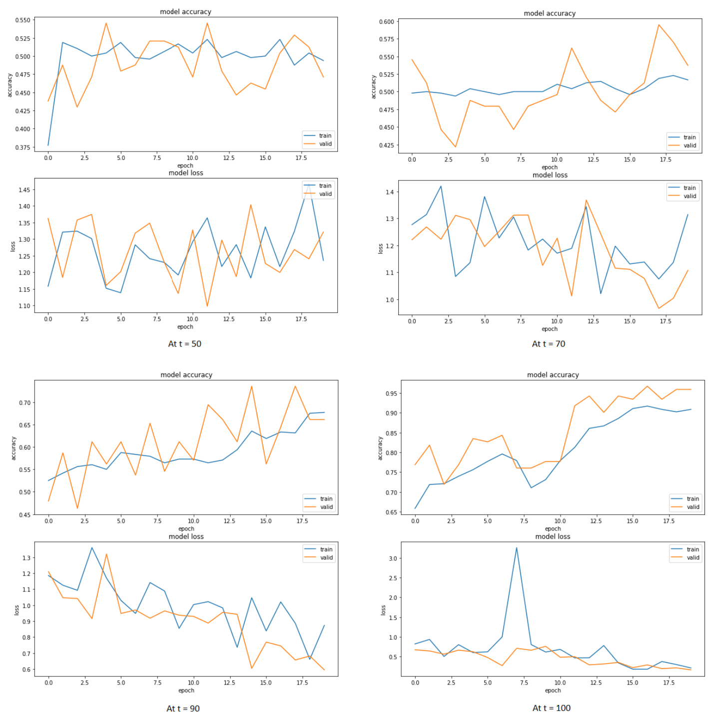

Figure 10: Distiller Model Accuracy and Loss on Unbalanced Dataset

5.10 Evaluation of Distiller Model using Balanced Dataset

The distiller model was trained using a balanced dataset of 1203 images of ocular diseases. Models are trained using batch sizes of 32 and epoch sizes of 20. A total of 30 batch sizes are used for training and 8 batch sizes are used for validation for each epoch. The temperature (t) values used to compile the model are t=50, t=70, t=90, and t=100. Average training accuracy is 53.5% at t=50, 61% at t=70, 71% at t=90, and 82.3% at t=100. The test accuracy is 51.4% at t=50, 60.9% at t=70, 63.7% at t=90, and 76.5% at t=100. The accuracy of the distiller model improves as the value of temperature is increased, as seen in figure 11. The accuracy of the ResNet50 model and the distiller model with a balanced dataset is identical at temperature t=100.

Figure 11: Distiller Model Accuracy and Loss on Balanced Dataset

6 Comparison and Results

6.1 Comparison of Models Between Balanced and Unbalanced Dataset

As can be seen from the table (figure 12), when the model has trained with a balanced dataset of ocular diseases its learning rate increases as compared to the unbalanced dataset. Additionally, it has greater accuracy in a balanced dataset than in an imbalanced one. The table (figure 12) indicates that as the temperature is increased, the distiller model's accuracy likewise rises. The distiller model is trying to achieve accuracy similar to the teacher model (ResNet50). The distiller model reached a testing accuracy equal to the ResNet50 model at t=90 with a difference of 2.80% in the unbalanced dataset, whereas at t=100 with a difference of only about 1.20% in the balanced dataset.

Models / Accuracy	Unbalanced Dataset		Balanced Dataset	
	Training Accuracy	Testing Accuracy	Training Accuracy	Testing Accuracy
ResNet50 (Teacher)	68.90%	69.10%	82.10%	77.70%
CNN12 (Student)	50.40%	49.50%	56.70%	52.60%
Distiller at t = 50	49.70%	47.10%	53.50%	51.40%
Distiller at t = 70	56.10%	56.80%	61.00%	60.90%
Distiller at t = 90	68.30%	71.90%	71.00%	63.70%
Distiller at t = 100	97.90%	96.60%	82.30%	76.50%

Figure 12: Comparison between Balanced and Unbalanced Dataset

The table (figure 12) shows that the test accuracy of a balanced dataset is typically 4% to 5% higher than that of an unbalanced dataset, implying that a balanced dataset is more accurate. The distiller model's test accuracy with balanced and unbalanced datasets crossing between the temperature ranges of t=70 and t=90 is depicted in figure 13. The figure 13 also illustrates how to determine how accurately a balanced dataset differs from an unbalanced one.

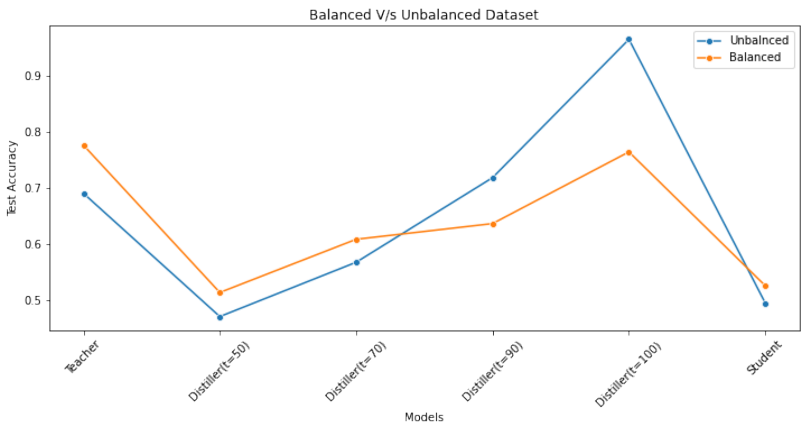

Figure 13: Comparing the Test Accuracy between Balanced and Unbalanced Dataset

6.2 Comparing Distiller Models to Earlier Research

The table (figure 14) shows a comparison between the accuracy found in this research and previous research that has classified fundus images into multiple classes. The knowledge distillation of the ResNet50 model with a balanced dataset resulted in the best test accuracy of 76.50% out of all the models used in this study. From the table (figure 14), it is clear that this research paper model outperforms models that were applied to various datasets in terms of accuracy. While other researchers' accuracy was closer to the range of 70% to 87%, only the (Raza et al.; 2021) Inception v4 model reached an accuracy of 96%. The accuracy of (Sarki et al.; 2022), (Gour and Khanna; 2021), (He et al.; 2021), and (Raza et al.; 2021) is higher than that of this research paper model. The number of images used to train the model may be the main cause of this. Compared to prior research, these models are trained on a comparatively small number of images. Even though there were only 1203 images used in this research, ResNet50's knowledge distillation outperformed (Haraburda and Dabała; 2022), (Mai et al.; 2021), and (Wang et al.; 2020) research models.

Researcher Name	Dataset Name	Size of Dataset	Models	Same/Different Dataset In this Research	Accuracy
Sarki et al. (2022)	Messidor, Messidor-2,DRISHTI-GS, and Kaggle	4348	CNN + RMSprop	Different	81.33%
Haraburda and Dabała (2022)	Pubicly Availble 18 Datasets	30,000	Synergic Deep Learning (DCNN)	Different	70.00%
Gour et al. (2021)	Ocular Disease Intelligent Recognition (ODIR)	5814	VGG16 + SGD	Different	87.16%
He et al. (2021)	Ocular Disease Intelligent Recognition (ODIR)	5814	Knowledge Distillation of ResNet101	Different	83.10%
Mai et al. (2021)	BOE and Cell Database of OCT images	1,09,309	Knowledge Distillation of ResNet50	Different	74.45%
Raza et al. (2021)	Kaggle	1803	Inception v4 + Augmentation	Different	96.00%
Wang et al. (2020)	Shanggong Medical Technology Co. Ltd. (ODIR)	4115	EfficentNetB3	Different	73.00%
This Research (2022)	Kaggle	601	Knowledge Distillation of ResNet50	Same	71.90%
This Research (2022)	Kaggle	1203	Knowledge Distillation of ResNet50	Same	76.50%

Figure 14: Comparing the Accuracy with Previous Research

7 Conclusion and Future Work

As can be seen from the comparison and result section, the best result of knowledge distillation of the ResNet50 model got at temperature 100 with a balanced dataset of ocular diseases. The obtained results at various temperature levels demonstrate that all of the research paper's objectives have been met. The CNN12 distiller network, which is less complex and needs less computing memory and speed, can be created, as per the research question. The best accuracy of this research is 76.50%, which is almost the same as or higher than other previous research. This research has led to the conclusion that knowledge distillation helps in the deployment of complex models on lightweight equipment. From the literature review section, it can be seen that there is little research performed on ocular diseases using the knowledge distillation method. But due to an increase in the lightweight smart device in today's era, the demand for knowledge distillation process is increased. It is clear from the literature review part that there is not much knowledge distillation method research on ocular diseases. But as the use of portable smart devices has grown, the demand for knowledge distillation process is increased. This study has shown that the CNN12 distilling model can be used in lightweight devices, and that lightweight, embedded, or smart edge devices may also be used to classify ocular diseases.

Even so, this research accuracy is good given the limited dataset and epoch size used. However, in the future, the model's accuracy can be improved by fitting it with a huge dataset and a high epoch size. In the future, a more complicated model can take the place of the ResNet50 model to further improve the model's accuracy.

8 Acknowledgement

In order to complete this research project, Dr.Jorge Basilio, Dr.Paul Hayes, and Dr.Rejwanul Haque had to supervise the work. It has been a pleasure working with Dr.Jorge Basilio, Dr.Paul Hayes, and Dr.Rejwanul Haque for the past three months. I appreciate their guidance and advice regarding this research. Special thanks to Jorge Basilio for supporting me throughout the entire research process over the last month.

References

Ahmad, Z., Yaacob, S., Ibrahim, R. and Fakhruddin, W. F. W. (2022). The review for visual analytics methodology, *2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, IEEE, pp. 1–10.

Ajith, T. A. (2020). Possible therapeutic effect of magnesium in ocular diseases, *Journal of basic and clinical physiology and pharmacology* **31**(2).

Akinyelu, A. A. and Blignaut, P. (2020). Convolutional neural network-based methods for eye gaze estimation: A survey, *IEEE Access* **8**: 142581–142605.

Armstrong, G. W. and Lorch, A. C. (2020). A (eye): a review of current applications of artificial intelligence and machine learning in ophthalmology, *International ophthalmology clinics* **60**(1): 57–71.

Arrieta Ramos, J. M. (2022). Semi-supervised deep learning for ocular image classification.

Delmas, D., Cornebise, C., Courtaut, F., Xiao, J. and Aires, V. (2021). New highlights of resveratrol: A review of properties against ocular diseases, *International Journal of Molecular Sciences* **22**(3): 1295.

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.

Guarnieri, A., Carnero, E., Bleau, A.-M., Alfonso-Bartolozzi, B. and Moreno-Montañés, J. (2020). Relationship between osdi questionnaire and ocular surface changes in glaucomatous patients, *International Ophthalmology* **40**(3): 741–751.

Haraburda, P. and Dabała, Ł. (2022). Eye diseases classification using deep learning, *International Conference on Image Analysis and Processing*, Springer, pp. 160–172.

He, J., Li, C., Ye, J., Qiao, Y. and Gu, L. (2021). Self-speculation of clinical features based on knowledge distillation for accurate ocular disease classification, *Biomedical Signal Processing and Control* **67**: 102491.

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.

Jiang, H., Yang, K., Gao, M., Zhang, D., Ma, H. and Qian, W. (2019). An interpretable ensemble deep learning model for diabetic retinopathy disease classification, *2019 41st annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, IEEE, pp. 2045–2048.

Kim, T., Oh, J., Kim, N., Cho, S. and Yun, S.-Y. (2021). Comparing kullback-leibler divergence and mean squared error loss in knowledge distillation, *arXiv preprint arXiv:2105.08919* .

Kumar, S., Adarsh, A., Kumar, B. and Singh, A. K. (2020). An automated early diabetic retinopathy detection through improved blood vessel and optic disc segmentation, *Optics & Laser Technology* **121**: 105815.

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.

Mai, S., Li, Q., Zhao, Q. and Gao, M. (2021). Few-shot transfer learning for hereditary retinal diseases recognition, *International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, pp. 97–107.

Milea, D., Najjar, R. P., Jiang, Z., Ting, D., Vasseneix, C., Xu, X., Aghsaei Fard, M., Fonseca, P., Vanikieti, K., Lagrèze, W. A. et al. (2020). Artificial intelligence to detect papilledema from ocular fundus photographs, *New England Journal of Medicine* **382**(18): 1687–1695.

Nazir, T., Nawaz, M., Rashid, J., Mahum, R., Masood, M., Mehmood, A., Ali, F., Kim, J., Kwon, H.-Y. and Hussain, A. (2021). Detection of diabetic eye disease from retinal images using a deep learning based centernet model, *Sensors* **21**(16): 5283.

Ou, X., Gao, L., Quan, X., Zhang, H., Yang, J. and Li, W. (2022). Bfenet: A two-stream interaction cnn method for multi-label ophthalmic diseases classification with bilateral fundus images, *Computer Methods and Programs in Biomedicine* **219**: 106739.

Pahuja, R., Sisodia, U., Tiwari, A., Sharma, S. and Nagrath, P. (2022). A dynamic approach of eye disease classification using deep learning and machine learning model, *Proceedings of Data Analytics and Management*, Springer, pp. 719–736.

Park, D. H., Connor, K. M. and Lambris, J. D. (2019). The challenges and promise of complement therapeutics for ocular diseases, *Frontiers in Immunology* **10**: 1007.

Pervaiz, S., Ul-Qayyum, Z., Bangyal, W. H., Gao, L. and Ahmad, J. (2021). A systematic literature review on particle swarm optimization techniques for medical diseases detection, *Computational and Mathematical Methods in Medicine* **2021**.

Raza, A., Khan, M. U., Saeed, Z., Samer, S., Mobeen, A. and Samer, A. (2021). Classification of eye diseases and detection of cataract using digital fundus imaging (dfi) and inception-v4 deep learning model, *2021 International Conference on Frontiers of Information Technology (FIT)*, IEEE, pp. 137–142.

Sarki, R., Ahmed, K., Wang, H. and Zhang, Y. (2020). Automatic detection of diabetic eye disease through deep learning using fundus images: a survey, *IEEE Access* **8**: 151133–151149.

Sarki, R., Ahmed, K., Wang, H., Zhang, Y. and Wang, K. (2022). Convolutional neural network for multi-class classification of diabetic eye disease, *EAI Endorsed Transactions on Scalable Information Systems* pp. e15–e15.

Saxena, A., Vyas, A., Parashar, L. and Singh, U. (2020). A glaucoma detection using convolutional neural network, *2020 International Conference on Electronics and Sustainable Communication Systems (ICESC)*, IEEE, pp. 815–820.

Spinozzi, E., Baldassarri, C., Acquaticci, L., Del Bello, F., Grifantini, M., Cappellacci, L. and Riccardo, P. (2021). Adenosine receptors as promising targets for the management of ocular diseases, *Medicinal Chemistry Research* **30**(2): 353–370.

Tong, Y., Lu, W., Yu, Y. and Shen, Y. (2020). Application of machine learning in ophthalmic imaging modalities, *Eye and Vision* **7**(1): 1–15.

Ulhaq, Z. S., Soraya, G. V., Wulandari, L. R. et al. (2020). The role of il-6-174 g/c polymorphism and intraocular il-6 levels in the pathogenesis of ocular diseases: a systematic review and meta-analysis, *Scientific reports* **10**(1): 1–20.

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.

Wang, L. and Yoon, K.-J. (2021). Knowledge distillation and student-teacher learning for visual intelligence: A review and new outlooks, *IEEE Transactions on Pattern Analysis and Machine Intelligence* .

Yang, J.-J., Li, J., Shen, R., Zeng, Y., He, J., Bi, J., Li, Y., Zhang, Q., Peng, L. and Wang, Q. (2016). Exploiting ensemble learning for automatic cataract detection and grading, *Computer methods and programs in biomedicine* **124**: 45–57.