T.C

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DETECTION OF RETINOPATHY DISEASES USING CONVOLUTIONAL NEURAL NETWORK BASED ON DISCRETE COSINE TRANSFORM

MASTER'S THESIS

Mouad KABBOURI

Department of Software Engineering

Artificial Intelligence and Data Science Program

SEPTEMBER, 2023

T.C

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DETECTION OF RETINOPATHY DISEASES USING CONVOLUTIONAL NEURAL NETWORK BASED ON DISCRETE COSINE TRANSFORM

MASTER'S THESIS

Mouad KABBOURI

(Y2013.140014)

Department of Software Engineering

Artificial Intelligence and Data Science Program

Thesis Advisor: Prof. Dr. Ali OKATAN

SEPTEMBER, 2023

ONAY FORMU

DECLARATION

I hereby declare with respect that the study “Detection of Retinopathy Diseases Using Convolutional Neural Network Based on Discrete Cosine Transform”, which I submitted as a Master / PhD thesis, is written without any assistance in violation of scientific ethics and traditions in all the processes from the Project phase to the conclusion of the thesis and that the works I have benefited are from those shown in the References. (30/09/2023)

Mouad KABBOURI

FOREWORD

"First and foremost, I extend my heartfelt gratitude to God for guiding me through this journey and providing me with the strength and perseverance to complete this thesis. I would also like to express my deepest appreciation to my family for their unwavering support and encouragement throughout my academic endeavors.

I am incredibly grateful to have had the opportunity to work with Dr. Ali OKATAN as my supervisor, who guided me through the research process with patience and expertise.

I would like to extend my sincere thanks to Istanbul Aydin University for providing me with the opportunity to pursue my master's degree and for creating an environment that has allowed me to meet and be inspired by some of the most talented and motivated individuals in my field.

I extend my profound gratitude to all those who have helped and supported me throughout the research and writing of this thesis."

September, 2023 Mouad KABBOURI

DETECTION OF RETINOPATHY DISEASES USING CONVOLUTIONAL NEURAL NETWORK BASED ON DISCRETE COSINE TRANSFORM

ABSTRACT

The aim of this study was to examine the relationship between self-esteem and basic psychological needs satisfaction in the adolescents from diffrent kind of high shcool. The universe of the research consists of students who have completed secondary education in Istanbul in 2016-2017 academic year. The sample of the research consists of 299 students selected from 4 secondary schools in Istanbul with random sampling. 64.9% (194) of the students are girls and 35.1% (105) are boys. 8.4% (25) of the students are 15 years old, 46.2% (138) are 16 years old, 43.5% (130) are 17 years old, and 2% (6) are 18 years old. According to school distributions; The rate of students with Multi-Program Anatolian High School is 25.1% (75), the rate of Social Science High School students is 25.4% (76), the rate of Science High School students is 24.1% (72) and the rate of Anatolian High School students is 25.4% (76) 'kind.

The data of the study were collected with the Rosenberg Self-Esteem Scale short form and Basic Psychological needs scale. The data obtained were analyzed in SPSS. Descriptive statistics: frequency tables consist of avarage, standart deviation information.

As a result of the study, it was found that autonomy, competence and being related sub-dimensions of basic psychological needs explained 46.7% of the variance of self-esteem. The findings are discussed in the light of previous research.

Keywords: Competence, Autonomy, Relatedness and Self-Respect

AYRIK KOSİNÜS DÖNÜŞÜMÜNE DAYALI KONVOLÜSYONEL SİNİR AĞI KULLANILARAK RETİNOPATİ HASTALIKLARININ TESPİTİ

ÖZET

Bu yüksek lisans tezi, ayrık kosinüs dönüşümüne (DCT) dayalı bir evrişimli sinir ağı (CNN) kullanarak retinopati hastalıklarını saptamak için yeni bir yaklaşım önermektedir.

Retinopati, erken teşhis ve tedavi edilmezse görme kaybına neden olabilen yaygın bir göz hastalığıdır. Önerilen yöntem, algılama doğruluğunu artırmak için CNN ve DCT'nin gücünü birleştirir. Giriş görüntüsü, gürültü miktarını azaltan ve görüntünün önemli özelliklerini vurgulayan DCT kullanılarak frekans alanına dönüştürülür. Ardından, dönüştürülen görüntü sınıflandırma için CNN'ye beslenir. Önerilen yöntemin performansı, halka açık bir retinal görüntü veri kümesi kullanılarak değerlendirilir.

Sonuçlar, önerilen yöntemin doğruluk ve hesaplama verimliliği açısından mevcut yöntemlerden daha iyi performans gösterdiğini göstermektedir. Önerilen yöntem, retinopati hastalıklarının erken tanı ve tedavisi için gerçek dünyadaki uygulamalarda kullanılma potansiyeline sahiptir.

Anahtar Kelimeler: retinopati hastalıkları, konvolüsyonel sinir ağı, ayrık kosinüs dönüşümü, erken tanı, tedavi

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

I. INTRODUCTION

A. Introduction

Retinopathy refers to a group of eye diseases that primarily affect the retina, a light-sensitive tissue situated Toward the rear of the eye, in charge of converting light into neural signals. The retina plays an essential role in proper vision functioning. Retinopathy diseases, including diabetic retinopathy, retinopathy of prematurity, and hypertensive retinopathy, can cause progressive damage if not recognized and treated immediately, it can cause vision loss or possibly blindness. Early detection and intervention are critical in mitigating the impact of retinopathy and preserving vision for patients.

The traditional diagnostic methods for retinopathy involve various techniques such as fundus examination, optical coherence tomography (OCT), and fluorescein angiography. Fundus examination involves using an ophthalmoscope to evaluate the retina's health, while OCT is a non-invasive imaging technique that captures high-resolution cross-sectional images of the retina. Fluorescein angiography involves injecting a dye into the bloodstream and using a specialized camera to photograph the blood vessels in the retina. These methods necessitate specialized equipment, skilled medical professionals, and can be time-consuming and expensive. Furthermore, the interpretation of the test results may be subjective, potentially leading to inconsistencies in diagnosis and delayed treatment.

The global prevalence of diabetes has been rising at an alarming rate, increasing the number of diabetic retinopathy cases. This rise emphasizes the need for more efficient, accurate, and accessible diagnostic tools to meet the growing demand for retinopathy screening and diagnosis. Deep learning and artificial intelligence (AI) developments have showed promise in overcoming these difficulties, especially in the realm of medical imaging.

Convolutional Neural Networks (CNNs) have emerged as a powerful tool in the field of medical imaging, particularly for the detection and classification of various diseases. CNNs are a type of deep learning model that mimics the human visual system, enabling the efficient processing of images for pattern recognition and feature extraction. Researchers have been employing CNNs for retinopathy detection, achieving significant progress. Some examples of CNN architectures used in retinopathy detection include AlexNet, VGG-16, and Inception. These models have demonstrated high accuracy and efficiency in detecting retinopathy lesions, such as microaneurysms, hemorrhages, and exudates.

However, there is still room for improvement in terms of accuracy, efficiency, and robustness of the models. One challenge faced by CNNs is the large amount of data required for training. Insufficient data may lead to overfitting, resulting in a model that performs poorly on new, unseen data. Data augmentation techniques, such as rotation, flipping, and scaling, are often employed to increase the dataset's size, but these methods may not always be sufficient to prevent overfitting. Furthermore, the high computational cost of training and deploying CNNs might be a barrier, particularly in low-resource situations.

Discrete Cosine Transform (DCT) is an established technique in image processing, widely used for image compression, enhancement, and feature extraction. DCT is a linear transformation that converts an image from the spatial domain to the frequency domain, representing the image's information using a series of cosine functions with varying frequencies. By analyzing the image in the frequency domain, DCT can help identify and extract relevant features while reducing the dimensionality of the data. Incorporating DCT into the CNN model for retinopathy detection may provide better feature representation, enhancing the model's performance and leading to more accurate and reliable diagnosis.

This literature review will delve into the various aspects of retinopathy diseases, their types, symptoms, risk factors, and the challenges associated with diagnosis. It will also explore the advancements in CNNs and their application in medical imaging, along with the role of DCT in image processing and feature extraction.

The next step in the study will involve the collection and preprocessing of fundus images for model development. The dataset will include fundus images from publicly available sources, such as Aptos 2019 blindness detection, which contain images with varying stages and types of retinopathy. The preprocessing step will involve techniques such as contrast enhancement, noise reduction, and vessel segmentation to ensure high-quality input for the model. Additionally, data augmentation techniques, such as rotation, flipping, and scaling, will be employed to increase the dataset size and promote model generalization.

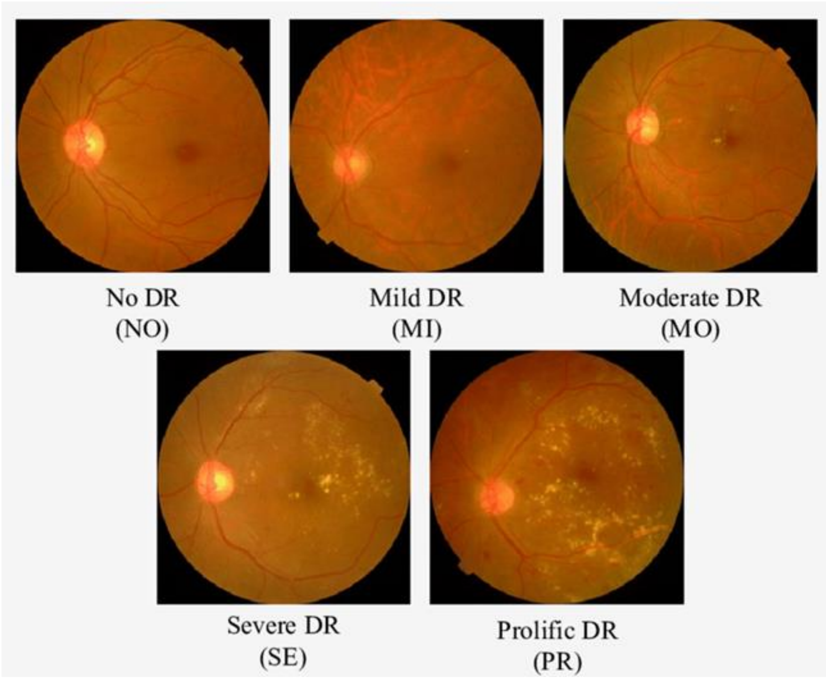


Figure 1: Sample images of the five DR classes collected from the APTOS 2019 dataset

The study will then focus on developing a CNN-based model for retinopathy detection that incorporates DCT for feature extraction. The development process will include the following steps:

Implementing DCT-based feature extraction to obtain more robust and discriminative features from the fundus images. This step will involve applying DCT to the preprocessed images, retaining a select number of DCT coefficients that capture the most relevant information, and reconstructing the images in the spatial domain. The resulting images will serve as input for the CNN model.

Designing, implementing, and optimizing the CNN model using the extracted features. The model architecture will be designed based on existing successful CNN architectures, such as Xception, and Resnet152v2, with modifications tailored to retinopathy detection. The model will be trained using a portion of the dataset, with the remaining data reserved for validation and testing.

To get optimal performance, fine-tune model Learning rate, also batch size, and number of convolutional layers are examples of hyperparameters. This step will involve using techniques such as grid search and cross-validation to systematically explore the hyperparameter space and identify the best combination of parameters.

This research is dedicated to evaluating the efficacy of the suggested model. Upon its completion and optimization, the model will be thoroughly examined in terms of various metrics. These include accuracy, sensitivity, specificity, F1-score, and the area beneath the receiver operating characteristic curve, utilizing a meticulously curated test dataset. By delving into these metrics, we aim to provide a comprehensive understanding of the model's capabilities. Furthermore, this in-depth evaluation will not only highlight the potential advantages for its implementation in clinical settings but will also offer insights into its effectiveness and precision when tasked with the identification of retinal diseases. This holistic approach ensures that the model's robustness and reliability are well-understood before any practical applications.

Moreover, the study will compare the performance of the proposed model with existing methods, including other CNN-based models and traditional diagnostic techniques, such as fundus examination. This comparison will help identify potential improvements and areas for future research.

Finally, the study will discuss the implications of the findings for medical professionals, patients, and the broader healthcare system. By developing a more accurate, efficient, and robust model for retinopathy detection, the study aims to contribute to the early diagnosis and treatment of retinopathy, ultimately helping to preserve vision for those affected by these diseases. The study will also explore the potential applications of the proposed model beyond retinopathy detection, such as the diagnosis of other eye diseases or medical imaging tasks.

In conclusion, This research seeks to establish a new model rooted in CNN techniques for identifying retinopathy that incorporates DCT for feature extraction, offering a more accurate and efficient diagnostic tool for retinopathy diseases. The outcomes of this research might have a substantial influence on the early detection and treatment of retinopathy. ultimately helping to preserve vision for those affected by these diseases. Moreover, the study will provide valuable insights into the application of CNNs and DCT in medical imaging, paving the way for further research and innovation in this area.

B. Background And Motivation

Retinopathy encompasses a group of eye diseases that primarily affect the retina, the light-sensitive tissue responsible for converting light into neural signals at the back of the eye. The retina is crucial for proper vision functioning, and damage to The retina can cause visual loss. or even blindness. Retinopathy diseases, including diabetic retinopathy, retinopathy of prematurity, and hypertensive retinopathy, are characterized by progressive damage to the retinal blood vessels, which can cause a variety of symptoms, such as blurred vision, floaters, and dark spots in the visual field.

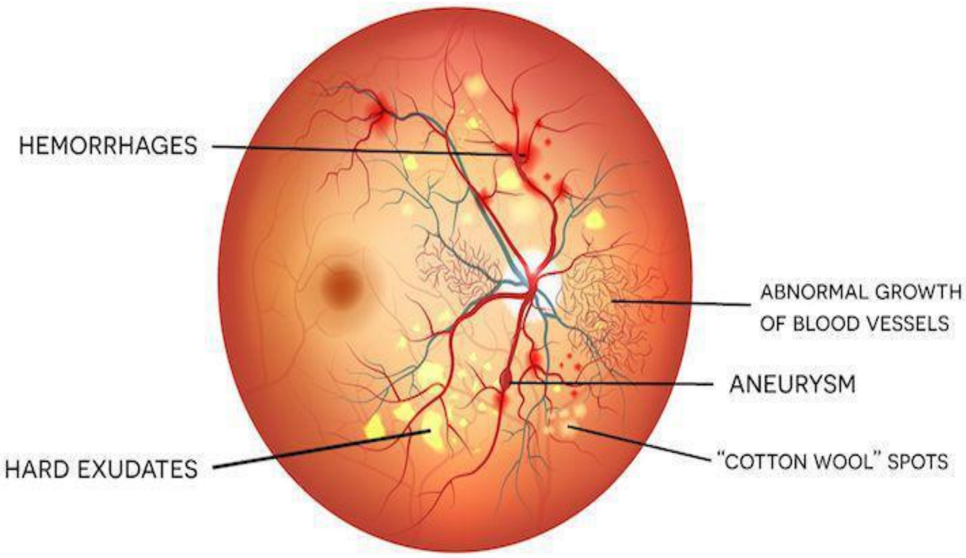


Figure 2: Sample images of the five DR classes collected from the APTOS 2019 dataset

Diabetic retinopathy, in particular, is a growing concern due to the global increase in diabetes cases. It is a major cause of vision loss among working-age adults and poses a significant burden on healthcare systems worldwide. Early detection and intervention are critical in mitigating the impact of retinopathy and preserving vision for patients. However, the traditional diagnostic methods, such as fundus examination, optical coherence tomography (OCT), and fluorescein angiography, have several limitations, including high costs, the need for specialized equipment and skilled medical professionals, and subjective interpretation of test results. These factors can lead to inconsistencies in diagnosis and delayed treatment.

The advances in artificial intelligence (AI) and deep learning techniques have shown promise in addressing the challenges associated with retinopathy diagnosis.

Convolutional Neural Networks (CNNs), a type of deep learning model that mimics the human visual system, have emerged as a powerful tool in medical imaging for the detection and classification of various diseases. Researchers have been employing CNNs for retinopathy detection, achieving significant progress. However, there are still areas for improvement in terms of accuracy, efficiency, and robustness of the models. The limitations of CNNs include the large amount of data required for training, the high computational cost, and the susceptibility to overfitting.

Discrete Cosine Transform (DCT), an established technique in image processing, has been widely used for image compression, enhancement, and feature extraction. DCT is a linear transformation that converts an image from the spatial domain to the frequency domain, representing the image's information using a series of cosine functions with varying frequencies. By analyzing the image in the frequency domain, DCT can help identify and extract relevant features while reducing the dimensionality of the data. Incorporating DCT into the CNN model for retinopathy detection may provide better feature representation, enhancing the model's performance and leading to more accurate and reliable diagnosis.

The motivation behind this study is to address the challenges associated with retinopathy detection and improve the early diagnosis and treatment of these diseases. By developing a novel CNN-based model that incorporates DCT for feature extraction, the study aims to provide a more accurate, efficient, and robust diagnostic tool for retinopathy diseases. The suggested approach has the potential to assist medical professionals in early detection and intervention, eventually assisting in the preservation of eyesight for patients suffering from retinopathy. Furthermore, the study will contribute to the growing body of research on the application of AI and deep learning techniques in medical imaging, paving the way for further innovation in this area.

The increasing prevalence of retinopathy diseases, particularly diabetic retinopathy, underscores the urgent need for more effective and accessible diagnostic tools. The limitations of traditional diagnostic methods, such as fundus examination, optical coherence tomography (OCT), and fluorescein angiography, highlight the potential benefits of leveraging advanced AI and deep learning techniques in retinopathy detection.

Convolutional neural networks (CNNs) have already shown amazing outcomes in a number of fields when used in medical imaging., including cancer detection, lesion segmentation, and organ identification. CNNs offer the ability to process complex images, identify patterns, and extract features, making them a promising tool for retinopathy detection. However, despite the achievements, there remain challenges and areas for improvement in the implementation of CNNs for retinopathy detection, such as the need for large training datasets, high computational costs, and model overfitting.

Incorporating Discrete Cosine Transform (DCT) into the CNN model for retinopathy detection has the potential to address some of these challenges. DCT can improve feature representation by converting pictures from the spatial to the frequency domains, enabling the extraction of more discriminative features and the reduction of data dimensionality. By combining the strengths of CNNs and DCT, the proposed model aims to achieve higher accuracy, efficiency, and robustness in retinopathy detection.

Moreover, the integration of DCT with CNNs has broader implications for the field of medical imaging. By demonstrating the effectiveness of DCT-based feature extraction in improving CNN performance, this study could inspire further research and innovation in other medical imaging applications. For instance, the proposed approach could be extended to the detection of other eye diseases, such as age-related macular degeneration and glaucoma, or applied to different imaging modalities, such as magnetic resonance imaging (MRI) and computed tomography (CT).

Another motivation for this study is the potential for enhancing the accessibility of retinopathy diagnosis. Developing a more accurate, efficient, and robust CNN-based model for retinopathy detection could help overcome some of the barriers associated with traditional diagnostic methods, such as the need for specialized equipment and skilled medical professionals. By simplifying the diagnostic process and reducing the associated costs, the proposed model could facilitate the widespread adoption of retinopathy screening and improve the availability of early diagnosis and treatment for those affected by these diseases.

Finally, the study aims to contribute to the growing body of research on the application of AI and deep learning techniques in healthcare. By developing and evaluating a novel CNN-based model for retinopathy detection that incorporates DCT for feature extraction, this study will add valuable insights to the understanding of the strengths and limitations of AI-based medical imaging approaches. These insights could inform future research, policy-making, and clinical practice, ultimately improving patient care and outcomes for those affected by retinopathy and other vision-threatening diseases.

C. Objectives Of The Study

The primary goal of this study is to develop a novel CNN-based model for retinopathy detection that incorporates DCT for feature extraction, thereby offering a more accurate, efficient, and robust diagnostic tool for retinopathy diseases. To achieve this goal, the study will pursue the following detailed objectives:

Conduct an extensive literature review to understand the present situation of retinopathy detection methods, the application of CNNs in medical imaging, and the role of DCT in image processing. This review will provide the foundation for the development of the proposed model and identify potential areas for improvement in existing approaches.

Collect and preprocess a dataset of fundus images for model development. The dataset will include images with varying stages and types of retinopathies, sourced from publicly available databases such as APTOS 2019 BLINDNESS. The preprocessing step will involve techniques such as contrast enhancement, noise reduction, and vessel segmentation to ensure high-quality input for the model. Data augmentation techniques will also be employed to increase the dataset size and promote model generalization.

Implement DCT-based feature extraction to obtain more robust and discriminative features from the fundus images. This step will involve applying DCT to the preprocessed images, retaining a select number of DCT coefficients that capture the most relevant information, and reconstructing the images in the spatial domain. The resulting images will serve as input for the CNN model.

Design, implement, and optimize the CNN model using the extracted features. The model architecture will be designed based on existing successful CNN architectures, such as Xception, Resnet152v2, with modifications tailored to retinopathy detection. The model will be trained using a portion of the dataset, with the remaining data reserved for validation and testing.

Fine-tune the model hyperparameters, such as learning rate, batch size, and the number of convolutional layers, to achieve optimal performance. This step will involve using techniques such as grid search and cross-validation to systematically explore the hyperparameter space and identify the best combination of parameters.

To assess the efficacy of the proposed model, we will utilize a carefully curated test dataset and analyze its outcomes based on several crucial performance metrics. These metrics include:

- Accuracy: A measure of the model's overall ability to correctly classify both the presence and absence of retinopathy diseases.
- Sensitivity (or Recall): This will shed light on the model's capability to rightly pinpoint true positive cases, crucial for early detection and treatment of retinopathy.
- Specificity: This metric will highlight the model's precision in correctly identifying true negative cases, ensuring patients without the disease aren't mistakenly diagnosed.
- F1-score: Serving as the harmonic mean of precision and recall, this score will encapsulate a balanced view of the model's performance.
- Area Under the ROC Curve (AUC-ROC): Representing the model's true positive rate against its false positive rate, the AUC provides a comprehensive metric for the model's diagnostic ability across varying thresholds.
- By examining the model through these lenses, we aim to delve into its competence in recognizing retinopathy diseases and to determine its suitability for implementation in a clinical environment.

Compare the performance of the proposed model with existing methods, including other CNN-based models and traditional diagnostic techniques, such as fundus examination, OCT, and fluorescein angiography. This comparison will help identify potential improvements and areas for future research.

Discuss the implications of the findings for medical professionals, patients, and the broader healthcare system. By developing a more accurate, efficient, and robust model for retinopathy detection, the study aims to contribute to the early diagnosis and treatment of retinopathy, ultimately helping to preserve vision for those affected by these diseases. The study will also explore the potential applications of the proposed model beyond retinopathy detection, such as the diagnosis of other eye diseases or medical imaging tasks.

Investigate the potential for transfer learning in the proposed model by exploring the effectiveness of using pre-trained CNN models as a starting point. This objective will assess whether leveraging existing CNN models trained on large-scale image datasets, such as ImageNet, can further improve the performance of the retinopathy detection model by providing an effective feature extraction base.

Assess the impact of different DCT coefficient selection strategies on the performance of the proposed model. This objective will explore various methods for retaining the most relevant DCT coefficients and examine their influence on the model's accuracy, efficiency, and robustness. The analysis will help identify the optimal DCT coefficient selection approach for retinopathy detection.

Evaluate the robustness of the proposed model against various image artifacts and distortions, such as blur, noise, and uneven illumination. This objective will test the model's performance under challenging conditions and help identify potential areas for improvement to ensure reliable retinopathy detection in real-world clinical settings.

Investigate the interpretability of the proposed model by examining the features and patterns learned by the CNN during training. This objective will provide insights into the model's decision-making process, which is crucial for building trust and acceptance among medical professionals and facilitating the adoption of AI-based diagnostic tools in clinical practice.

Explore the potential for integrating the proposed model into a telemedicine platform to enhance the accessibility of retinopathy screening and diagnosis. This objective will assess the feasibility of deploying the model in remote or resource-limited settings, enabling more patients to access early diagnosis and treatment for retinopathy diseases.

Conduct a cost-benefit analysis of the proposed model compared to traditional diagnostic methods, such as fundus examination, OCT, and fluorescein angiography. This objective will provide an understanding of the potential economic benefits associated with adopting the proposed model in clinical settings, considering factors such as equipment costs, personnel training, and time efficiency.

Assess the scalability of the proposed model to handle large-scale screening programs and multi-center studies. This objective will evaluate the model's capacity to process and analyze large volumes of fundus images efficiently and accurately, which is essential for its successful implementation in large-scale retinopathy screening initiatives.

In summary, the study aims To build and assess a novel CNN-based model for retinopathy detection that incorporates DCT for feature extraction. The detailed objectives outlined above will guide the research process, ensuring a comprehensive and rigorous approach to addressing the challenges associated with retinopathy diagnosis and improving patient outcomes.

D. Scope And Limitations

The scope and limitations of this study are outlined below to provide a clear understanding of the research boundaries and the potential constraints that may impact the study's findings.

1. Scope

The study focuses on retinopathy detection, specifically targeting diseases such as diabetic retinopathy, retinopathy of prematurity, and hypertensive retinopathy. The proposed model will be designed to identify and classify different stages and types of retinopathies based on fundus images.

The research will utilize publicly available retinopathy dataset, such as APTOS 2019 Blindness Detection dataset. The dataset will include images with varying degrees of retinopathy severity and different demographic profiles to ensure the model's generalizability.

The study will investigate the integration of DCT-based feature extraction into the CNN model to improve retinopathy detection performance. The focus will be on evaluating different DCT coefficient selection strategies and their impact on the model's accuracy, efficiency, and robustness.

The proposed model's performance will be evaluated against existing retinopathy detection methods, including other CNN-based models and traditional diagnostic techniques such as fundus examination, OCT, and fluorescein angiography. This comparison will provide insights into the model's effectiveness and potential advantages over existing approaches.

The research will explore the potential applications of the proposed model beyond retinopathy detection, such as the diagnosis of other eye diseases or medical imaging tasks. The study will also assess the feasibility of integrating the model into telemedicine platforms to enhance the accessibility of retinopathy screening and diagnosis.

The study will explore the potential benefits of transfer learning in the proposed model by examining the impact of using pre-trained CNN models, such as those trained on ImageNet, as a starting point for feature extraction. This investigation will assess whether transfer learning can further improve the model's performance in retinopathy detection.

The research will also address the scalability of the proposed model, evaluating its ability to handle large volumes of fundus images efficiently and accurately. This assessment will be crucial for the successful implementation of the model in large-scale retinopathy screening initiatives and multi-center studies.

The study will consider the practical aspects of implementing the proposed model in real-world clinical settings, such as the ease of integration with existing diagnostic workflows, the impact on medical professionals' workload, and the cost-effectiveness compared to traditional diagnostic methods.

The research will also address the robustness of the proposed model against various image artifacts and distortions, such as blur, noise, and uneven illumination. This evaluation will give details on the model's capacity to perform reliably under challenging conditions, which is essential for its practical application in clinical settings.

2. Limitations

The study's findings will be dependent on the quality and representativeness of the fundus image datasets used for model development and evaluation. Potential limitations in the datasets, such as the presence of artifacts or a lack of diversity in patient demographics, may impact the model's performance and generalizability.

The proposed model's performance may be influenced by various factors, such as the choice of CNN architecture, hyperparameter settings, and DCT coefficient selection strategies. While the study will explore these factors systematically, the findings may not necessarily represent the optimal solution for retinopathy detection.

The research will primarily focus on fundus images for retinopathy detection. While this imaging modality is widely used in clinical practice, the study's findings may not be directly applicable to other imaging techniques.

The study will not address potential ethical, legal, or regulatory the difficulties connected with the adoption of AI-based diagnostic technologies in healthcare, such as data privacy, informed consent, or liability. These aspects are important considerations for the practical implementation of the proposed model but are beyond the scope of this research.

The proposed model's interpretability and explain ability may be limited, given the complexity of deep learning models and the inherent difficulty in understanding their decision-making processes. While the study will explore the features and patterns learned by the CNN, the findings may not provide a complete understanding of the model's rationale for retinopathy detection.

The computational resources required for the development and optimization of the proposed model may be substantial, particularly when training the CNN on large datasets and exploring the hyperparameter space. While the study will attempt to optimize computational efficiency, the availability of resources may impose limitations on the model's performance and the extent of hyperparameter tuning.

The study's focus on retinopathy detection using fundus images may limit the applicability of the findings to other imaging modalities or other eye diseases. However, the proposed model's underlying concepts, such as the integration of DCT-based feature extraction with CNNs, may still be relevant and valuable for future research in these areas.

The proposed model's clinical adoption may be influenced by factors beyond its performance, such as the acceptance and trust of medical professionals, the availability of necessary infrastructure, and the legal and regulatory environment. While the study will attempt to address some of these concerns, it is important to acknowledge that the successful implementation of the model in clinical practice will require a multi-faceted approach involving various stakeholders.

The study will not investigate the potential impact of the proposed model on patient outcomes, such as the timely initiation of treatment or the prevention of vision loss. While the ultimate goal of the research is to contribute to the early diagnosis and treatment of retinopathy diseases, the direct assessment of patient outcomes is beyond the scope of this study.

By acknowledging the scope and limitations of this study, the research aims to provide a comprehensive and rigorous analysis of the proposed CNN-based model for retinopathy detection that incorporates DCT for feature extraction while being transparent about the potential constraints that may impact the findings.

II. LITERATURE REVIEW

This chapter presents an in-depth review of the pertinent literature surrounding retinopathy diseases, existing diagnostic methodologies, the application of Convolutional Neural Networks (CNNs) in medical imaging, the role of Discrete Cosine Transform (DCT) in image processing, and prior research on retinopathy detection employing both CNNs and DCT. By delving into these areas, the review aims to furnish the essential foundation and context required for developing the proposed retinopathy detection model. The literature review will be structured as follows:

A comprehensive examination of retinopathy diseases, focusing on their various types, symptoms, and risk factors. This section will elucidate the underlying biological mechanisms and manifestations of retinopathy diseases, highlighting their significance in the field of ophthalmology and their impact on patients' quality of life. A thorough analysis of the current diagnostic methods used in retinopathy detection, encompassing both traditional techniques and novel approaches. This part of the review will assess the strengths and weaknesses of each method, as well as the challenges faced by medical professionals in accurately diagnosing retinopathy diseases.

An exploration of the role of Convolutional Neural Networks (CNNs) in medical imaging, including their architecture, training processes, and various applications. This section will discuss the advantages of CNNs in image analysis tasks and their potential to revolutionize the field of medical imaging, particularly in the context of retinopathy detection.

An in-depth discussion of Discrete Cosine Transform (DCT) and its applications in image processing, with a focus on its use in feature extraction, compression, and noise reduction. This part of the review will provide an understanding of the mathematical principles underlying DCT and its effectiveness in representing images using a compact set of coefficients.

A critical evaluation of the existing research on retinopathy detection using CNNs and DCT, identifying the key advancements, challenges, and opportunities in this domain. This section will synthesize the current state of knowledge on the integration of CNNs and DCT in retinopathy detection, highlighting the potential benefits of combining these techniques to develop a more accurate, efficient, and robust diagnostic model.

By providing a well-rounded and detailed examination of the relevant literature, this chapter will establish a solid foundation for the proposed study on retinopathy detection using a CNN-based model that incorporates DCT for feature extraction. The insights gained from this review will inform the research design, methodology, and analysis, Contributing to the creation of a more effective and trustworthy diagnostic tool for retinal illnesses in the future.

A. Retinopathy diseases: types, symptoms, and risk factors

Retinopathies, a group of disorders impacting the retina - a light-sensitive tissue integral to our vision - is the central focus of my thesis. The ability of these disorders to drastically impair vision, even leading to blindness, underlines the significance of understanding their various forms, symptoms, related risk factors, and particularly their trajectory from a normal stage, through mild, moderate, severe stages, culminating in the proliferative stage.

The first type I'm examining is Diabetic retinopathy, which is primarily a complication of diabetes and is a leading cause of blindness among adults. It results from sustained high blood sugar levels damaging the blood vessels in the retina. The disease presents in two main stages: non-proliferative diabetic retinopathy (NPDR), which comprises the mild to severe stages, and proliferative diabetic retinopathy (PDR), representing the most severe stage. The typical symptoms, such as blurred vision, floaters, and dark spots, along with difficulties seeing in low light, can escalate over time. Risk factors include the duration of diabetes, inadequate glycemic control, hypertension, elevated cholesterol levels, and kidney disease.

Secondly, I've looked into Retinopathy of Prematurity (ROP). This condition primarily affects premature infants, especially those with low birth weight and gestational age. The premature and abnormal growth of retinal blood vessels can lead to retinal scarring, detachment, and eventually, vision loss. The progression of ROP spans five stages, from mild to severe. Some cases of ROP may resolve independently; however, more severe cases require intervention such as laser therapy, cryotherapy, or vitrectomy. The main risk factors are premature birth, low birth weight, the necessity for oxygen therapy, and other systemic health complications.

Lastly, I've delved into Hypertensive retinopathy. Chronic high blood pressure can lead to damage to the retina's blood vessels, causing hypertensive retinopathy. Initial damage can present as mild swelling and hemorrhages. As the condition advances, cotton wool spots, or areas of localized retinal ischemia, emerge. In severe cases, hypertensive choroidopathy and hypertensive optic neuropathy can develop, causing further vision loss. Symptoms can be subtle or non-existent, underscoring the importance of regular eye examinations for early detection. Risk factors include poorly controlled hypertension, duration of hypertension, older age, and other cardiovascular risks.

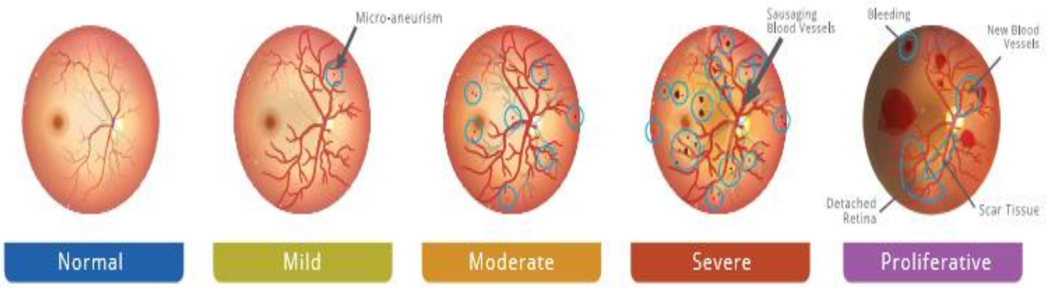


Figure 3: Stages of diabetic retinopathy

Table 1: Severity stages

Stage	Severity	Description
Normal	-	Retina with no apparent signs of retinopathy.
Mild	Mild Retinopathy	Early signs of retinopathy, often no symptoms.
Moderate	Moderate Retinopathy	Further weakening of retinal blood vessels. Some visual symptoms might be observed.
Severe	Severe Retinopathy	Pronounced damage to retinal blood vessels. Symptoms may be more noticeable and impactful.
Proliferative	Proliferative Retinopathy	Growth of new, abnormal blood vessels. Can lead to severe vision loss or blindness due to complications such as retinal detachment.

B. Current Diagnostic Methods For Retinopathy

Accurate and timely diagnosis of retinopathy is crucial for initiating appropriate treatment and managing the progression of the disease. This section provides a detailed overview of the current diagnostic methods used for retinopathy, discussing their principles, advantages, and limitations.

1. Fundus Examination

Fundus examination is a fundamental diagnostic tool for retinopathy assessment. It involves the visual inspection of the retina using specialized equipment such as an ophthalmoscope or a fundus camera. The ophthalmoscope projects a beam of light into the eye, allowing the examiner to observe the retina directly. In contrast, fundus cameras capture digital images of the retina, which can be stored, analyzed, and shared with other healthcare professionals. Fundus examination can reveal signs of retinopathy, such as hemorrhages, microaneurysms, cotton wool spots, and abnormal blood vessel growth. While fundus examination is relatively simple and non-invasive, its accuracy depends on the examiner's skill and experience. Moreover, the quality of the images captured may be affected by factors such as media opacity (e.g., cataracts) and patient cooperation.

2. Optical coherence tomography (OCT)

OCT is a non-invasive imaging technique that uses low-coherence light to generate high-resolution, cross-sectional images of the retina. By measuring the echo time delay of the reflected light, OCT can provide detailed information about the retinal layers' structure and thickness. OCT is particularly useful for detecting and monitoring retinal edema, which is a common feature of diabetic retinopathy and other retinopathy diseases. OCT also enables the assessment of other retinal pathologies, such as macular holes, epiretinal membranes, and choroidal neovascularization. Despite its high resolution and ability to detect subtle changes in retinal structure, OCT is limited by its reliance on patient cooperation and its sensitivity to eye movements.

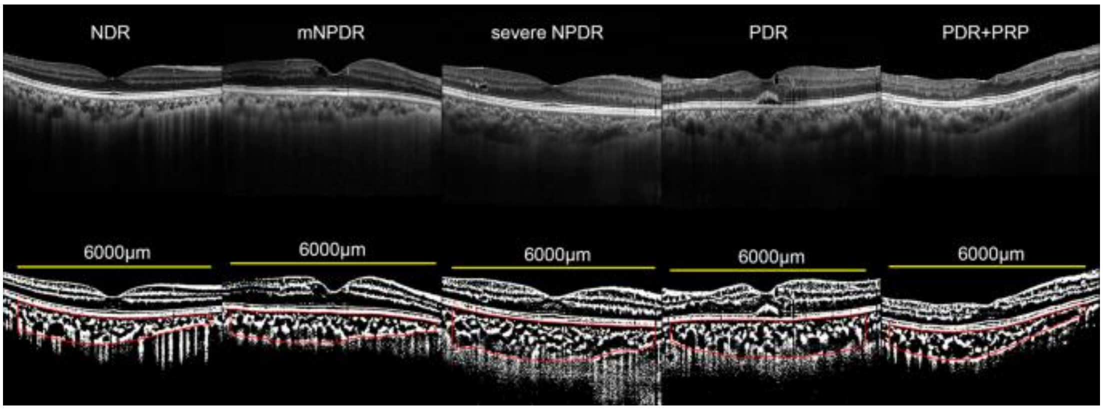


Figure 4: Differences in the choroidal structures at different stages of diabetic retinopathy(https://doi.org/10.1038/s41598-019-52750-0)

3. Fluorescein Angiography

Fluorescein angiography is an invasive diagnostic method that involves the intravenous injection of a fluorescent dye (fluorescein) followed by a series of retinal images captured as the dye circulates through the blood vessels. The images can reveal abnormalities in retinal blood vessels, such as leakage, blockages, and neovascularization, which are indicative of retinopathy. Fluorescein angiography is particularly useful for diagnosing and monitoring proliferative diabetic retinopathy and assessing the need for laser photocoagulation or other treatments. However, this technique carries potential risks and side effects, such as allergic reactions to the dye, nausea, and transient changes in blood pressure. Additionally, the quality of the images may be affected by factors such as patient movement and media opacity.

4. Automated Retinal İmage Analysis

With advances in computer vision and artificial intelligence, automated retinal image analysis has emerged as a promising diagnostic tool for retinopathy. These methods frequently entail the use of machine learning approaches such as Convolutional Neural Networks. to analyze retinal images captured by fundus cameras or other imaging devices. Automated retinal image analysis has the potential to improve the accuracy, efficiency, and accessibility of retinopathy screening, particularly in primary care settings and remote areas with limited access to specialized eye care. However, the performance of these algorithms depends on the quality and representativeness of the training data, and their generalizability to diverse populations and imaging devices remains an area of ongoing research.

In conclusion, current diagnostic methods for retinopathy include fundus examination, optical coherence tomography, fluorescein angiography, and automated retinal image analysis. Each method has its strengths and limitations, and their use in clinical practice often depends on factors such as the specific type and severity.

C. Convolutional Neural Networks (Cnns) İn Medical İmaging

In recent times, Convolutional Neural Networks (CNNs) have garnered considerable interest owing to their outstanding capabilities in a range of image processing activities. including object detection, segmentation, and classification. CNNs have been increasingly applied to medical imaging, where they show great potential to enhance the accuracy and efficiency of illness diagnosis and prognosis. This section will provide a detailed overview of CNNs, their architecture, Machine learning methodologies, such as Convolutional Neural Networks, are commonly used in these procedures.

1. CNN Architecture

A typical CNN architecture consists of several layers, including input, convolutional, activation, pooling, and fully connected layers, which are organized in a hierarchical manner. The input layer takes raw pixel values from an image, while the convolutional layers apply filters to extract local features, such as edges, corners, and textures. Activation layers introduce non-linearity to the network, usually through functions like ReLU (Rectified Linear Unit) or sigmoid. Pooling layers reduce the spatial dimensions of the feature maps, thereby decreasing computational complexity and improving the network's translational invariance. Finally, fully connected layers aggregate and classify the extracted features into distinct categories, such as healthy or diseased tissue.

2. CNN Training

CNNs are trained using a large dataset of labeled images through a process called supervised learning. The training process involves adjusting the weights and biases of the network to minimize the difference between the predicted output and the ground truth labels. This optimization is usually performed using stochastic gradient descent or other variants. During training, the network learns to recognize relevant patterns and features in the images, which can then be used for classification or segmentation tasks.

3. Applications Of Cnns İn Medical İmaging

CNNs have been successfully applied to various medical imaging modalities, such as magnetic resonance imaging (MRI), computed tomography (CT), ultrasound, and optical coherence tomography. Some common applications of CNNs in medical imaging include:

Tumor detection and segmentation: CNNs have been used for the accurate identification and delineation of tumors in different organs, such as the brain, lungs, and breast, improving early diagnosis and treatment planning.

Organ and tissue segmentation: CNNs can automatically segment organs and tissues in medical images, which is essential for tasks like volume estimation, radiation therapy planning, and surgical navigation.

Medical image classification involves determining and assigning labels to medical images from a fixed set. The task involves the extraction of features from the image, and assigning labels using the extracted features.

4. Disease Classification

CNNs have demonstrated high accuracy in classifying various diseases and abnormalities, such as diabetic retinopathy, age-related macular degeneration, and Alzheimer's disease, based on medical images.

5. CNNs İn Retinopathy Detection

In the context of retinopathy detection, several studies have employed CNNs to analyze retinal images for the presence of disease-specific features, such as microaneurysms, hemorrhages, and neovascularization. CNNs have shown promising results in identifying and classifying retinopathy diseases, with some studies achieving performance levels comparable to or even surpassing human experts. CNN-based models can also be integrated into automated retinal image analysis systems, potentially improving the accessibility and efficiency of retinopathy screening and diagnosis.

In summary, Convolutional Neural Networks have emerged as a potent medical imaging tool, demonstrating remarkable performance in tasks such as object detection, segmentation, and classification. Their application to retinopathy detection has shown great promise, with the potential to revolutionize disease diagnosis and prognosis. However, challenges remain, for example, big, diversified, and well-annotated datasets are required for training. as well as ensuring the generalizability and interpretability of the models. Ongoing research seeks to address these challenges and further harness the potential of CNNs in medical imaging and retinopathy detection.

D. Discrete Cosine Transform (DCT) in İmage Processing

Discrete Cosine Transform (DCT) is a widely used technique in image processing, known for its ability to transform a signal or an image from the spatial domain into the frequency domain, thereby facilitating efficient compression, analysis, and reconstruction of the original data. This section provides a detailed overview of DCT, its underlying principles, and its applications in image processing, with a focus on its potential use in retinopathy detection.

1. Description of DCT

DCT is a linear, separable, and orthogonal transform that operates on a block of data (usually of size NxN) and converts it into a set of coefficients representing the frequencies present in the input data. The basis functions of DCT are real-valued cosine functions, which have desirable properties such as energy compaction and decorrelation. These properties enable DCT to efficiently represent the input data with a relatively small number of coefficients, especially when the data exhibits strong spatial correlations, as is common in natural images.

The One-Dimensional DCT:

The common DCT definition of a 1-D sequence (One Dimensional) of length N is:

$$c(u) = \alpha(\mathrm{u}) \sum_{x=0}^{N=1} f(x) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \quad (1)$$

For u = 0,1,2, …., $N = 1$. Similarly, the inverse transformation is defined as:

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \quad (2)$$

For x = 0,1, 2, … , N=1. In both equations (1) and (2) α(u) is defined as:

$$\alpha(\mathrm{u}) = \begin{cases} \sqrt{\frac{1}{N}}, & for\ \ u = 0 \\ \sqrt{\frac{2}{N}}, & for\ u \neq 0 \end{cases} \quad (3)$$

It is clear from (1) that for $u = 0 c(u = 0) = \sqrt{\frac{1}{N}} \sum_{x=0}^{N-1} f(x)$ Thus, the first transform coefficient is the average value of the sample sequence.

2. The Two-Dimensional DCT

The objective of this document is to study the efficacy of DCT on images. This necessitates the extension of ideas presented in the last section to a two-dimensional space. The 2-D DCT is a direct extension of the 1-D case and is:

$$c(u, v) = \alpha(\mathrm{u})\alpha(\mathrm{v}) \sum_{\mathrm{x}=0}^{\mathrm{N}-1} \sum_{y=0}^{N-1} f(x, y) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right] \quad (4)$$

For $u, v = 0,1,2, \ldots, N-1$ and $\alpha(\mathrm{u})$ and $\alpha(\mathrm{v})$ are defined in (3). The inverse transform is defined as:

$$f(x, y) = \sum_{\mathrm{u}=0}^{\mathrm{N}-1} \sum_{v=0}^{N-1} \alpha(\mathrm{u})\alpha(\mathrm{v}) c(u, v) \cos\left[\frac{\pi(2x+1)u}{2N}\right] \cos\left[\frac{\pi(2y+1)v}{2N}\right], \quad (5)$$

For $x, y = 0,1,2, \dots, N-1$. The 2-D basis functions can be generated by multiplying the horizontally oriented 1-D basis functions with vertically oriented set of the same functions (W. B. Pennebaker and J. L. Mitchell, 1993). In the presented context, the foundation functions corresponding to 8N are depicted. A closer observation reveals that these foundation functions display a systematic escalation in frequency, evident in both vertical and horizontal orientations. When examining the top left foundation function, it becomes apparent that its existence is attributed to the multiplication of the DC component, as showcased in Figure 5, with its respective transpose. Owing to this particular interaction, this function maintains a uniform value throughout. Because of its consistent nature and origins from the DC component, it is commonly labeled as the DC coefficient. This coefficient plays a pivotal role in understanding the underlying principles of the model and its interpretation in the broader scheme of the study.

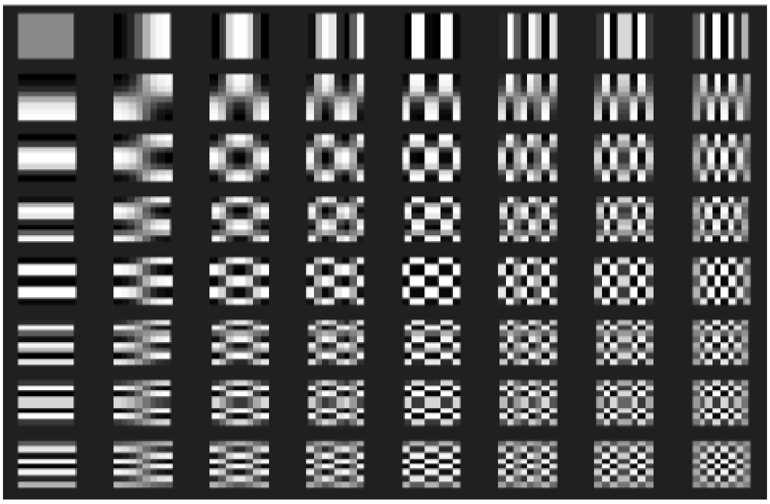


Figure 5: Two dimensional DCT basis functions (N = 8). Neutral gray represents zero, white represents positive amplitudes, and black represents negative amplitude (W. B. Pennebaker and J. L. Mitchell, 1993).

3. DCT İn İmage Compression

One of the most popular applications of DCT is in image compression, where it serves as the core component of widely used standards such as JPEG. In this context, DCT is applied to small, non-overlapping blocks of an image, generating a set of frequency coefficients for each block. Most of the energy in natural images is concentrated in low-frequency components, which can be retained, while high-frequency components can be quantized and compressed to achieve significant data reduction with minimal loss of visual quality. During decompression, the inverse DCT (IDCT) is applied to reconstruct the original image from the compressed data.

4. DCT İn Feature Extraction

DCT can also be used for feature extraction in image analysis tasks, such as pattern recognition, classification, and segmentation. By transforming an image into the frequency domain, DCT can help capture relevant features and patterns that are less sensitive to noise and other artifacts. These features can then be used as inputs to machine learning algorithms, such as support vector machines, decision trees, or deep learning models, for various image analysis tasks. The dimensionality of the extracted features can be reduced by retaining only a subset of the most significant DCT coefficients, thereby reducing computational complexity and improving the efficiency of the subsequent processing steps.

5. DCT İn Retinopathy Detection

In the context of retinopathy detection, DCT can be employed as a pre-processing step to extract relevant features from retinal images, such as texture patterns, local contrast, and edge information. These features can then be fed into machine learning models, such as CNNs, to improve their performance in tasks like retinopathy classification and lesion detection. By incorporating DCT-based features into the analysis pipeline, it is possible to enhance the robustness and efficiency of retinopathy detection algorithms, particularly in situations where the input images are of varying quality, resolution, or contrast.

In summary, Discrete Cosine Transform is a versatile and powerful technique in image processing, with applications ranging from image compression to feature extraction. Its ability to efficiently represent and analyze images in the frequency domain makes it particularly suitable for tasks like retinopathy detection, where the extraction of meaningful features and patterns is critical for accurate and efficient disease diagnosis. By combining DCT with machine learning approaches like CNNs, The quality and accessibility of eye care services can be enhanced by creating improved retinopathy detection technologies that can handle the difficulties provided by various and noisy retinal pictures.

E. Related Works On Retinopathy Detection

Retinopathy detection has been one of the areas in medical diagnostics that have experienced considerable advancements due to the joint use of Convolutional Neural Networks (CNNs) and Discrete Cosine Transform (DCT). These two methodologies, when incorporated together, promise to enhance the accuracy, efficiency, and robustness of diagnostic algorithms.

The utilization of deep learning in the realm of medical imaging offers enhanced diagnostic precision and treatment planning capabilities, as indicated by research from:

Ali and colleagues. Techniques like IMNets are pivotal tools in this advancement. Nevertheless, it's vital to complement machine learning with the expertise of certified healthcare practitioners. IMNets, or Incremental Modular Networks, uniquely build deep learning models by adding modules when required rather than relying on a singular expansive network. This allows the model to specialize in certain patterns crucial for a specific task, instead of trying to understand everything simultaneously.

A hybrid methodology involving deep learning was used to classify Diabetic Retinopathy in a study. By utilizing VGG16 and VGG19 neural networks, images were recognized and their features were noted. The analysis used four severity scales: proliferative, mild, moderate, and severe. Results from this research demonstrated an 85% recall, a 90.4% F1 score, and an overall 90.60% accuracy. Gunasekaran and his team emphasized the potential of retinopathy images in diagnosing diabetes, despite the inherent challenges. A deep recurrent neural network (RNN) was employed in their study, achieving a commendable 95.5% accuracy rate in predicting Diabetic Retinopathy.

Khan and colleagues employed architectures like VGG-net, ResNet, and InceptionV3, leveraging transfer learning. For preprocessing, they applied the Gaussian method to reduce noise, aiming for enhanced outcomes. Their research encompassed five DR variants, with InceptionV3 outperforming other models, exhibiting 81.2% training accuracy and 79.4% testing accuracy.

Fang and team introduced the DAG network model, integrating features from fundus images for DR classification. After extracting three critical features, the DAG network undertook the task of learning and merging these features. The model's efficiency was assessed using data from a hospital and the DIARETDB1 dataset.

Elloumi and associates proposed a method to screen DR fundus images taken with smartphones. Despite the potential quality limitations of these images, the NasnetMobile technique was employed for feature extraction. The research used a dataset of 440 images, and the model demonstrated impressive results across various metrics.

Kanakaprabha and colleagues evaluated various deep learning algorithms for DR prediction, including CNN, VGG16, VGG19, and others. Sridhar recommended a CNN-based model for DR detection, trained on a public dataset from Kaggle, achieving superior accuracy.

Das and his team proposed a system for DR classification based on the characteristics of segmented fundus images. With the DIARETDB1 dataset, their model exhibited 97.2% precision and 98.7% accuracy. Vives-Boix and team implemented convolutional neural network-based meta-plasticity to recognize DR in images, attaining noteworthy results.

Luo and associates addressed the limitations of current DR classification methods by proposing a multi-view fundus image technique combined with convolutional neural networks. The model's performance was compared against existing DR detection systems. Adriman analyzed the efficacy of several deep learning networks for DR detection, using the APTOS 2019 Blindness Detection dataset.

Multiple other studies by researchers like Fatima, Qureshi, Kalyani, Gayathri, Bodapati, Math, and Gao have explored various methodologies, networks, and datasets to improve DR detection and categorization. Kobat and team highlighted the impact of DR on retinal blood vessels. Using a pre-trained DenseNET model, they segmented digitized fundus images, achieving an 84.90% accuracy over multiple tests.

III. METHODOLOGY

This section presents a detailed description of the methodology that will be employed in this study for the development of a Convolutional Neural Network (CNN)-based model to effectively detect retinopathy diseases by leveraging the Discrete Cosine Transform (DCT). The methodology is systematically divided into multiple steps, encompassing data collection and preprocessing, DCT-based feature extraction, design of the CNN architecture, model training and validation, and ultimately, the evaluation of the model's performance.

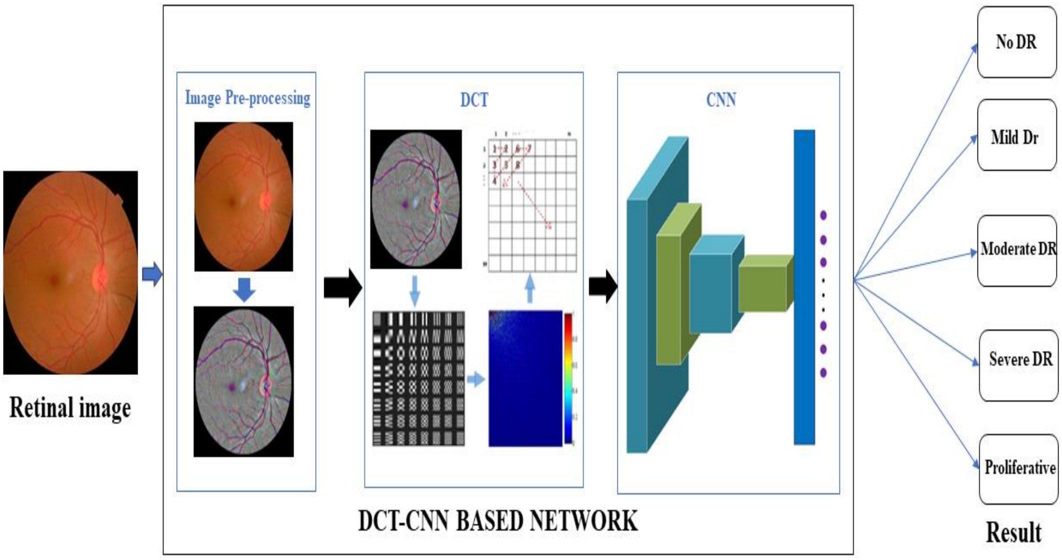


Figure 6: Proposed Diabetic Retinopathy Classification Framework

A. Data Collection and Preprocessing

The initial phase of the methodology necessitates the gathering of a comprehensive dataset comprising retinal images, accompanied by their respective ground truth labels for retinopathy diseases. This dataset can be obtained from publicly accessible sources like APTOS 2019 Blindness Detection dataset and Kaggle. These datasets primarily consist of retinal images captured through fundus photography in 5 categories and are annotated by experts to indicate the presence and severity of retinopathy.

It is critical to preprocess the acquired photos in order to preserve consistency and improve the quality of the supplied data. The following steps may be included in the preprocessing phase:

- Image Resizing: Standardize the dimensions of all images in the dataset by resizing them to a consistent resolution.
- Image Normalization: The photos' pixel values should be adjusted to a consistent range, usually between 0 and 1.
- Image Enhancement: Employ techniques such as contrast stretching, histogram equalization, or adaptive filtering to improve the visibility of retinal structures and lesions.
- Data Augmentation: Generate additional training samples by applying random transformations like rotation, scaling, flipping, or cropping to the original images. This helps enrich the dataset and enhances the model's generalizability.

1. Feature Extraction Using DCT

Once the images have been preprocessed, the DCT will be utilized to extract relevant features from the retinal images. This entails dividing each image into non-overlapping blocks of fixed dimensions (e.g., 8x8 pixels) and calculating the DCT coefficients for every block. The most significant DCT coefficients, representing the low-frequency components of the image, can be preserved as features for subsequent analysis.

2. CNN Architecture Design

The subsequent step involves designing the architecture of the CNN model that will be employed for retinopathy detection. The model may comprise multiple layers, including convolutional, activation, pooling, and fully connected layers. The DCT features extracted in the previous step can be integrated into the CNN architecture in a variety of ways, such as adding them as additional input channels, feeding them to a separate branch of the network, or using them as intermediate targets for auxiliary prediction tasks. The selection of network architecture, layer configurations, and training parameters will be guided by the literature review and informed by empirical experimentation.

3. Training And Validation

Using a supervised learning technique, the CNN model will be trained on the preprocessed retinal visuals and their related ground truth labels. The training process entails iteratively adjusting the weights and biases of the network to minimize the discrepancy between the predicted output and the ground truth labels, typically employing stochastic gradient descent or other optimization algorithms. The model's performance will be monitored using a validation set comprising a subset of images not used for training. By doing this, overfitting is avoided and it is ensured that the model generalizes effectively to new data.

4. Performance Evaluation

Finally, the performance of the CNN model will be assessed using a test set of retinal images not utilized during training or validation. Various metrics can be employed to evaluate the model's performance, such as accuracy, sensitivity, specificity, F1 score, and area under the receiver operating characteristic (ROC) curve. The outcomes of the evaluation will be compared with existing methods for retinopathy detection, as documented in the literature, to ascertain the effectiveness and potential advantages of the proposed approach.

B. Data Collection And Preprocessing

Constructing a proficient Convolutional Neural Network (CNN) model to detect retinopathy diseases via Discrete Cosine Transform (DCT) requires an initial phase of comprehensive data collection and meticulous preprocessing. This crucial phase sets the precedent for subsequent stages as it directly impacts the efficacy of the model by defining the quality and relevance of the gathered data. This section offers an in-depth perspective on the initial steps of data collection and preprocessing, emphasizing the importance of dataset selection and outlining vital preprocessing stages with corresponding justifications.

1. Data Collection

The creation of a robust model capable of learning the complexities of retinopathy diseases necessitates an extensive dataset of retinal images coupled with corresponding accurate labels. This dataset should incorporate a broad range of images, portraying various types and stages of retinopathy diseases. For this project, we used APTOS 2019 Blindness Detection dataset. This dataset is a product of a mission by Aravind Eye Hospital in India to identify and prevent diabetic retinopathy, especially in rural areas where medical screening is challenging. It includes thousands of retinal images acquired from these regions, aiming to expedite disease detection, potentially preventing irreversible blindness. The APTOS 2019 dataset includes two categories: 'train_images' with 3662 images and 'test_images' with 1928 images, all in PNG format as you see in figure bellow.

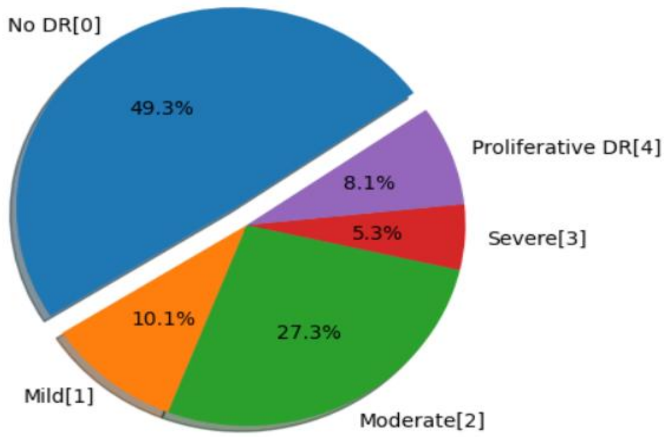


Figure 7: Pie chart analysis of number of images on each target label

2. Data Preprocessing

We have implemented both Convolutional Neural Network based model and DCT model and subsequently tested with both quantized and un-quantized coefficients of PNG compressed data for image classification. The compressed images were partially decoded and the DCT coefficients obtained were fed as input to the neural network after Up-sampling and Downsampling the Y, Cb, Cr components, and concatenated them as one unit. The process of Up-sampling and Down -sampling is further called as transformation. The chosen dataset undergoes preprocessing to enhance data quality and ensure consistency across the dataset. Preprocessing plays an instrumental role in guiding the model to concentrate on crucial features while diminishing the influence of noise and other non-relevant factors. The preprocessing stage includes:

- Image Resizing: The images are resized to standard dimensions of 224x224 pixels, ensuring homogeneity in the input data and simplifying subsequent processing stages.
- Image Enhancement: Techniques such as contrast stretching, histogram equalization, or adaptive filtering are applied to augment the visibility of retinal structures and lesions, aiding the model in learning more effectively.

Additionally, we employ the function crop_image_from_gray() to crop irrelevant or dark regions from the image and the function preprocess_image() to prepare each image for input into the model.

The data collection and preprocessing phase is an indispensable stage of the methodology, setting the foundation for the development of an efficient CNN-based retinopathy detection model. By providing the model with access to a diverse, high-quality set of retinal images from the APTOS 2019 dataset, it can discern the most relevant features for detecting retinopathy diseases, thereby achieving optimal performance.

This dataset represents five classes or levels of diabetic retinopathy severity: 'No_DR', 'Mild', 'Moderate', 'Severe', and 'Proliferate_DR'. Understanding these classes is pivotal for interpreting the model's output, providing vital insights into the severity of diabetic retinopathy, thereby informing necessary medical interventions.

C. Dataset Description And Partitioning

The success of a CNN-based retinopathy detection model hinges upon the quality and diversity of the dataset used for training and evaluation. This section delves into the description of the selected dataset and explains the partitioning strategy employed to divide the dataset into training, validation, and test subsets.

- Dataset Description: The selected dataset should encompass a wide range of retinal images, capturing various types and stages of retinopathy diseases, to ensure a comprehensive learning experience for the model. The retinal images should be acquired through fundus photography, a non-invasive technique that captures high-resolution images of the retina. These images should come with expert annotations, indicating the presence and severity of retinopathy diseases, as well as any specific lesions that might be present. Ideally, the dataset should include a balanced representation of healthy and diseased retinas, while also accounting for potential demographic factors, such as age, gender, and ethnicity, to ensure a robust and generalizable model.

- Partitioning Strategy: After acquiring a suitable dataset and preprocessing the images, the next crucial step is to partition the dataset into three distinct subsets: training, validation, and test sets. This partitioning strategy is crucial to ensure that the model's performance is evaluated objectively and to prevent overfitting. The partitioning can be carried out using different approaches, such as random sampling, stratified sampling, or k-fold cross-validation, depending on the size and characteristics of the dataset.
- Training Set: This subset of the dataset will be used to train the CNN model, allowing it to learn the patterns and features associated with retinopathy diseases. The size of the training set should be large enough to enable the model to learn the intricacies of the problem effectively. Typically, 60-70% of the total dataset is allocated for training.
- Validation Set: The validation set plays a vital role in monitoring the model's performance during the training process. It comprises a subset of images that are not used for training, enabling the assessment of the model's generalization capabilities on unseen data. The validation set helps in fine-tuning the model's architecture and training parameters, ensuring optimal performance. Approximately 10-20% of the total dataset is set aside for validation purposes.
- Test Set: The test set serves as the final evaluation metric for the model's performance. It contains a subset of images not used during training or validation and provides an unbiased assessment of the model's effectiveness in detecting retinopathy diseases on new, unseen data. The test set should ideally represent 10-30% of the total dataset.

In conclusion, this section has described the dataset requirements and partitioning strategy essential for the development of a successful CNN-based retinopathy detection model. By selecting a diverse and representative dataset and partitioning it into appropriate training, validation, and test subsets, the model can be trained, fine-tuned, and evaluated effectively, ensuring optimal performance and generalizability.

D. Image Preprocessing Techniques

Image preprocessing is a crucial step in the development of a CNN-based retinopathy detection model, as it prepares the dataset for effective learning by enhancing the image quality, ensuring consistency, and increasing the model's robustness. This section focuses on the specific image preprocessing techniques employed in this study.

1. Image Resizing

To ensure consistency across the dataset, all retinal images should be resized to a fixed resolution. This process allows the model to accept images of varying dimensions while ensuring that the input data is homogeneous.

In the preprocessing pipeline, a constant IMG_SIZE is set to 256, representing the uniform size (both in terms of width and height) to which all input images will be resized. This standardization is crucial for ensuring consistent input dimensions for subsequent machine learning models, given that neural networks, in particular, require fixed-size input tensors. The preprocess_image function is tasked with several preprocessing steps, including reading an image from its file path, converting its color representation, cropping away potential dark borders, and resizing it to the predetermined IMG_SIZE. To facilitate the storage of preprocessed images for training, the number of images in the dataset, denoted as N, is inferred from the first dimension of the train_Data array. Subsequently, an empty array x_train of shape (N, IMG_SIZE, IMG_SIZE, 3) is initialized. This structure is designed to accommodate N RGB images, each of dimension 256x256. The chosen data type for this array is np.uint8, optimal for representing standard 8-bit image pixel intensities ranging from 0 to 255. Through these steps, the preprocessing pipeline ensures a consistent format and size for all images, making them ready for training and evaluation in machine learning models.

2. Image Enhancement

Enhancing the images improves the visibility of retinal structures and lesions, allowing the model to learn more effectively from the dataset. Various image enhancement techniques can be employed, such as:

i. Contrast Stretching

This technique improves the contrast in the image by stretching the range of intensity values to cover the entire dynamic range.

In the updated preprocessing pipeline, the Python Imaging Library (PIL) is employed to read and handle images, known for its extensive file format support and efficient opening of large images. Specifically, the Image.open() method from PIL is utilized to load the image, which is subsequently converted to a numpy array for further processing.

Once the image data is available as a numpy array, its color space is transformed from RGB to BGR using OpenCV's cvtColor method. This transformation facilitates subsequent image operations that are to be carried out using OpenCV functions, which conventionally expect images in the BGR format.

Post color space conversion, the previously discussed crop_image_from_gray function is employed to eliminate dark borders from the image. The cropped image is then resized to the predetermined dimensions, given by the constant IMG_SIZE.

An essential robustness measure has been introduced to check if the image is empty after preprocessing steps, as certain operations might lead to an empty image matrix. If the image is found to be empty, a ValueError is raised with a descriptive error message, enabling developers or users to identify potential issues with the input image or the processing steps.

To enhance the image's features and diminish noise, a Gaussian blur is applied using OpenCV's GaussianBlur function, with sigmaX as the standard deviation. A technique to sharpen the image is then employed, where a weighted sum of the original image and its blurred version is calculated. This amplifies the details and intensifies the contrast, making features more pronounced.

Finally, before returning the preprocessed image, its color space is converted back to RGB, aligning with common image representation standards and ensuring compatibility with potential downstream tasks or visualizations.

This comprehensive preprocessing pipeline ensures that input images are not only standardized in size but are also enhanced for clarity and feature visibility, making them optimally suited for machine learning or computer vision tasks.

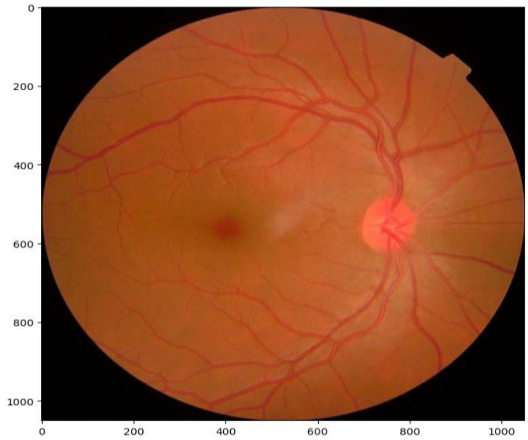
3. Data Augmentation

This process entails creating extra training instances by introducing varied alterations to the initial images. This step is particularly valuable when dealing with a limited number of training samples or when aiming to improve the model's performance on diverse and noisy data.

For our developed method, here is an explanation of data augmentation, the employed ImageDataGenerator introduces the following transformations:

- zoom_range=0.6: This parameter allows for a random zoom on the image. The value 0.6 means that the image can be zoomed in or out by up to 60%. This helps the model to recognize objects or features of various scales.
- fill_mode='constant' and cval=0.: When an image is rotated or shifted, some of its parts might fall outside the original frame, leaving empty spaces in the image. The fill_mode parameter determines how to fill these spaces. The chosen mode, 'constant', fills these gaps with a constant value, which is defined as 0. by the cval parameter
- horizontal_flip=True and vertical_flip=True: These parameters enable the random flipping of images both horizontally and vertically. This introduces variability in the spatial orientation of features, aiding the model in being orientation-invariant.
- rotation_range=360: Images can be randomly rotated within a range of 0 to 360 degrees, offering full rotational variability.
- width_shift_range=0.1 and height_shift_range=0.1: These parameters allow the image to be randomly shifted vertically and horizontally by up to 10% of its width and height, respectively
- rescale=1./255: All pixel values in the image are scaled to the range [0,1] by dividing them by 255. This normalization is essential for most deep learning models as it ensures that input values are small, making the optimization process smoother.

In summary, the chosen data augmentation techniques ensure During training, the model is provided with a wide range of modifications of the original photos, enhancing its capacity to generalize to new data. The application of such transformations becomes crucial, especially in medical imaging or scenarios where capturing the full diversity of real-world data is challenging, and here is the images after preprocessing:



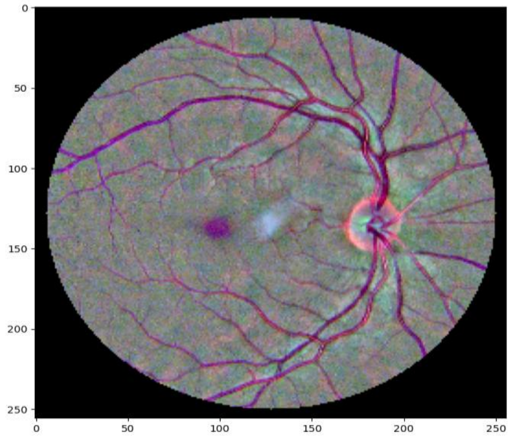


Figure 8: Image before preprocessing Figure 9: Image after preprocessing

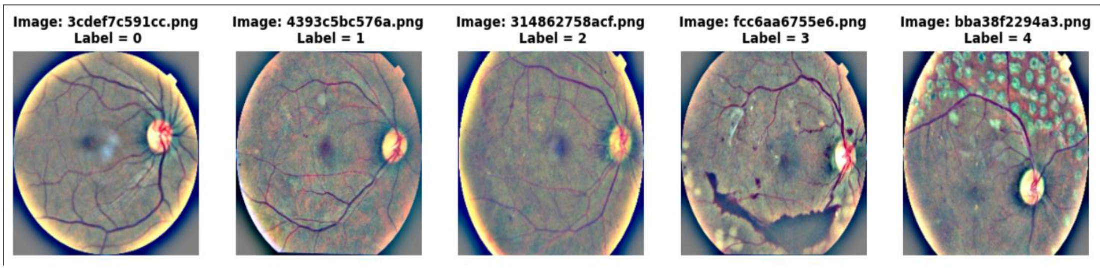


Figure 10: preprocessed image

E. Feature extraction using Discrete Cosine Transform (DCT)

Feature extraction is a vital step in the development of a CNN-based retinopathy detection model, as it enables the model to focus on the most relevant features while reducing computational complexity. The Discrete Cosine Transform (DCT) is a widely used technique for feature extraction in image processing applications. This section provides a detailed overview of the DCT-based feature extraction process employed in this study.

1. Discrete Cosine Transform (DCT)

The DCT is a linear mathematical transformation that converts a given image from the spatial domain to the frequency domain. The frequency domain representation of the image emphasizes the energy distribution across different frequency bands, allowing the model to focus on the most relevant and informative features. DCT is particularly suitable for retinopathy detection, as it captures the essential structural information present in the retinal images, such as blood vessels and lesions, while eliminating high-frequency noise.

2. DCT Application

To apply DCT to the preprocessed retinal images, the images are first divided into non-overlapping blocks of fixed size (e.g., 8x8 or 16x16 pixels). Next, the DCT is applied independently to each block, resulting in a set of DCT coefficients that represent the energy distribution across different spatial frequencies. These DCT coefficients form a matrix of the same size as the original block, with the top-left element (DC coefficient) representing the average energy of the block and the remaining elements (AC coefficients) corresponding to the energy distribution across different frequency bands.

3. DCT Feature Selection

Since the DCT coefficients capture the energy distribution across various frequencies, it is essential to select the most informative features that contribute to accurate retinopathy detection. Typically, lower-frequency coefficients (located in the top-left corner of the DCT matrix) contain the most relevant structural information, while higher-frequency coefficients (located in the bottom-right corner) represent noise and other less relevant details. A common approach is to select a subset of the low-frequency DCT coefficients, such as the first N coefficients in a zigzag pattern, as features for the CNN model.

Feature Vector Formation: Once the most relevant DCT coefficients are selected, they can be organized into a feature vector for each image block. These feature vectors serve as the input for the CNN model, allowing it to learn the relationships between the extracted features and the presence or severity of retinopathy diseases.

By employing DCT-based feature extraction, the proposed retinopathy detection model can efficiently focus on the most informative and relevant features present in the retinal images. This approach significantly reduces the computational complexity while preserving the essential structural information required for accurate retinopathy detection.

F. Design And İmplementation Of The Convolutional Neural Network (CNN) Model

The Convolutional Neural Network (CNN) lies at the core of this study, serving as the primary tool for retinopathy detection based on the extracted DCT features. This section delves into the design and implementation of the CNN model, detailing the various components, architecture, and training process.

1. CNN Architecture

The architecture of the CNN model consists of several layers, each with a specific purpose, working together to learn the patterns and relationships between the DCT features and retinopathy diseases. The primary components of the architecture include:

a. Input Layer

The input layer receives the DCT feature vectors extracted from the retinal images, serving as the starting point for the CNN model.

b. Convolutional Layers

Integral to the architecture of Convolutional Neural Networks (CNNs), these layers specialize in identifying and learning spatial hierarchies and local patterns within input data, often images. The primary mechanism of action in these layers is the application of convolutional filters. These filters, or small receptive fields, move across the input data (e.g., an image) in a sliding window fashion. As they traverse, they perform mathematical convolutions, effectively highlighting particular features in the image such as edges, textures, and shapes. As a result, they generate what are termed as "feature maps" or "activation maps." Each feature map is a representation of the input data, focusing on one specific feature or pattern detected by the filter. Over subsequent layers, these filters can recognize increasingly complex patterns, enabling the model to understand intricate details and spatial relationships in the data. The adaptability of convolutional layers in recognizing various patterns makes them a cornerstone of deep learning, especially in tasks that involve visual data processing.

c. Activation Functions

Non-linear activation functions, such as ReLU (Rectified Linear Unit), introduce non-linearity to the model, allowing it to learn complex relationships between the input features and the target output.

d. Pooling Layers

Pooling layers reduce the spatial dimensions of the feature maps while preserving the most relevant information. This process reduces computational complexity and promotes translation invariance, making the model more robust.

e. Fully Connected Layers

Fully connected layers are responsible for combining the information from previous layers and producing the final output. They map the high-level features learned by the convolutional and pooling layers to the desired output, such as retinopathy disease classification.

f. Output Layer

The output layer generates the final predictions of the model, typically using a softmax activation function for multi-class classification problems. The output layer should have as many neurons as there are classes in the problem (e.g., healthy, mild retinopathy, severe retinopathy).

2. Model Implementation

The CNN model can be implemented using popular deep learning frameworks such as TensorFlow, PyTorch, or Keras. These frameworks provide high-level APIs for defining the model architecture, specifying the training process, and evaluating the model's performance.

3. Training Process

The training process involves adjusting the model's weights to minimize the difference between its predictions and the ground truth labels. This process requires the following components:

a. Loss Function

The loss function quantifies the discrepancy between the model's predictions and the actual labels. Common loss functions for classification tasks include categorical cross-entropy or binary cross-entropy.

b. Optimizer

The optimizer is responsible for updating the model's weights based on the calculated loss. Popular optimizers include Stochastic Gradient Descent (SGD), Adam, and RMSProp.

c. Hyperparameters

Hyperparameters control various aspects of the training process, such as the learning rate, batch size, and the number of training epochs. These values need to be tuned to achieve optimal model performance.

4. Model Validation and Evaluation

During the training process, the model's performance should be monitored using the validation set. This evaluation helps to prevent overfitting and allows for hyperparameter tuning. Once the training is complete, the model's final performance can be assessed using the test set, providing an unbiased estimation of its retinopathy detection capabilities.

By designing and implementing a robust CNN model and training it on the preprocessed retinal images with DCT features, the study aims to develop an effective tool for retinopathy detection. The architecture, implementation, training process, and evaluation strategy detailed in this section ensure that the resulting model is both accurate and generalizable, with the potential to significantly improve retinopathy detection in real-world clinical settings. By leveraging the power of convolutional neural networks and the efficiency of Discrete Cosine Transform-based feature extraction, this study aims to provide a reliable, efficient, and automated solution for early diagnosis and monitoring of retinopathy diseases, ultimately contributing to improved patient outcomes and the overall management of these conditions.

G. Architecture Of The CNN Model

The architecture of the CNN model plays a crucial role in determining its effectiveness for retinopathy detection. It comprises several interconnected layers designed to work together to learn the patterns and relationships between the DCT features and retinopathy diseases. This section elaborates on the architecture of the CNN model, describing each layer and its purpose in detail.

1. Input Layer

The input layer serves as the entry point for the DCT feature vectors extracted from the retinal images. It receives these feature vectors and passes them to the subsequent layers for further processing. The dimension of the input layer should match the size of the feature vectors.

2. Convolutional Layers

Convolutional layers are the fundamental building blocks of Convolutional Neural Networks (CNNs). Their primary role is to discern local patterns, structures, and spatial relationships within the input data, often images or videos. In the intricate architecture of a CNN, several convolutional layers can be layered atop one another. This layering allows the network to learn patterns of increasing complexity and abstraction as data progresses through the layers.

In each convolutional layer, a collection of filters, also referred to as kernels, is applied. These filters are not static; they are trainable parameters, meaning they adjust and refine during the learning process to better identify specific patterns in the input. As these filters slide over the input data, they produce what's known as "feature maps" or "activation maps." Each feature map represents the response of the input data to a particular filter, emphasizing where certain patterns or features are detected.

3. Activation Functions

Activation functions introduce non-linearity into the model, allowing it to learn complex relationships between input features and output predictions. ReLU (Rectified Linear Unit) is a popular choice for activation functions in CNNs due to its computational efficiency and ability to mitigate the vanishing gradient problem. Activation functions are applied to the output of convolutional layers.

4. Pooling Layers

Pooling layers are interspersed between convolutional layers to reduce the spatial dimensions of the feature maps, thereby reducing computational complexity and promoting translation invariance. Common types of pooling layers include max pooling and average pooling, which retain the maximum or average value, respectively, within a specified neighborhood of each feature map.

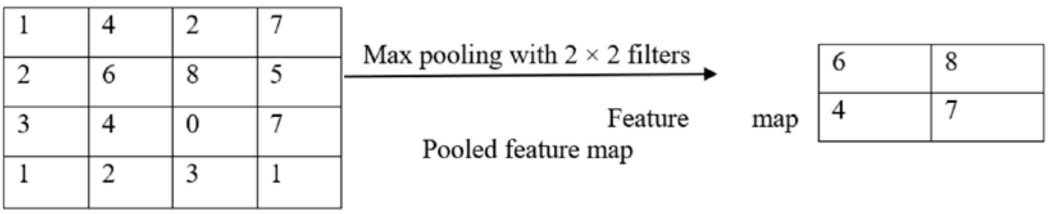


Figure 11: Max pooling layer of the CNN

5. Dropout Layers

Dropout layers can be introduced to the architecture to regularize the model and prevent overfitting. They randomly "drop" a proportion of the neurons during training, effectively reducing the model's capacity and forcing it to learn more robust features. The dropout rate is a hyperparameter that controls the proportion of neurons dropped during training.

6. Fully Connected Layers

The fully connected layers combine the high-level features learned by the convolutional and pooling layers to produce the final output. These layers establish dense connections between all neurons in consecutive layers, enabling the model to learn global patterns and make predictions based on the extracted features.

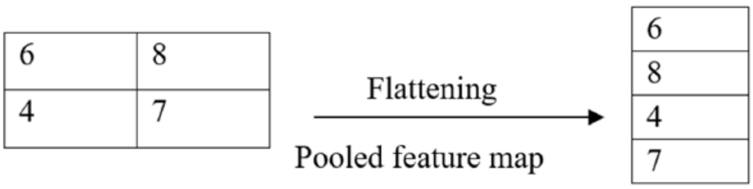


Figure 12: Flattening Layer

7. Output Layer

The output layer generates the final predictions of the model. For multi-class classification problems, a softmax activation function is typically used to produce probability distributions across the different classes. The number of neurons in the output layer should equal the number of classes in the problem (e.g., healthy, mild retinopathy, severe retinopathy).

The architecture of the CNN model is designed to effectively learn the patterns and relationships between the DCT features and retinopathy diseases. By carefully designing and configuring the layers in the architecture, the study aims to create a powerful and accurate tool for retinopathy detection, with the potential to improve early diagnosis and patient outcomes.

H. Activation Functions And Loss Function

Activation functions and loss functions are essential components of the CNN model, as they govern the model's learning process and determine how well it captures the relationships between the input features and output predictions. This section provides a detailed discussion of the activation functions used within the model and the choice of an appropriate loss function.

1. Activation Functions

Activation functions are applied to the output of neurons in a neural network, introducing non-linearity and enabling the model to learn complex relationships between input features and output predictions. The following activation functions are commonly used in CNN models:

- ReLU (Rectified Linear Unit): ReLU is a popular activation function in CNNs due to its computational efficiency and ability to mitigate the vanishing gradient problem. ReLU introduces non-linearity by setting all negative values to zero, while preserving positive values. This activation function is typically applied to the output of convolutional layers and fully connected layers.

- Leaky ReLU: Leaky ReLU is a variant of the ReLU function that addresses the "dying ReLU" problem, in which neurons with negative inputs become inactive and cease to learn. Leaky ReLU assigns a small positive slope to negative inputs, ensuring that neurons remain active even when their inputs are negative. This activation function can be used as an alternative to the standard ReLU.

- Sigmoid: The sigmoid activation function maps input values to the range (0, 1), providing a smooth, continuous output. While sigmoid functions were widely used in early neural networks, they are less common in modern CNNs due to their susceptibility to the vanishing gradient problem.

2. Loss Function

The loss function quantifies the discrepancy between the model's predictions and the actual ground truth labels, guiding the optimization process during training. For the retinopathy detection problem, which involves multi-class classification, the following loss functions are relevant:

3. Categorical Cross-Entropy

Categorical cross-entropy is a widely used loss function for multi-class classification problems. It measures the difference between the predicted probability distribution and the true probability distribution, penalizing the model when its predictions are far from the actual labels. This loss function is well-suited for problems with multiple, mutually exclusive classes, such as retinopathy detection.

In summary, the careful selection of activation functions and the appropriate loss function play a crucial role in the effectiveness of the CNN model for retinopathy detection. Activation functions ensure the model's capacity to collect complicated information, non-linear relationships between the DCT features and retinopathy diseases, while the loss function guides the optimization process to minimize prediction errors using proper activation function efficacy and the categorical cross-entropy loss function, the study aims to develop a reliable and efficient tool for the detection and classification of retinopathy diseases, ultimately contributing to improved patient outcomes and more effective management of these conditions in clinical settings.

İ. Model Training, Validation, And Optimization

To ensure the effectiveness of the CNN model for retinopathy detection, a systematic approach to model training, validation, and optimization is essential. This section outlines the key steps involved in containing the model, validating its performance, and fine-tuning the model's hyperparameters to achieve optimal results.

Training and Validation Data: The dataset is split into separate training and validation sets, typically using an 80/20. The training set is used to adjust the model's weights during the learning process, while the validation set helps assess the model's performance on unseen data and monitor for overfitting. This partitioning enables an unbiased evaluation of the model's generalization capabilities.

Batch Size and Epochs: The batch size refers to the number of samples used for each weight update during training. Smaller batch sizes typically lead to more stable training, while larger batch sizes allow for faster computation. The total number of iterations the model developed performs across the whole training dataset is called or we can say represented by the number of epochs. The choice of batch size and epochs should strike a balance between training stability, computational efficiency, and achieving convergence.

Learning Rate and Optimizer: The learning rate is a critical hyperparameter that controls the step size during gradient descent optimization. A suitable learning rate enables the model to converge efficiently without oscillating or overshooting the optimal weights. Popular optimizers for CNNs include Stochastic Gradient Descent (SGD), Adam, and RMSprop, which automatically adjust the learning rate during training for improved convergence.

Regularization Techniques: Regularization techniques, such as L1 and L2 regularization, dropout layers, and early stopping, can help prevent overfitting by adding constraints to the model's complexity. These techniques encourage the model to learn more robust features and generalize better to new data.

Hyperparameter Tuning: The model's hyperparameters, including the number of layers, filter sizes, learning rate, and regularization strength, may need to be adjusted to obtain the best results. The most effective combination of hyperparameters for a given issue may be found by exploring the hyperparameter space using techniques like grid search, random search, and Bayesian optimization.

By following a structured approach to model training, validation, and optimization, the study aims to develop a CNN model for retinopathy detection that achieves high accuracy and generalizes well to new data. This process will help ensure that the proposed model is a reliable and effective tool for detecting and classifying retinopathy diseases, contributing to better patient outcomes and more efficient management of these conditions.

To provide a more detailed account of the model training, validation, and optimization, the following sub-sections delve deeper into each of these critical aspects:

1. Data Augmentation

Data augmentation techniques can be applied to the training set to artificially increase the dataset size and improve the model's ability to generalize to unseen data. Common data augmentation techniques for images include rotation, flipping, zooming, and adding noise. By augmenting the dataset, the model can learn more robust features that account for various transformations and variations in retinal images.

2. Cross-Validation

K-fold cross-validation can be employed to further enhance the evaluation of the model's performance. In k-fold cross-validation, the dataset is divided into k equal-sized subsets. The model is then trained and validated k times, using a different subset as the validation set in each iteration. The average performance across all folds is used to estimate the model's performance. This method gives a more accurate evaluation of the model's capacity for generalization, as it considers multiple training and validation set combinations.

3. Weight Initialization

The initial values of the model's weights can impact the speed of convergence and the quality of the final solution. Techniques such as Glorot initialization, He initialization, and LeCun initialization can be employed to provide better initial weight values based on the specific activation functions and network architecture. Proper weight initialization can lead to faster convergence and improved model performance.

4. Learning Rate Scheduling

Instead of using a fixed learning rate throughout the training process, learning rate scheduling can be applied to adapt the learning rate based on the training progress. Methods such as step decay, exponential decay, and cosine annealing can be used to reduce the learning rate over time, ensuring that the model converges more smoothly to the optimal solution.

5. Model Ensemble

Combining multiple models or training runs in an ensemble can improve the overall performance and robustness of the retinopathy detection system. Averaging the predictions from multiple models or applying more advanced ensemble techniques, such as stacking or bagging, can lead to more accurate and stable predictions by leveraging the strengths of each individual model.

By incorporating these additional considerations into the model training, validation, and optimization process, the study aims to develop a highly accurate and reliable CNN model for retinopathy detection. Addressing aspects such as data augmentation, cross-validation, weight initialization, learning rate scheduling, and model ensembles helps to create a robust and efficient solution capable of effectively detecting and classifying retinopathy diseases, thereby contributing to improved patient outcomes and better management of these conditions.

J. Performance Evaluation

The evaluation of the CNN model's performance is a crucial step in assessing its effectiveness in retinopathy detection. This section outlines the key elements involved in evaluating the model, including the use of test data, evaluation metrics, and the comparison of results to existing methods.

1. Metrics: Accuracy, Sensitivity, Specificity, And F1-Score

1. Test Data

To obtain an unbiased estimate of the model's performance, a separate test dataset, not used during training or validation, is employed. This dataset contains retinal images with ground truth labels, enabling the evaluation of the model's predictions against the true disease classifications. By using test data, the study can assess the model's ability to generalize to new, unseen data, which is critical for its practical application in a clinical setting.

2. Evaluation Metrics

To measure the model's performance, various evaluation metrics are employed to provide insights into different aspects of its classification capabilities. The following metrics are commonly used for evaluating multi-class classification models:

- Accuracy: The proportion of correctly classified samples out of the total number of samples. Accuracy provides an overall measure in the case of class imbalance, the performance of the model may be less informative.

- Precision: The proportion of true positives out of the total number of predicted positives. Precision measures the model's ability to correctly identify positive cases while minimizing false positives.

- Recall (also known as Sensitivity or True Positive Rate): This metric represents the proportion of actual positives (true positives) that the model successfully identified as such, compared to the total number of actual positives in the dataset. In simpler terms, it gauges how well the model is at detecting the positive instances when they are actually there. It's a crucial metric, especially in scenarios where missing a positive instance (like a disease or malfunction) could have dire consequences.

- F1-Score: This is a metric that seeks to strike a balance between precision (how many of the items identified as positive are actually positive) and recall (how many of the actual positives were identified). It is the harmonic mean of these two values. By combining precision and recall into a single number, the F1-score offers a comprehensive measure of the model's accuracy and completeness. An F1-score closer to 1 indicates a better balance between precision and recall, whereas a score closer to 0 suggests that the model struggles either with precision or recall or both.

- Confusion Matrix: The values in the confusion matrix are used to calculate various performance metrics such as accuracy, precision, recall, and F1 score, which help to understand the effectiveness of the classifier. The true positives (TP) are the number of instances that are correctly classified as positive. False positives (FP) are instances that are incorrectly classified as positive when they are actually negative. True negatives (TN) are instances that are correctly classified as negatives, while false negatives (FN) are instances that are incorrectly classified as negatives when they are actually positives. True Positives: Positive instances that were correctly classified.

  - ❖ False Positives: Incorrectly classified positive instances.
  - ❖ True Negatives: Correctly classified negative instances.
  - ❖ False Negatives: Incorrectly classified negative instances.

Accuracy is defined as the total number of correct predictions divided by the total number of predictions made by the model. Precision is the ratio of true positive predictions to the total number of positive predictions made by the model. Recall, also known as sensitivity or the "true positive rate," is the ratio of true positive predictions to the total number of actual positive instances in the data. The F1-score is the harmonic mean of precision and recall and provides a single metric that balances both metrics. Specificity is the ratio of true negative predictions to the total number of actual negative instances in the data. These metrics provide valuable information about the performance of a classification model and help identify areas for improvement.

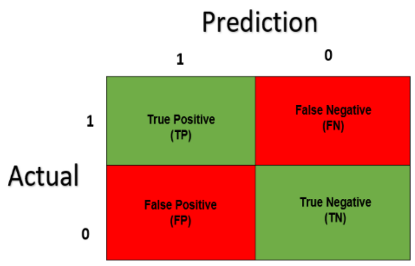


Figure 13: Confusion Matrix for Binary Datasets

The four parameters calculated to compare the performance of each model are:

- **Accuracy:** Represents the amount of tests accurately classified by the classifier.

$$\text{Accuracy} = \frac{(\text{True Positives} + \text{True Negatives})}{(\text{True Positives} + \text{True Negatives} + \text{False Positives} + \text{False Negatives})}$$

- **Precision:** Refers to the accuracy of the classifier, specifically the rate at which the samples are correctly identified as positive among all the samples classified as positive in the test set.

$$\text{Precision} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Positives})}$$

- **Recall:** Reflects the integrity of the model, also known as the detection rate or sensitivity of the model. It measures the successful labeling of positive samples in the test set.

$$\text{Recall} = \frac{\text{True Positives}}{(\text{True Positives} + \text{False Negatives})}$$

- **F1-Score:** The F1-Score denotes the harmonic average concerning Precision and Recall, serving as a crucial metric that ensures an even-handed and comprehensive analysis of both the occurrences of false positives and false negatives in a given dataset. This balance is vital for understanding the true efficacy of a classification system.

$$\text{F1}_{\text{Score}} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

3. Comparison to Existing Methods

To establish the value of the proposed CNN model in retinopathy detection, it is essential to compare its performance to existing methods, such as other machine learning algorithms or manual diagnosis by medical experts. This comparison helps to gauge the model's relative effectiveness and identify any advantages or drawbacks compared to current practices.

4. Analysis of Results

By evaluating the model's performance using the selected metrics and comparing it to existing methods, the study can identify the model's strengths and weaknesses, as well as potential areas for improvement. This analysis aids in refining the model and ensuring that it is a valuable tool for detecting and classifying retinopathy diseases in a clinical context.

By following a comprehensive approach to performance evaluation, the study aims to assess the CNN model's effectiveness in retinopathy detection and its potential for practical application. Through the use of test data, evaluation metrics, comparison to existing methods, and thorough analysis of results, the performance evaluation process provides valuable insights into the model's capabilities and potential areas for enhancement.

IV. EXPERIMENTAL RESULTS

The model was trained and evaluated using a robust computational setup. The experiments were carried out on a 3.2 GHz machine with 16 GB of memory and a GTX 1650ti GPU. The development environment was Jupyter Notebook, provided by Anaconda, ensuring reliable, reproducible results.

A. Convolutional Neural Network Model

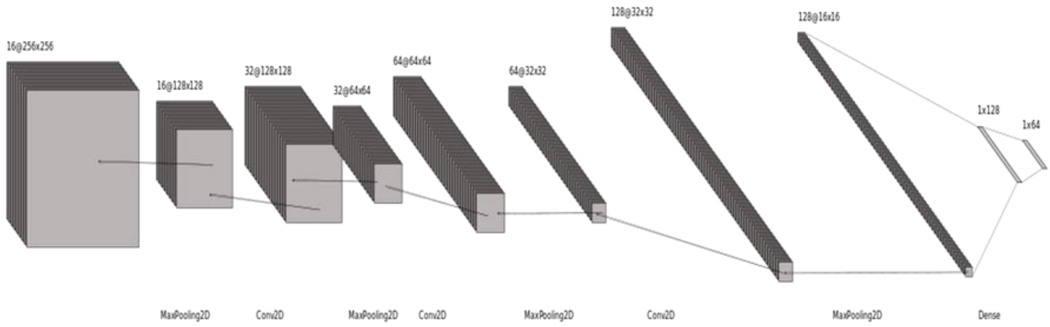


Figure 14: CNN model Architecture

The size of each input image is (256, 256, 3). When using Keras, images are typically processed in batches of a certain size. As such, an additional dimension is introduced to account for the batch size. The actual batch size can vary depending on the dataset being used, which is why this dimension is represented by the value "None." Hence, the initial shape of the input is represented as (None, 256, 256, 3).

When a convolution operation is performed on a (256, 256) image with a filter size, and considering standard parameters like a stride and dilation rate of 1, along with the 'same' padding, the output dimensions remain (256, 256). Given that there are 16 filters in the first convolutional layer, the output shape becomes (256, 256, 16). Following this, the MaxPooling layer, which typically has a stride of 2, reduces the spatial dimensions of the image by half. Thus, the output shape post-pooling becomes (128, 128, 16).

The next Conv2D layer with 32 filters processes the pooled output, retaining the spatial dimensions at (128, 128) but changing the depth to 32 due to the number of filters. This results in an output shape of (128, 128, 32). Post this convolution operation, another MaxPooling layer further halves the spatial dimensions, yielding an output shape of (64, 64, 32).

Continuing in a similar fashion, another convolution is performed on this output using 64 filters, yielding an output shape of (64, 64, 64). Yet another MaxPooling operation on this output then gives a shape of (32, 32, 64).
Thereafter, a convolution with 128 filters retains the spatial dimensions but changes the depth to 128, producing an output shape of (32, 32, 128). Post-convolution, the subsequent MaxPooling layer reduces the dimensions to (16, 16, 128).

Towards the end of this architecture, a Dropout layer is used, which doesn't alter the shape of the tensor but helps in regularizing the model to prevent overfitting. This explains the structure and the progression of shapes within the network, eventually leading to a model with a total of 16,823,733 parameters, all of which are trainable. The number of parameters for a Conv2D layer is given by the following Equation (1):

Parameter= (kernel height*kernel width*input channel*output channel) + (output channels that use bias)

```
 Layer (type)                Output Shape              Param #
=================================================================
 conv2d (Conv2D)             (None, 256, 256, 16)      208

 max_pooling2d (MaxPooling2D  (None, 128, 128, 16)     0
 )

 conv2d_1 (Conv2D)           (None, 128, 128, 32)      2080

 max_pooling2d_1 (MaxPooling  (None, 64, 64, 32)       0
 2D)

 conv2d_2 (Conv2D)           (None, 64, 64, 64)        8256

 max_pooling2d_2 (MaxPooling  (None, 32, 32, 64)       0
 2D)

 conv2d_3 (Conv2D)           (None, 32, 32, 128)       32896

 max_pooling2d_3 (MaxPooling  (None, 16, 16, 128)      0
 2D)

 dropout (Dropout)           (None, 16, 16, 128)       0
...
Total params: 16,823,733
Trainable params: 16,823,733
Non-trainable params: 0
```

Figure 15: The model summary of the created CNN network.

B. Model Trained Using Our Developed Model

The developed model was trained and evaluated over 50 epochs. The model performance metrics, including loss (cross-entropy loss) and accuracy, were observed for both the training and validation set for each epoch.

```
model=Sequential()
model.add(Conv2D(filters=16,kernel_size=2,padding="same",activation="relu",input_shape=(256,256,3)))
model.add(MaxPooling2D(pool_size=2))
model.add(Conv2D(filters=32,kernel_size=2,padding="same",activation="relu"))
model.add(MaxPooling2D(pool_size=2))
model.add(Conv2D(filters=64,kernel_size=2,padding="same",activation="relu"))
model.add(MaxPooling2D(pool_size=2))
model.add(Conv2D(filters=128,kernel_size=2,padding="same",activation="relu"))
model.add(MaxPooling2D(pool_size=2))
model.add(Dropout(0.2))
model.add(Flatten())
model.add(Dense(512,activation="relu"))
model.add(Dropout(0.2))
model.add(Dense(5,activation="softmax"))
model.summary()
```

Figure 16: shows the model summary of the created CNN network

The initial epoch revealed a training accuracy of 66.05%. On the validation set, the accuracy was 70.31%. This established the baseline performance of our model. As the epochs progressed, we noticed an increase in accuracy and a decrease in loss on both the training and validation sets, indicating the learning capability of our model. By the end of the second epoch, the model achieved a training accuracy of 72.43%, and the validation accuracy improved to 70.74%.

By the 7th epoch, the validation accuracy had reached 72.44%, and the model's loss had reduced to 0.494. At the 8th epoch, we observed a significant jump in the validation accuracy, reaching 74.72%. This suggested that the model was generalizing well and was capable of classifying unseen data accurately.

However, in the following epochs, there was a slight increase in the validation loss, despite the model's improved accuracy, suggesting a bit of overfitting. After the 10th epoch, the training accuracy reached 83.78%, and the validation accuracy stood at 78.01%.

C. Model Trained Using Pretrained Model Xception

In my research project, I leveraged the power of the pre-trained Xception model for an image classification task. After applying my preprocessing pipeline and running the model, the Xception model demonstrated strong performance. Specifically, during the 21st epoch, the Xception model achieved an impressive accuracy of 74.52% on the validation set. This result underscores the efficacy of the Xception model in this particular image classification task, illustrating its robust generalization capabilities.

```
Model: "Xception"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 xception (Functional)       (None, 8, 8, 2048)        20861480

 global_average_pooling2d (G  (None, 2048)             0
 lobalAveragePooling2D)

 dense (Dense)               (None, 256)               524544

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 5)                 1285

=================================================================
Total params: 21,387,309
Trainable params: 525,829
Non-trainable params: 20,861,480
_________________________________________________________________
```

Figure 17: shows the model summary of Pretrained model Xception

D. Model Trained Using Pretrained Model ResNet152V2

In a parallel analysis during my research, I applied the same image preprocessing procedure to another renowned pre-trained model, ResNet152V2, in the same image classification task. The performance of the ResNet152V2 model was robust and achieved a notable accuracy. Precisely, during the evaluation stage, the ResNet152V2 model reached an accuracy of 73.07% on the validation set. This result highlights the strong performance and the model's capability to effectively generalize in this specific image classification context.

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_4 (Conv2D)            (None, 254, 254, 64)      1792
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 127, 127, 64)      0
_________________________________________________________________
conv2d_5 (Conv2D)            (None, 125, 125, 32)      18464
_________________________________________________________________
max_pooling2d_4 (MaxPooling2 (None, 62, 62, 32)        0
_________________________________________________________________
flatten_1 (Flatten)          (None, 123008)            0
_________________________________________________________________
dense_3 (Dense)              (None, 256)               31490304
_________________________________________________________________
dropout_2 (Dropout)          (None, 256)               0
_________________________________________________________________
dense_4 (Dense)              (None, 128)               32896
_________________________________________________________________
dropout_3 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_5 (Dense)              (None, 5)                 645
=================================================================
Total params: 31,544,101
```

Figure 18: shows the model summary of the pretrained model Resnet152v2

E. Comparison of Performance

Diabetic retinopathy diagnosis has been significantly enhanced using advanced methodologies such as deep learning, machine learning, and image processing algorithms. To optimize the clarity of images, we employed image processing and augmentation techniques that notably improved their brightness and contrast. For a comparative analysis of performance across different methods, accuracies were systematically tabulated. Notably, leveraging transfer learning techniques with deep learning classifiers yielded superior accuracy, especially for datasets with fewer records. A comparative analysis with existing research findings is detailed in Table.

Table 2: Comparison between all the used methods

Method	Validation Accuracy	Validation Loss
CNN	78.01%	0.4358%
Xception	74.52%	0.5381%
ResNet152V2	73.07%	0.2249%

In conclusion, our developed model demonstrated promising performance with accuracy of 78%, achieving a respectable accuracy in classifying the severity of retinopathy in the APTOS 2019 Blindness Detection dataset. Future work may include tuning the model to mitigate overfitting, enhancing its generalizability, and potentially improving the model's performance further.

V. CONCLUSION

One of the illnesses with the fastest recent growth rates is diabetes. A patient with diabetes has a 30% chance of developing diabetic retinopathy, according to several surveys. There are several phases of DR, from moderate to severe, and PDR (Proliferative Diabetic Retinopathy) is the last stage. If the condition is not discovered in the earlier stages, it progresses to blindness, floaters, and impaired vision in its later stages. There is still a need for simple access to such models despite the fact that several computer vision-based strategies for the automatic identification of DR employing hands-on engineering and end-to-end learning approaches have been offered.

In this regard, the study described in this paper suggests a novel method for early DR detection based on a Convolutional Neural Network (CNN) model based on Discrete Cosine Transform (DCT) that is specially designed for mobile devices.

In developing this model, we utilized one publicly accessible Kaggle dataset for training and validating our CNN-based DCT network. We found that image pre-processing is an essential step in this process.

Our results indicate that the CNN-based DCT model successfully classifies the early stages of DR. This aligns with our central aim of encouraging early detection before the progression to more advanced stages such as PDR. However, it's important to note that the model does encounter some difficulty in classifying the final stages of DR.

Early DR stages are notoriously difficult to classify, as confirmed by related work in the field. However, our proposed CNN-based DCT model exhibits encouraging results for early stage classification. It particularly excels at distinguishing Non_DR from the mild to moderate stages, thanks to the pre-processing technique employed.

Nonetheless, our study did uncover a limitation in the model's performance when it came to classifying the last two stages of the disease. We believe this could be due to the fact that the characteristic features of DR's later stages do not replace the earlier stages' features but rather augment them. This means that features from the earlier stages still exist in the later stages, complicating their differentiation. The model's performance for these advanced phases will be improved in the future.

VI. REFERENCES

ADRIMAN, R.; MUCHTAR, K.; Maulina, N. **Performance Evaluation of Binary Classification of Diabetic Retinopathy through Deep Learning Techniques Using Texture Feature**. Procedia Comput. Sci. 2021, 179, 88–94.

ALI, R.; HARDIE, R.C.; NARAYANAN, B.N.; KEBEDE, T.M. IMNets: **Deep Learning Using an Incremental Modular Network Synthesis Approach for Medical Imaging Applications.** Appl. Sci. 2022, 12, 5500.

BODAPATI, J.D.; SHAIK, N.S.; NARALASETTI, V. **Composite Deep Neural Network with Gated-Attention Mechanism for Diabetic Retinopathy Severity Classification**. J. Ambient Intell. Humaniz. Comput. 2021, 12, 9825–9839.

DAS, S.; KHARBANDA, K.; **Suchetha, M.; Raman, R.; Dhas, E. Deep Learning Architecture Based on Segmented Fundus Image Features for Classification of Diabetic Retinopathy. Biomed**. Signal Process. Control 2021, 68, 102600.

Diabetic Retinopathy (resized), **Resized version of the Diabetic Retinopathy Kaggle competition dataset** https://www.kaggle.com/tanlikesmath/diabetic-retinopathy-resized,2019.

Diabetic Retinopathy Detection SHAP, Data Preprocessing https://www.kaggle.com/code/simrahikram/diabetic-retinopathy-detection-shap/notebook

DR. OKATAN ALI, KARLIK B, DEMIREZEN F, **Detection of retinopathy diseases using an artificial neural network based on the discrete cosine transform January 2009 Neural Network World** 19(2):215-221.

ELLOUMI, Y.; ABROUG, N.; BEDOUI, M.H. **End-to-End Mobile System for Diabetic Retinopathy Screening Based on Lightweight Deep Neural Network. Lect. Notes Comput. Sci**. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinform.) 2022, 13205 LNCS, 66–77.

FANG, L.; QIAO, H. **Diabetic Retinopathy Classification Using a Novel DAG Network Based on Multi-Feature of Fundus Images. Biomed**. Signal Process. Control 2022, 77, 103810.

FATIMA; IMRAN, M.; ULLAH, A.; ARIF, M.; **Noor, R. A Unified Technique for Entropy Enhancement Based Diabetic Retinopathy Detection Using Hybrid Neural Network. Comput**. Biol. Med. 2022, 145, 105424.

GAO, Z.; JIN, K.; YAN, Y.; LIU, X.; SHI, Y.; GE, Y.; PAN, X.; LU, Y.; WU, J.; WANG, Y.; ET AL. **End-to-End Diabetic Retinopathy Grading Based on Fundus Fluorescein Angiography Images Using Deep Learning**. Graefe's Arch. Clin. Exp. Ophthalmol. 2022, 260, 1663–1673.

GAYATHRI, S.; GOPI, V.P.; PALANISAMY, P. **Diabetic Retinopathy Classification Based on Multipath CNN and Machine Learning Classifiers**. Phys. Eng. Sci. Med. 2021, 44, 639–653.

GUNASEKARAN, K.; PITCHAI, R.; CHAITANYA, G.K.; SELVARAJ, D.; ANNIE SHERYL, S.; ALMOALLIM, H.S.; ALHARBI, S.A.; RAGHAVAN, S.S.; TESEMMA, B.G. **A Deep Learning Framework for Earlier Prediction of Diabetic Retinopathy from Fundus Photographs**. Biomed Res. Int. 2022, 2022, 1–15.

K. SIMONYAN AND A. ZISSERMAN. **Very deep convolutional networks for large-scale image recognition**. arXiv preprintarXiv:1409.1556, 2014.

KALYANI, G.; JANAKIRAMAIAH, B.; KARUNA, A.; PRASAD, L.V**. Narasimha Diabetic Retinopathy Detection and Classification Using Capsule Networks**. Complex Intell. Syst. 2021, 2021, 1–14.

KANAKAPRABHA, S.; RADHA, D.; SANTHANALAKSHMI, S. **Diabetic Retinopathy Detection Using Deep Learning Models**. Int. 2022, 302, 75–90.

KHAN, A.; KULKARNI, N.; KUMAR, A.; KAMAT, A**. D-CNN and Image Processing Based Approach for Diabetic Retinopathy Classification**. Appl. Inf. Process. Syst. 2022, 1354, 283–291.

KOBAT, S.G.; BAYGIN, N.; YUSUFOGLU, E.; BAYGIN, M.; BARUA, P.D.; DOGAN, S.; YAMAN, O.; CELIKER, U.; YILDIRIM, H.; TAN, R.S.; et al. **Automated Diabetic Retinopathy Detection Using Horizontal and Vertical Patch Division-Based Pre-Trained DenseNET with Digital Fundus Images**. Diagnostics 2022, 12, 1975.

LUO, X.; PU, Z.; XU, Y.; WONG, W.K.; SU, J.; DOU, X.; YE, B.; HU, J.; MOU, L. MVDRNet: **Multi-View Diabetic Retinopathy Detection by Combining DCNNs and Attention Mechanisms**. Pattern Recognit. 2021, 120, 108104.

MATH, L.; FATIMA, R. **Adaptive Machine Learning Classification for Diabetic Retinopathy. Multimed**. Tools Appl. 2021, 80, 5173–5186.

MENAOUER, B.; DERMANE, Z.; EL HOUDA KEBIR, N.; MATTA, N**. Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach**. SN Comput. Sci. 2022, 3, 357.

QURESHI, I.; MA, J.; ABBAS, Q. **Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning**. Multimed Tools Appl 2021, 80, 11691–11721.

RAGAB, M.; ALJEDAIBI, W.H.; NAHHAS, A.F.; **Alzahrani, I.R. Computer Aided Diagnosis of Diabetic Retinopathy Grading Using Spiking Neural Network**. Comput. Electr. Eng. 2022, 101, 108014.

SRIDHAR, S.; PRADEEPKANDHASAMY, J.; SINTHUJA, M.; STERLIN MINISH, T.N. **Diabetic Retinopathy Detection Using Convolutional Nueral Networks Algorithm**. Mater. Today Proc. 2021.

TAJBAKHSH N, SHIN JY, GURUDU SR, HURST RT, KENDALL CB, **Gotway MB, et al. Convolutional neural networks for medical image analysis: full training or fine tuning? IEEE Trans Med Imaging**. 2016;35(5):1299–312.

VIVES-BOIX, V.; Ruiz-Fernández, D. **Diabetic Retinopathy Detection through Convolutional Neural Networks with Synaptic Metaplasticity**. Comput. Methods Programs Biomed. 2021, 206, 106094.

JOURNALS

ALI, R.; HARDIE, R.C.; NARAYANAN, B.N.; KEBEDE, T.M. IMNETS: **Deep Learning Using an Incremental Modular Network Synthesis Approach for Medical Imaging Applications.** Appl. Sci. 2022, 12, 5500.

DAS, S.; KHARBANDA, K.; SUCHETHA, M.; RAMAN, R.; DHAS, E. **Deep Learning Architecture Based on Segmented Fundus Image Features for Classification of Diabetic Retinopathy**. Biomed. Signal Process. Control 2021, 68, 102600.

ELLOUMI, Y.; ABROUG, N.; BEDOUI, M.H. **End-to-End Mobile System for Diabetic Retinopathy Screening Based on Lightweight Deep Neural Network.** Lect. Notes Comput. Sci. (Incl. Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinform.) 2022, 13205 LNCS, 66–77.

FANG, L.; QIAO, H. **Diabetic Retinopathy Classification Using a Novel DAG Network Based on Multi-Feature of Fundus Images.** Biomed. Signal Process. Control 2022, 77, 103810.

GAO, Z.; JIN, K.; YAN, Y.; LIU, X.; SHI, Y.; GE, Y.; PAN, X.; LU, Y.; WU, J.; WANG, Y.; ET AL. **End-to-End Diabetic Retinopathy Grading Based on Fundus Fluorescein Angiography Images Using Deep Learning.** Graefe's Arch. Clin. Exp. Ophthalmol. 2022, 260, 1663–1673.

GUNASEKARAN, K.; PITCHAI, R.; CHAITANYA, G.K.; SELVARAJ, D.; ANNIE SHERYL, S.; ALMOALLIM, H.S.; ALHARBI, S.A.; RAGHAVAN, S.S.; TESEMMA, B.G. **A Deep Learning Framework for Earlier Prediction of Diabetic Retinopathy from Fundus Photographs.** Biomed Res. Int. 2022, 2022, 1–15.

LUO, X.; PU, Z.; XU, Y.; WONG, W.K.; SU, J.; DOU, X.; YE, B.; HU, J.; MOU, L. MVDRNet: **Multi-View Diabetic Retinopathy Detection by Combining DCNNs and Attention Mechanisms.** Pattern Recognit. 2021, 120, 108104.

MENAOUER, B.; DERMANE, Z.; EL HOUDA KEBIR, N.; **Matta, N. Diabetic Retinopathy Classification Using Hybrid Deep Learning Approach**. SN Comput. Sci. 2022, 3, 357.

RAGAB, M.; ALJEDAIBI, W.H.; NAHHAS, A.F.; ALZAHRANI, I.R. **Computer Aided Diagnosis of Diabetic Retinopathy Grading Using Spiking Neural Network.** Comput. Electr. Eng. 2022, 101, 108014.

VIVES-BOIX, V.; RUIZ-FERNÁNDEZ, D. **Diabetic Retinopathy Detection through Convolutional Neural Networks with Synaptic Metaplasticity. Comput.** Methods Programs Biomed. 2021, 206, 106094.

PAPERS

BODAPATI, J.D.; SHAIK, N.S.; NARALASETTI, V. **Composite Deep Neural Network with Gated-Attention Mechanism for Diabetic Retinopathy**

Severity Classification. J. Ambient Intell. Humaniz. Comput. 2021, 12, 9825–9839.

GAYATHRI, S.; GOPI, V.P.; PALANISAMY, P. **Diabetic Retinopathy Classification Based on Multipath CNN and Machine Learning Classifiers.** Phys. Eng. Sci. Med. 2021, 44, 639–653.

KALYANI, G.; JANAKIRAMAIAH, B.; KARUNA, A.; PRASAD, L.V. **Narasimha Diabetic Retinopathy Detection and Classification Using Capsule Networks.** Complex Intell. Syst. 2021, 2021, 1–14.

KANAKAPRABHA, S.; RADHA, D.; SANTHANALAKSHMI, S**. Diabetic Retinopathy Detection Using Deep Learning Models**. Int. 2022, 302, 75–90.

KHAN, A.; KULKARNI, N.; KUMAR, A.; KAMAT, A. **D-CNN and Image Processing Based Approach for Diabetic Retinopathy Classification.** Appl. Inf. Processing. Syst. 2022, 1354, 283–291.

MATH, L.; FATIMA, R. **Adaptive Machine Learning Classification for Diabetic Retinopathy.** Multimed. Tools Appl. 2021, 80, 5173–5186.

SRIDHAR, S.; PRADEEPKANDHASAMY, J.; SINTHUJA, M.; **Sterlin Minish, T.N. Diabetic Retinopathy Detection Using Convolutional Nueral Networks Algorithm.** Mater. Today Proc. 2021.

ARTICLES

DR. OKATAN ALI, KARLIK B, DEMIREZEN F, **Detection of retinopathy diseases using an artificial neural network based on the discrete cosine transform**. January 2009 Neural Network World 19(2):215-221.

K. SIMONYAN AND A. ZISSERMAN. **Very deep convolutional networks for large-scale image recognition.** arXiv preprint arXiv:1409.1556, 2014.

TAJBAKHSH N, SHIN JY, GURUDU SR, HURST RT, KENDALL CB, GOTWAY MB, ET AL. **Convolutional neural networks for medical image analysis: full training or fine tuning?** IEEE Trans Med Imaging. 2016;35(5):1299–312.

RESUME

MOUAD KABBOURI

Artificial Intelligence and Data Science Engineering

As a highly motivated and ambitious graduate student specializing in Artificial Intelligence and Data Science, I am committed to pushing the boundaries of what technology can achieve. Equipped with a deep understanding of theoretical principles and practical applications, I harbor a profound passion for innovation and data-driven solutions. Throughout my academic journey, I have consistently exhibited top-tier analytical skills, adaptability, and a team-centric mindset. My objective is to harness my expertise to contribute to the evolution of advanced AI systems and data analytics solutions that tackle global challenges. Dedicated to lifelong learning and staying updated with the latest in AI research and developments, I am on the lookout for opportunities to implement my knowledge in a progressive and visionary organization.

Education

January 2021 - September 2023

MSc. Artificial Intelligence and Data Science Engineering
Thesis in Machine Learning and Deep Learning Techniques.
ISTANBUL AYDIN University - Turkey

- CGPA: 3.10

September 2017 - June 2020

Bachelor in Computer engineering and telecommunications networks.
Institut supérieure VINCI d'ingénierie informatique et de réseaux de télécommunications, Rabat Morocco.

September 2015 - June 2016

SCIENTIFIC BACCALAUREATE Option Physics. Tamesna, Morocco

Expertise

- Deep Learning Architectures
- Natural Language Processing (NLP)

- Predictive Analytics
- Data Visualization
- Big Data Technologies
- Reinforcement Learning
- Optimization Techniques

Skills

- Data Preprocessing
- Statistical Analysis
- Data Visualization
- Model Development
- Database Management
- Research & Problem Solving
- Version Control
- Soft Skills

Languages

- Arabic
- English
- French