T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DIABETIC RETINOPATHY DETECTION USING META LEARNING AND DEEP LEARNING TECHNIQUES

MASTER'S THESIS

Muhammad Ammar Khan

Department of Software Engineering

Artificial Intelligence and Data Science Program

JUNE, 2023

T.C.

ISTANBUL AYDIN UNIVERSITY

INSTITUTE OF GRADUATE STUDIES

DIABETIC RETINOPATHY DETECTION USING META LEARNING AND DEEP LEARNING TECHNIQUES

MASTER'S THESIS

Muhammad Ammar Khan

(Y2113.011005)

Department of Software Engineering

Artificial Intelligence and Data Science Program

Thesis Advisor: Prof. Dr. Ali Okatan

JUNE, 2023

APPROVAL PAGE

DECLARATION

I hereby declare with the respect that the study "Diabetic Retinopathy Detection using Meta learning and Deep learning techniques", which I submitted as a Master 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. (15/06/2023)

Muhammad Ammar Khan

FOREWORD

I would like to express my gratitude to Dr. Ali Okatan, who served as my adviser for my dissertation, for providing me with all of the necessary advice. I am grateful to him for all of the assistance he has provided during this time. The contributions of his thoughts and support were unquestionably essential, and he made it possible for me to improve my thesis and complete it on time. He has always done his best to guide me in the right direction whenever he felt I was stuck.

In addition, I want to take this opportunity to thank my family and friends for providing me with unwavering support and inspiration during the process of completing my thesis. This success would not have been attainable without the involvement of the individuals I have mentioned. Thank you.

June, 2023 Muhammad Ammar Khan

DIABETIC RETINOPATHY DETECTION USING META LEARNING AND DEEP LEARNING TECHNIQUES

ABSTRACT

In the world of ocular health, diabetic retinopathy is a common condition that, if not recognized and treated promptly, can cause vision loss. This research paper introduces a meta learning stacking approach for the diagnosis and referral of ocular defects. Our approach demonstrates exceptional efficacy in detecting uncommon conditions by utilizing a combination of previously trained convolutional neural networks (CNNs) and stacking meta-learning techniques. The following novel approach improves the accuracy of results while significantly reducing the time required compared to conventional deep learning methods. The method demonstrates the promise of stacking meta learning in addressing data scarcity and improving early diagnosis of sight-threatening diseases by achieving an outstanding accuracy of 93%. Additionally, the solution beats problems brought on by a lack of readily available data which needs to be preprocessed. When compared to other deep learning models frequently used in ocular abnormality detection. These findings underscore the potential impact of our approach as an advanced computer-aided diagnosis tool for ocular anomalies, paving the way for significant advancements in the field. These valuable insights provide a solid foundation for future research, driving innovation and progress in computer-aided diagnosis tools for ocular health.

Keywords: Meta learning; Deep Learning; Retinopathy Detection; Diabetic Retinopathy

META ÖĞRENME VE DERİN ÖĞRENME TEKNİKLERİ KULLANARAK DİYABETİK RETİNOPATİ TESPİTİ

ÖZET

Göz sağlığı dünyasında, diyabetik retinopati, hemen tanınmaz ve tedavi edilmezse görme kaybına neden olabilen yaygın bir durumdur. Bu çalışmada, oküler kusurların teşhisi ve sevki için bir meta öğrenme yığınlama yaklaşımı sunuyoruz. Yaklaşımımız, önceden eğitilmiş evrişimli sinir ağları (CNN'ler) ve istifleme meta-öğrenme tekniklerinin bir kombinasyonunu kullanarak nadir görülen koşulları tespit etmede olağanüstü etkinlik göstermektedir. Bu yeni yaklaşım, geleneksel derin öğrenme yöntemlerine kıyasla gereken süreyi önemli ölçüde azaltırken sonuçların doğruluğunu artırır. Yöntem, veri kıtlığının ele alınmasında ve %93'lük olağanüstü bir doğruluk elde ederek görmeyi tehdit eden hastalıkların erken teşhisini iyileştirmede meta öğrenmeyi yığınlama vaadini gösteriyor. Ek olarak, çözüm, önceden işlenmesi gereken hazır veri eksikliğinden kaynaklanan sorunları yener. Oküler anormallik tespitinde sıklıkla kullanılan diğer derin öğrenme modelleriyle karşılaştırıldığında. Bu bulgular, oküler anomaliler için gelişmiş bir bilgisayar destekli tanı aracı olarak yaklaşımımızın potansiyel etkisinin altını çiziyor ve bu alanda önemli ilerlemelerin önünü açıyor. Bu değerli içgörüler, oküler sağlık için bilgisayar destekli teşhis araçlarında yenilik ve ilerleme sağlayan gelecekteki araştırmalar için sağlam bir temel sağlar.

Anahtar Kelimeler: Meta öğrenme; Derin Öğrenme; Retinopati Tespiti; Diyabetik Retinopati

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

AI	: Artificial Intelligence
DR	: Diabetic Retinopathy
APTOS	: Asia Pacific Tele-Ophthalmology Society
BN	: Batch Normalization
CNN	: Convolutional Neural Network
RNN	: Recurrent Neural Networks
EyePACS	: Eye Picture Archive Communication System
Conv	: Convolution
ConvNet	: Convolutional Network
IRMA	: Intraretinal Microvascular Abnormalities
DME	: Diabetic Macular Edema
PDR	: Proliferative Diabetic Retinopathy
HCI	: Human-Computer Interaction
LMTCNN	: Lightweight Multi-Task Convolutional Neural Network
MBConv	: Mobile Inverted Bottleneck Conv
NB	: Naive Bayes
NIR	: Near Infrared
PCA	: Principal Component Analysis
PRAUC	: Precision-Recall of Area Under the Curve
RBF	: Radial Basis Function
ReLU	: Rectified Linear Unit
RF	: Random Forest
TP	: True Positive
TN	: True Negatives
FP	: False positive
FN	: False negative
TPR	: True positive rate
FPR	: False positive rate
AUC	: Area Under The Curve

TPR	: True positive rate
AUCROC	: Area Under the Curve of Receiver Operating Characteristic
ROC	: Receiver Operating Curve

LIST OF FIGURES

LIST OF TABLES

LIST OF EQUATIONS

I. INTRODUCTION

Diabetic retinopathy (DR) emerges as a significant complication of Diabetes Mellitus, impacting nearly 93 million individuals globally. This condition is characterized by the progressive disruption of retinal vasculature due to chronic hyperglycemia, often resulting in severe visual impairment and the loss of sight among adults within the working age range. Criticality of the timely detection and appropriate intervention of DR cannot be overemphasized, considering its potential to prevent vision loss. Regular eye screening is thus essential for individuals diagnosed with diabetes. However, the traditional approach to DR diagnosis, which involves fundus examination by trained ophthalmologists, poses significant challenges. These challenges include the requirement of specialized expertise, time constraints, and issues pertaining to accessibility, especially in areas that are densely populated or geographically isolated, where individuals may face limited access to specialized healthcare resources.

In light of these challenges, there has been a surge of interest in leveraging Artificial intelligence (AI) and deep learning (DL) technologies to develop computer-assisted diagnostic systems for the purpose of detecting and classifying diabetic retinopathy (DR). These deep learning models, especially those employing deep neural networks, have exhibited promising outcomes in the analysis of fundus images for the automated classification, detection and grading the severity of diabetic retinopathy (DR). They possess the potential to accurately identify retinal lesions and evaluate the progression of DR, thereby providing valuable insights that can facilitate timely medical intervention and treatment. Furthermore, these DL-based models also present the opportunity to bridge the gap in access to specialized eye care services, especially in regions where there is a shortage of ophthalmologists.

Convolutional Neural Networks (CNNs) have emerged as one of the prominent deep learning techniques used in this context, particularly for visual image assessment. CNNs operate by accepting an input image and applying trainable parameters such as priority, weights, and biases. These parameters can be learned and tailored to various sections of the image, enabling differentiation between them. Unlike other classification algorithms, CNNs require significantly less preprocessing, making them a convenient and efficient solution for tasks such as detecting and classifying diabetic retinopathy (DR).

In this research, our primary focus is to offer a comprehensive and thorough analysis. review of the meta learning based approach for DR detection and classification. Our objective is to highlight the advancements, persistent challenges, and potential future trajectories in this rapidly evolving research field. Our aim is to deliver a detailed examination of various aspects of this field, including technical implementations, dataset characteristics, preprocessing techniques, and the performance of different DL models in DR detection and classification.

Furthermore, we seek to objectively evaluate the strengths and weaknesses of the various DL models that have been used in the study of DR, providing insight into their relative merits and potential clinical applicability. An essential component of this analysis is an examination of the ethical considerations and interpretability of these DL models. We emphasize the importance of fairness, transparency, and robustness in these AI-powered systems, given the sensitivity and significance of the tasks they are employed for.

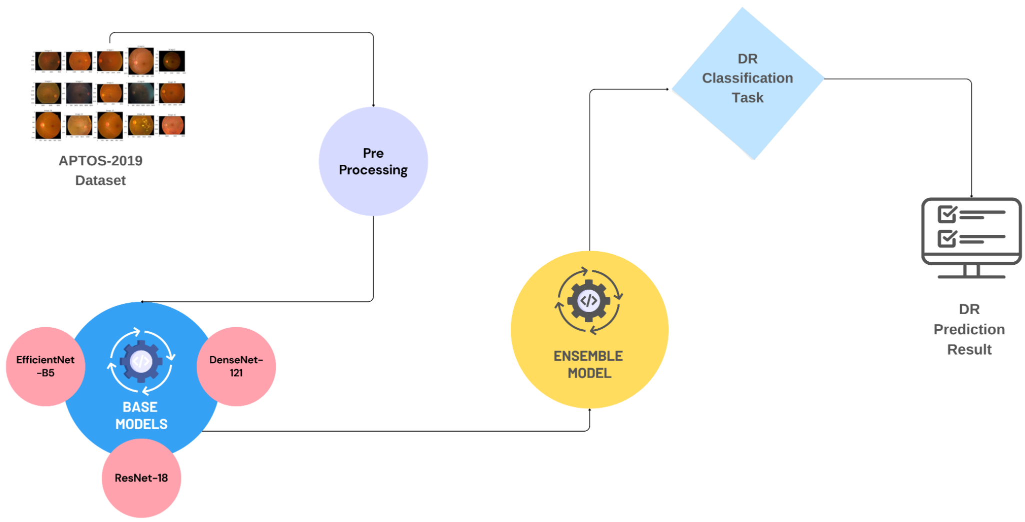

Figure 1 System flow diagram. During the pre-processing stage, we normalized and resized all of the images to a ratio of 224 by 224 pixels. After that, the data were sent to pre-trained base models, and then its output was sent to an ensemble learning model and the output of

the stacking model was used for Diabetic Retinopathy Classification. Metrics scores were evaluated against the ensemble model in the last step.

Deep learning is an artificial intelligence (AI) approach that endeavors to mimic the functionality of the human brain by learning from experience. It is presently capable of performing computer vision tasks and performing state-of-the-art image recognition. In other words, it is an approach that facilitates learning by attempting to uncover the hidden representation of data. Deep learning methods are currently extensively utilized for various tasks, including classification, automated feature extraction, object identification, and so on, because of their high classification accuracy and emphasis on computers learning and developing on their own via computer algorithm analysis.

A. Motivations

The motivations behind this paper are centered around addressing the challenges associated with the classification for the purpose of Diabetic Retinopathy (DR) detection using deep learning techniques. The key motivations include:

Diabetic Retinopathy is a critical eye condition that has the potential to cause vision impairment and even blindness if not detected and treated early. The motivation of this paper is to develop a reliable and accurate method for classifying the severity of DR using deep learning models. By improving the diagnosis process, timely interventions and treatments can be initiated, ultimately improving patient outcomes. Deep learning has demonstrated immense potential in various fields, including medical imaging. The motivation behind this paper is to leverage the capabilities of deep learning models, particularly convolutional neural networks (CNNs), in order to achieve precise classification of Diabetic Retinopathy (DR). Deep learning models possess an inherent capability to automatically learn and capture intricate patterns and representations from image data, making them well-suited for analyzing medical images and detecting subtle abnormalities.

Ensemble learning is a powerful technique that enhances overall performance by combining multiple models. The motivation of this paper is to leverage ensemble learning, along with the stacking approach, to further enhance the classification accuracy of DR. By combining the predictions of multiple base models, the stacking model aims to capture a more comprehensive representation of the data and achieve higher accuracy. Developing a classification model is another motivation from the paper that can generalize well to unseen data and can be easily transferred to other datasets or medical imaging tasks. The robustness and transferability of the proposed approach would enable its application to different datasets or even different medical conditions, providing a framework for efficient and accurate diagnosis across various medical domains.

Overall, the motivations of this paper revolve around improving the diagnosis of DR, leveraging the capabilities of deep learning models, utilizing ensemble learning techniques, and ensuring the generalization and transferability of the proposed approach to enhance its applicability in medical imaging tasks.

B. Aims and Objectives

This research utilizes meta-learning models trained on retinal images for diagnosing Diabetic Retinopathy (DR). The primary objectives are outlined as follows:

- To provide a comprehensive analysis demonstrating the potential of meta-learning and ensemble learning techniques in recognizing Diabetic Retinopathy (DR) using retinal images.
- To apply various ensemble learning strategies, such as stacking, to enhance the predictive accuracy of meta-learning models for DR detection.
- To devise strategies for reducing false positives and negatives, thereby improving the precision and recall of the ensemble meta-learning models.
- To assess the robustness and generalizability of the proposed ensemble meta-learning method using a variety of performance indicators such as F1-score, Accuracy and area under the ROC curve.
- To investigate the potential benefits of modifying the problem from multi-class to binary classification, thereby enhansing the overall performance of ensemble meta learning models in DR detection.

C. Thesis Outline

The thesis has the structure shown below:

- Chapter 2 is a literature review that focuses on reviewing previous work done by authors.
- Chapter 3 focuses on the research and methodology and also discuss some main terminologies
- Chapter 4 presents the implementation details are explained in this chapter, every base model used is explained as well
- Chapter 5 gives the results and the discussion about the research. It also gives information about experimental research with the data.
- In Chapter 5 which is titled as "Conclusions," we explain the effectiveness of using an ensemble and the potential areas for improvement, future work and the implications of the study.

II. BACKGROUND STUDIES AND RELATED WORKS

Gender classification has been extensively researched using a variety of techniques and methodologies, some of the more recent of which are described below:

In this comprehensive review, (Nikos et al., 2021) delve into the application of AI, specifically deep learning and transfer learning methods, for the detection and classification of Diabetic Retinopathy from fundus retina images. They scrutinize the entire detection process, encompassing the choice of datasets, preprocessing techniques, and the development of models for disease diagnosis, grading, and lesion localization. The paper concludes by emphasizing significant insights and proposing future research paths. The goal of the study (Gayathri et al., 2020) is to introduce an innovative system for Diabetic Retinopathy (DR) detection using CNNs commonly employed for feature extraction from retinal fundus images and machine learning classifiers for disease detection. The classifiers, including SVM, AdaBoost, Naive Bayes, Random Forest, and J48 classifiers have been extensively evaluated using the IDRiD, MESSIDOR, and KAGGLE datasets. After evaluating the following results, it indicated that the J48 classifier combined with the novel CNN feature extraction outperforms others in terms of accuracy and Kappa-score. Furthermore, they demonstrated the model's robustness across multiple datasets, reinforcing the potential of this approach in revolutionizing DR diagnosis and contributing towards mitigating vision loss worldwide.

Utilizing image enhancement and estimated hemorrhage location mapping, the technique streamlines traditional screening methods. The study introduces a novel deep learning approach for the detection of early Diabetic Retinopathy (DR), focusing on hemorrhage detection. The research proposes a unique hemorrhage network for classification, benchmarked against well-established models such as LeNet-5, AlexNet, ResNet50, and VGG-16. Despite its relative simplicity, the proposed network demonstrated competitive accuracy, while significantly reducing training time, underscoring the importance of suitable architecture and parameters in developing efficient deep learning models for DR detection. In this study by (H. Jiang et al., 2020), a new approach to diabetic retinopathy (DR) detection is proposed, leveraging deep learning for multi-label classification and lesion detection in DR fundus images. Recognizing the importance of diverse lesions for accurate DR diagnosis, the method not only categorizes DR images, but also localizes various lesions by utilizing the Gradient-weighted Class Activation Mapping (Grad-CAM) technique. Innovative image labeling reduces annotation efforts and enhances efficiency, with diverse lesions considered distinct labels. The model, based on a customized ResNet architecture, was developed using 3228 fundus images across five pre-defined labels. Evaluation on test images revealed a sensitivity with a classification accuracy of 93.9% and a specificity of 94.4% for Diabetic Retinopathy (DR), successfully outlining lesion regions on the images.

A research conducted by S. Mishra (Mishra et al., 2020) the team applied Deep Learning, a subset of Artificial Intelligence, to automate the detection and classification of Diabetic Retinopathy (DR) stages, a prevalent eye disease in diabetic individuals. To streamline this traditionally manual and time-intensive process, the researchers utilized the DenseNet model, training it on a large dataset of around 3662 high-resolution fundus images sourced from Kaggle's APTOS competition. Five stages of DR, labeled 0 through 4, are considered. The study leverages patient fundus eye images as input, with the trained DenseNet architecture extracting image features, and subsequent activation functions producing outputs. The implemented DenseNet architecture achieved an accuracy of 0.96 and a quadratic weighted kappa coefficient score of 0.89 in DR detection. Furthermore, the study presents a comparative analysis of two Convolutional Neural Network architectures, namely VGG16 and DenseNet121.

The study by S. Aziz et al. (Aziz et al., 2023) includes enhancement of image quality and identification of potential hemorrhage locations in the preprocessing phase, with modified gamma correction addressing uneven image brightness levels. Their algorithms employed in this study include Gaussian match filtering, entropy thresholding, and mathematical morphology for candidate location estimation. The proposed hemorrhage network is devised for classifying hemorrhages and is benchmarked against renowned models including LeNet-5, AlexNet, ResNet50, and VGG-16, have been evaluated on two datasets, considering performance metrics such as sensitivity, specificity, precision, and accuracy. Despite being a shallower network, the proposed model demonstrated competitive results while reducing training time. The study concludes with insights on the importance of choosing an appropriate deep model architecture and parameters for optimum performance, as merely increasing the network layers does not necessarily guarantee superior results.

Chandran et al. (Chandran et al., 2016) used a Random Forest approach for the problem in the paper the authors present a novel algorithm that has been developed for the extraction of texture and vesselness features in a patch-wise manner. The Gabor wavelet transform is utilized for the vessel map for extracting these features. The features extracted from image patches are then subjected to classification through a random forest classifier. By leveraging the outputs of this classifier and a rule-based decision system, the images are effectively categorized into stages, namely PDR (Proliferative Diabetic Retinopathy), NPDR (Non-Proliferative Diabetic Retinopathy), and Normal. The utilization of a random forest classifier enhances the accuracy and reliability of the system, making it a valuable tool for diabetic retinopathy screening.

While research (Sarki et al., 2020) focuses on the automated classification of diabetic Retinopathy (DR) using fundus images of the retina. The objective is to develop a system that has the capability to accurately classify both mild and multi-class DED. The study utilizes deep learning techniques, particularly (CNNs), which employs top pretrained models trained on ImageNet. Performance enhancing techniques such as fine tuning, optimization, and contrast enhancement are implemented. The results show that the VGG16 model achieved maximum accuracy for multi-class classification is 88.3%, while for mild multi-class classification it is 85.95%. This automated classification system has the potential to reduce the workload and time required for ophthalmologists in diagnosing DED.

In (Nazir et al., 2021), proposes a method which focuses on the automated system capable of detecting and classifying diabetic retinopathy (DR) and diabetic macular edema (DME) using retinal images. The proposed method consists of two main steps: The process involves dataset preparation and feature extraction, followed by the improvement of a customized deep learning-based CenterNet model specifically designed for eye disease classification. The approach involves generating annotations for suspected samples to precisely locate the regions of interest, and training the CenterNet model using DenseNet-100 for feature extraction and disease lesion localization and classification. The method was evaluated on challenging datasets, achieving high accuracy rates of 97.93% on the APTOS-2019 dataset and 98.10% on the IDRiD dataset. Validation of Cross dataset with benchmark datasets further demonstrated the best performance of the proposed approach, confirming its effectiveness and generalizability. The results highlight the effectiveness of the CenterNet model in localizing and classifying disease lesions, making it a valuable tool for automated DR and DME detection and recognition.

The study by (Zhang et al., 2019) proposes DeepDR, an automated system capable of identification and for the grading of diabetic retinopathy (DR) using fundus images. DeepDR utilizes learning techniques like transfer and ensemble, incorporating state-of-the-art NNs and custom DNNs. The system is developed using a high-quality dataset of medical images related to diabetic retinopathy (DR), which has been carefully labeled by clinical ophthalmologists. The optimal model is constructed by exploring the relationship between the classifiers of the components and labels of the class and evaluating different combinations of classifiers. DeepDR demonstrates satisfactory performance in detecting DR, the proposed approach aims to deliver reproducible and consistent results with high specificity and sensitivity. The study highlights the importance of the numbers of component classifications in achieving optimal model performance, offering valuable insights for ophthalmologists in the diagnostic process of DR.

This study by (Chen et al., 2018) presents a computer guided diagnosis method for (DR) using a DL algorithm. The proposed method capable of diagnosing diabetic retinopathy (DR) and classifying fundus images of retina into five grades. A novel pre processing algorithm is introduced to enhance image quality and improve uniformity in this study., as well as a transfer learning approach to improve performance. The system is evaluated on a test set, demonstrating promising results. The method addresses the challenges of early DR detection and the need for expert diagnosis, offering potential benefits in the diagnosis and management of DR.

The study by (Enkvetchakul et al., 2022) focuses on the rapid diagnosis of diabetic retinopathy using artificial intelligence techniques. Meta-learning

Convolutional Neural Networks (CNNs) are employed to recognize diabetic retinopathy from retinal images. Data resampling methods and data augmentation techniques are applied to improve training performance. The study compares two ensemble learning methods and finds that the meta-learner method achieves the highest accuracy. The proposed approach exhibits promising results, indicating its effectiveness and potential for further development with an accuracy of 86.32%. The results show the impact of using AI for efficient and accurate diabetic retinopathy diagnosis.

Table 1 presents a summary of recent studies on retinopathy detection and their corresponding accuracies. It provides an overview of various algorithms that have been utilized in recent research. While most contributions in the field employ multiple models and report results for each model, the table only includes the best result reported in each paper. The datasets used in these studies may vary, and authors often present multiple findings from different datasets within a single publication, but only the better result for the publication is listed here. The table offers insights into the commonly used models for diabetic retinopathy detection and highlights the models that perform well in this domain.

Table 1 A summary of diabetic retinopathy researches

Study	Classifier	Result/Accuracy
(Jiang et al., 2020)	ResNet50	94%
(Chandran et al., 2016)	Random Forest	89%
(Aziz et al., 2023)	LeNet-5, AlexNet, ResNet50, VGG-16, HemNet	96%
(Chandran et al., 2019)	Inception, Resnet and Inception-Resnet	94%
(Mishra et al., 2015)	VGG16 and DenseNet121+QWK+ImageNet,	96%

(Tao et al., 2019)	ResNet, VGG16, GoogLeNet, DenseNet	82%
(Zhang et al., 2019)	Resnet50 , Xception, InceptionV3, IncepresV2	97%
(Zhang et al., 2018)	SVM	94%
(H. Chen et al., 2018)	Inceptionv3, Xception, DenseNet	80%
(Ardiyanto et al., 2017)	ResNet, ResNet-20, Deep-DR-Net + SGD	95%
(X. Wang et al., 2018)	AlexNet, VGG16, InceptionNet V3	63%
(Rishab & Theodore, 2017)	Custom CNN & Decision Tree	94%
(S. Qummar et al., 2019)	Ensemble	86%
(Enkvetchakul et al., 2022)	Meta Learning	86%

Diabetic retinopathy detection poses a complex challenge for computer-based systems, particularly due to factors such as disease progression and image quality. While DL techniques have shown promising outcomes in this particular domain, indicating the potential and effectiveness of the approaches employed, the use of a meta learning approach for DR detection remains largely unexplored. The following study proposes a novel meta learning technique to improve the performance and generalizability of deep learning networks for DR detection. By leveraging meta learning techniques, our approach aims to improve the accuracy of the model and the robustness across diverse datasets.

III. RESEARCH AND METHODOLOGY

In this section, we discuss various aspects related to diabetic retinopathy detection. We provide an introduction to diabetic retinopathy, highlight the dataset used for training and evaluation, explore the application of deep learning techniques, specifically Convolutional Neural Networks (CNNs), has been utilized to enhance the detection process, discuss the utilization of transfer learning, and touch upon the integration of meta-learning approaches. These topics collectively contribute to a comprehensive understanding of diabetic retinopathy detection, encompassing dataset selection, the performance of DL, the effectiveness of CNNs, the leveraging of pre-trained models using transfer learning and the potential benefits of meta-learning for improved detection performance.

A. Diabetic Retinopathy

In the initial stages of DR specific manifestations on the retina can be observed, indicating disease progression. Microaneurysms, resulting from pericytes' degeneration and loss, cause the dilation of capillary walls. Intraretinal hemorrhages occur when capillary or microaneurysm walls rupture. Non-proliferative diabetic retinopathy (NPDR) includes additional abnormalities including soft and hard exudates, (IRMA), venous beading, and or reduplication. IRMAs, characterized as huge-caliber tortuous vessels, often emerge in areas of ischemia as a response to attempted vascular remodeling. Neovascularization, distinguished by the growth of new retinal vessels due to ischemia, is a notable manifestation, indicating (PDR). Furthermore, (DME), characterized by fluid accumulation in the macula, can further impair vision. Timely management and detection of these various manifestations are critical and important to prevent vision loss and provide appropriate treatment interventions. Figure 2 shows us representation of different manifestations and complications associated with diabetic retinopathy.

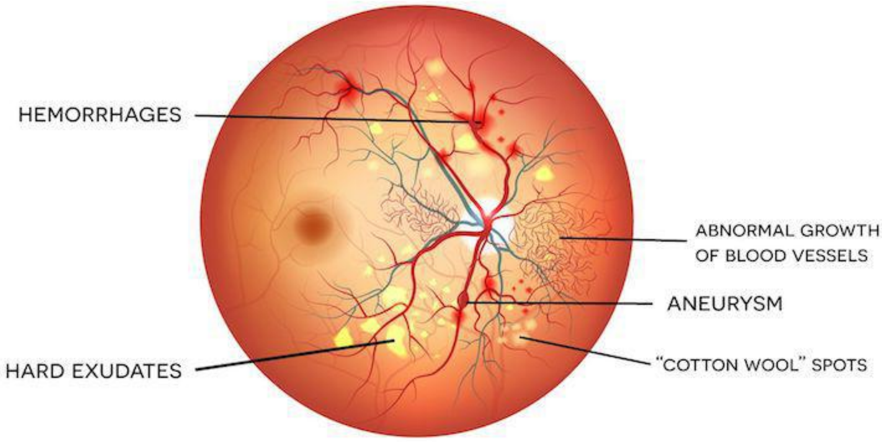

Figure 2 A visualization of the number of serious issues caused by diabetic retinopathy. A person with diabetic retinopathy may experience the following complications that affect the blood vessels supplying nutrients to the retina.

Diabetic retinopathy at any particular stage, the occurrence of (DME) can significantly contribute to visual impairment. Notable abnormalities associated with DME include exudates within a distance of one disc diameter from the center of the fovea, presence of exudates within the macula and retinal thickening within a disc diameter of the fovea's center, and the presence of microaneurysms or hemorrhages within the same region. Different grading protocols have been devised to evaluate the clinical severity of DR. Although the (ETDRS) grading system is widely recognized as the gold standard in assessing diabetic retinopathy, it has proven difficult to implement in everyday clinical practice. Alternative severity scales have been suggested as a result in various nations to improve patient screening and streamline communication among healthcare professionals. One such scale is the The International Clinical Diabetic Retinopathy Disease Severity Scale is considered a globally recognized standard for classifying the severity of diabetic retinopathy as shown in Table 2, which classifies diabetic retinopathy into five severity levels based on specific findings observed during dilated ophthalmoscopy. These findings include The presence of microaneurysms, intraretinal hemorrhages, venous beading, and (IRMA), and neovascularization.

Table 2 The table below shows The International Clinical Diabetic Retinopathy and DME Disease Severity Scales are extensively utilized in clinical practice to evaluate the severity and progression of DR and DME.

Stage	Dilated Ophthalmoscopy Observable Findings	Severity
I	No abnormalities	No DR
II	Micro-aneurysms only	Mild non-proliferative DR
III	Any of the following: - micro-aneurysms - retinal dot and blot haemorrhages - hard exudates or cotton wool spots No signs of severe non-proliferative diabetic retinopathy	Moderate non-proliferative DR
IV	Any of the following: - more than 20 intra-retinal hemorrhages in each of 4 quadrants - definite venous beading in 2 or more quadrants - prominent intra-retinal microvascular abnormality (IRMA) in 1 or more quadrants No signs of proliferative retinopathy	Severe non-proliferative DR
V	One or both of the following: - Neovascularization - Vitreous/pre-retinal hemorrhage	Proliferative DR

B. Dataset Analyzation

The dataset for the Kaggle APTOS 2019 competition was sourced significantly from the Aravind Eye Hospital in rural India. The dataset comprises a total of 5,590 retinal images specifically collected for automated Diabetic Retinopathy Detection. This dataset is considered the third-largest of its kind, providing valuable data for training and evaluating diagnostic models. However, it must be acknowledged that the dataset exhibits imbalance of the class, particularly in Severe NPDR class, which is represented by only 193 images.

To ensure the dataset's diversity and representativeness, technicians from the Aravind Eye Hospital visited various rural areas of the country to capture the images. This approach introduces variations in camera settings across multiple centers, similar to the Kaggle EyePACS dataset. It must be acknowledged that as these retinal pictures were gathered in real-world clinical settings, they may contain noise, such as artifacts, focus issues, and differences in exposure levels (over or under).

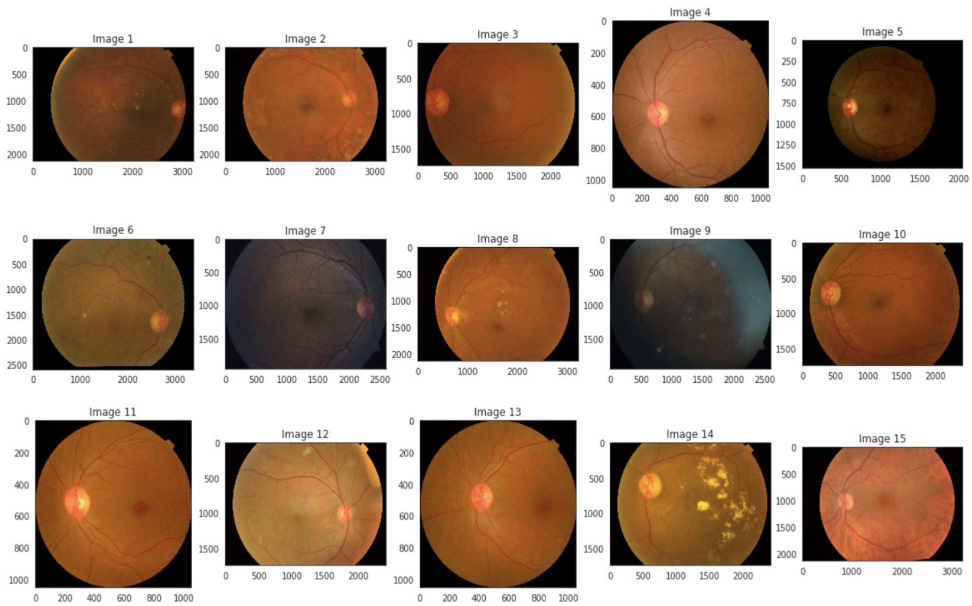

Figure 3 Here are 15 randomly selected retinal images from the APTOS-2019 dataset, which includes a collection of images for automated Diabetic Retinopathy Detection.

The labels associated with the dataset adhere to the widely recognized Diabetic Retinopathy and DME Disease Severity Scales as shown in the table 2. The dataset provides images categorized into five main classes: 0 grades as No DR, 1 grades as Mild, 2 grades as Moderate, 3 grades as Severe, and 4 grades as Proliferative DR . These categories serve as a reference for evaluating the severity and advancement of DR and DME. The classes 1, 3, and 4 each have less than half the amount of data as class 2. For the training set, the dataset comprises 3662 images and for the validation set, the dataset comprises 1928 images. The dataset has five classes, with class 0 having the highest number of samples at 1805, followed by class 2 with 999 samples. Class 1, class 4, and class 3 have 370, 295, and 193 samples, respectively. The figure 3 illustrating the label class distribution clearly visualizes the unbalanced nature of the dataset, with a notable disparity in the count of samples across various classes.

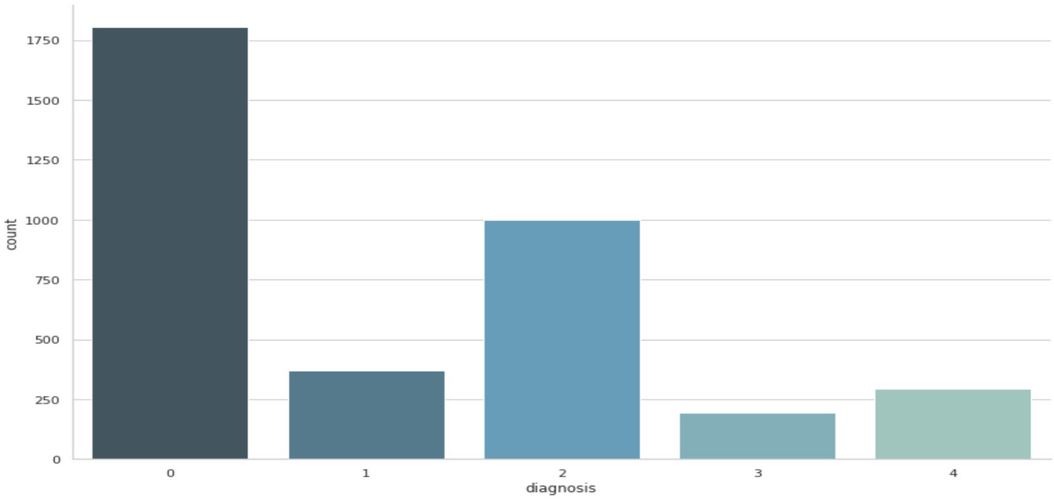

Figure 4 The figure illustrates the imbalanced label class distribution in the dataset, with class 0 being the most prevalent and classes 1, 3, and 4 having fewer samples.

C. Terminologies

In the field of AI and ML, several key terminologies like DL, CNN, Transfer Learning, and Meta Learning play a top role in advancing the capabilities of models and algorithms. Each term refers to a specific concept or technique within the broader domain of machine learning.

1. Deep Learning

Deep learning has brought about a transformative impact in diverse domains, encompassing computer vision, natural language processing, speech recognition, and robotics. Deep learning has made remarkable breakthroughs in numerous tasks, including image classification, object detection, machine translation, and voice synthesis. One of the prominent advantages of deep learning is its capability to effectively handle large and high-dimensional datasets, allowing it to capture intricate patterns and dependencies that might be difficult for traditional machine learning algorithms to discern.

Deep neural networks, the fundamental building blocks of deep learning, comprise multiple layers of interconnected neurons. Each neuron performs a weighted computation on its inputs and passes the result through a non-linear activation function. The network's depth allows it to acquire hierarchical representations, where lower layers capture low-level features, while higher layers capture more abstract and complex features. This hierarchical representation learning empowers deep learning models to achieve best performances in a wide range of tasks, often surpassing human-level performance in specific domains.

Deep learning models have also benefited from architectural innovations, including (CNNs) for image analysis, (RNNs) for sequential data processing, and transformers for NL understanding. The mentioned architectures, combined with techniques like transfer learning, regularization, and ensemble learning, have further improved the performance and generalization capabilities exhibited by DL models.

2. Convolutional Neural Network

CNNs are a powerful class of deep learning models specifically designed for the tasks of image processing. They are inspired by the structure and functionality of the visual cortex present in the human brain. CNNs are adept at capturing spatial dependencies and extracting meaningful features from images. This is accomplished by employing convolutional layers that apply filters to the input data, allowing the network to learn local patterns and spatial relationships. By stacking multiple convolutional layers, CNNs can progressively learn more complex representations of the input data, enabling them to effectively handle intricate visual tasks like classifications, detections, and segmentations of the images.

The notable advantage of CNNs is their capability to automatically learn relevant features from raw image data. Unlike traditional machine learning approaches that often necessitate manual feature engineering, CNNs eliminate the need for such time-consuming processes and have the potential to capture comprehensive information. In contrast, CNNs learn feature mapping from the data directly, allowing them to adapt and discover high-level features that are most discriminative for the task at hand. This capability is particularly valuable in computer vision, where images contain a vast amount of information that may not be easily discernible to human observers. CNNs excel at automatically extracting relevant visual features, enabling them to acquire high performance in various image-related tasks.

3. Transfer Learning

Transfer learning describes the method of using the knowledge gained by using pre-trained models on one task to enhance performance on a related but distinct task.

The ultimate goal of the transfer learning is to move the learnt parameters or representations from the source to the target. In transfer learning, a pre-trained model collects common patterns and features that are helpful for a variety of tasks and is often trained on a big dataset. The pre-trained model is applied to new data more successfully and possibly performs better than when learned from the start by adjusting or tuning it for the desired task. Transfer learning becomes very beneficial when intended tasks have few labeled data points or when the target task and source task have similar properties. It aims to leverage knowledge from pre-trained models to improve performance on a related task.

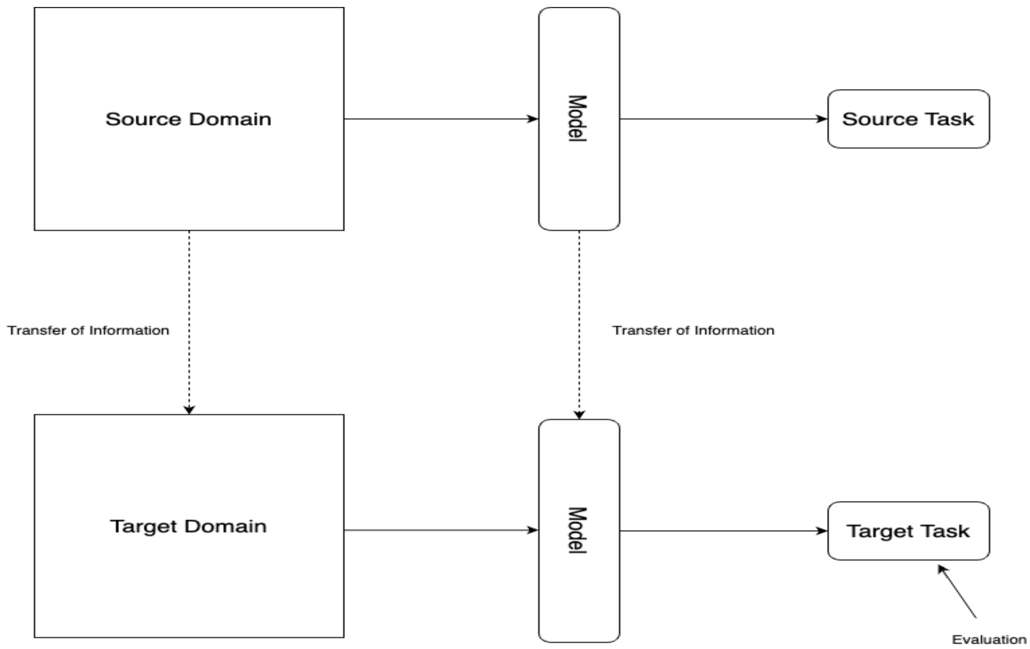

Figure 5 The basic structure of transfer learning involves transferring the learned representations from a source task to a new target task, allowing for improved performance and generalization.

4. Meta Learning

Meta learning, a technique also referred as "learning to learn," focuses on developing algorithms and techniques that enable ML models to acquire the ability to learn and adapt quickly from a limited set of training examples. Meta learning involves training models on multiple related tasks or datasets to learn higher-level knowledge or strategies that can be applied to new, unseen tasks. By learning how to learn, meta learning algorithms aim to improve the generalization and adaptation capabilities of ML models. These techniques often involve training models with meta-objectives, such as optimizing the learning process itself or learning to dynamically adjust model parameters based on task-specific characteristics.

By combining transfer learning and meta learning, researchers and practitioners aim to improve the efficiency, performance and generalization of ML models. Transfer learning provides a way to leverage existing knowledge and representations, while meta learning equips models with the ability to quickly adapt and generalize to new tasks. Both approaches have demonstrated significant advancements in various domains, including computer vision, natural language processing, and reinforcement learning tasks. These techniques continue to drive innovation in the field of ML, enabling models to tackle complex real-world problems with limited labeled data and enhance their overall learning capabilities.

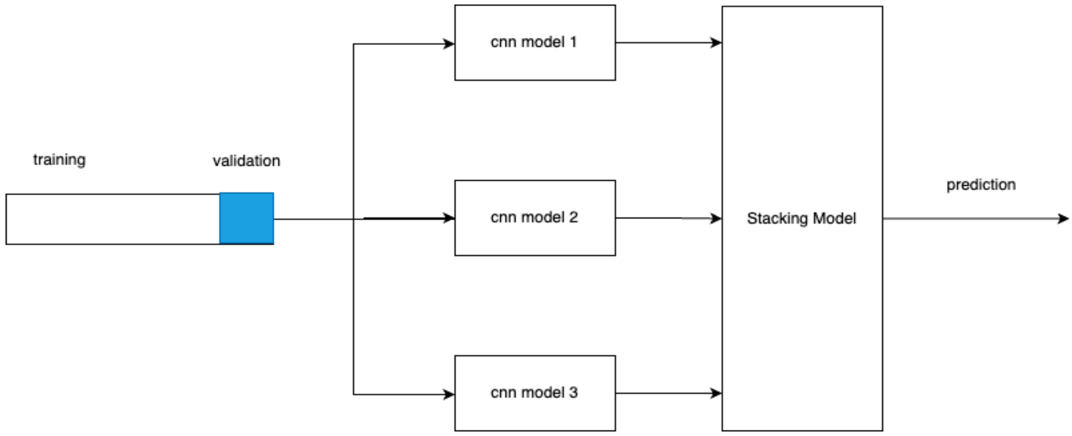

Figure 6 The figure illustrates the basic structure of meta-learning, showcasing the process of learning to learn for improved task adaptation and generalization.

IV. IMPLEMENTATION DETAILS

In this research, the implementation details of diabetic retinopathy (DR) detection and classification were divided into several phases. The first phase focused on setting up the base models, including EfficientNet, ResNet, and DenseNet. These models were chosen for their effectiveness in image classification tasks and their ability to extract relevant features from retinal images. Each base model was trained independently to learn the specific patterns and characteristics associated with DR. In the subsequent phases, ensemble and stacking techniques were employed to further improve the models' performance. Ensemble learning involved combining the predictions of multiple base models, while stacking utilized the predictions as input features for a meta-classifier. A thorough training and validation split was performed to ensure accurate evaluation and prevent overfitting. Through these steps, the research aimed to achieve accurate detection and classification of DR, facilitating timely diagnosis and treatment for affected individuals.

A. Base Models

For training the stacking model in ensemble learning we first need to train and prepare the base models. After the dataset has been trained on the model, we use the outputs of each of these base models as the input of the stacking model. We have used only three base models for this research.

1. EfficientNet

EfficientNet models have been developed with the objective of achieving more reliable and accurate approaches in computer vision and deep learning. These models achieve this by scaling depth, width, and resolution while effectively reducing the model size. The EfficientNet family consists of 8 models, ranging from B0 to B7, where the number of parameters does not increase with the model number, but the accuracy significantly improves. In contrast to other CNN models that employ ReLU as the activation function, EfficientNet introduces a novel activation function known as Swish, which combines linear and sigmoid activations. The architecture of EfficientNet is characterized by the inverted bottleneck MBConv, which utilizes direct connections and channel compression, resulting in fewer connected channels compared to expansion layers. EfficientNet models demonstrate higher accuracy and efficiency compared to existing CNN models, with EfficientNet-B7 achieving state-of-the-art performance with 84.4% top-1 and 97.1% top-5 accuracy on ImageNet dataset.

The architecture of EfficientNet also incorporates in-depth separable convolutions, reducing computation by a factor of k when the kernel size is k^2. To ensure consistent scaling of depth, width, and resolution, the compound coefficient φ is employed following the principles of compound scaling.

$$depth,\ d = \alpha\varphi$$
$$width, w = \beta\varphi$$
$$resolution, r = \gamma\varphi$$
(1)

EfficientNet models aim to achieve high accuracy while still using computational resources efficiently. Efficient Net-B5 has been used as the first base model. The scalable architecture of EfficientNet models makes it simple to change the model size in accordance with the resources at accessibility or your specific requirements.

$$\alpha\geq 1, \beta\geq 1, \gamma\geq 1 \quad (2)$$

In the equation 2 α, β, γ are constants, by using grid search these can be observed, the number of resources available for model scaling is controlled by φ which is a user-defined coefficient, on the other hand, α, β, γ determine shows the width, depth, and resolution of the architecture, respectively.

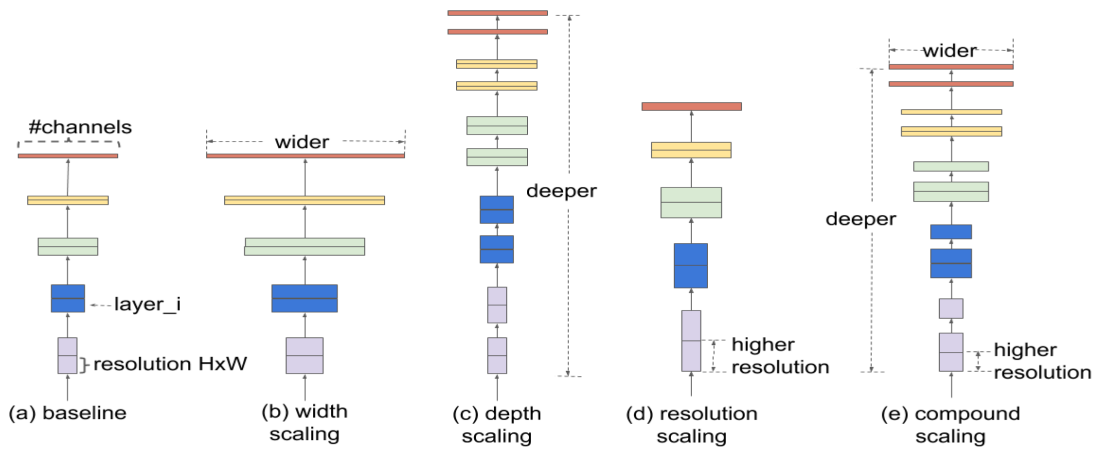

Figure 7 The figure illustrates the basic architecture of EfficientNet.

EfficientNet-B5 is the efficient DL model that combines accuracy and computational efficiency. As part of the EfficientNet family, it is designed to scale the width, depth and resolution of the architecture to achieve optimal performance. With its larger size, EfficientNet-B5 captures intricate patterns and features in data while maintaining computational efficiency. It has been trained on large-scale datasets and utilizes techniques like depthwise separable convolutions to reduce complexity. EfficientNet-B5 has achieved impressive results in various computer vision tasks, making it best suitable for applications which need accurate visual recognition.

With its larger size compared to earlier versions, EfficientNet-B5 exhibits a higher capacity to capture intricate details and learn complex representations from data. It has been trained on extensive datasets, enabling it to generalize well to various visual recognition tasks. The model incorporates depthwise separable convolutions, which separate the spatial and channel-wise transformations, reducing computational complexity without sacrificing accuracy. The reason for using efficientNet-B5 is that it's a powerful deep learning model that strikes a balance between accuracy and computational efficiency. Its scalable architecture, combined with innovative design choices, allows it to achieve remarkable performance across different visual recognition tasks. By leveraging EfficientNet-B5, researchers and practitioners can benefit from its high accuracy and efficient resource utilization, thus making it a well suited option for use in the research of DR.

2. ResNet

The ResNet architecture introduced Residual Blocks as a solution to the challenge of vanishing/exploding gradients in deep neural networks. By incorporating skip connections, ResNet enables layer activations to be directly connected to subsequent layers, creating residual blocks. This approach allows the network to learn residual mappings, focusing on the difference between the predicted and true outputs rather than directly modeling the desired mapping. The skip connections enable the network to handle the degradation of performance that can occur with increasing depth, making it possible to train extremely deep neural networks without encountering gradient-related issues.

Implementations of ResNet are readily available in popular deep learning frameworks such as Keras, which offers pre-trained ResNet models with different depths and variations. The ResNet models available in Keras Applications include ResNet50, ResNet101, ResNet152, ResNet50V2, ResNet101V2, and ResNet152V2. These models differ in terms of the number of layers and the specific architectural variations. Additionally, the V2 versions of ResNet incorporate batch normalization before each weight layer, providing further improvements in training stability and performance. The availability of these pre-trained models simplifies the integration and deployment of ResNet architectures in various computer vision applications, enabling researchers and practitioners to leverage the power of ResNet for their specific tasks.

The Residual Network, commonly known as ResNet, was originally introduced in 2015 by Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun in their influential research paper focused on computer vision (Tongfeng et al., 2020). These Networks are a sort of classic NN that serves as the basis for a wide variety of computer vision applications. ResNet pioneered the notion of skip connections in order to solve the issue of vanishing gradients. This is achieved by letting the gradient flow in a different shortcut direction and allowing the model to learn an identity mapping. This ensures that the higher layer performs as well as possible at the lower layer, and possibly even better. In the absence of the skip connection, the input "Y" undergoes a multiplication by the layer's weights, followed by incorporating a bias term. The activation function "f()" is subsequently applied to the output of the layer, yielding "H(y)" Equation (3).

$$H(y) = f(wy+b) \quad \text{or} \quad H(y) = f(y) \tag{3}$$

In summary, the inclusion of residual blocks in ResNet simplifies the learning process for layers to capture identity functions. This enhances the effectiveness of deep neural networks with a larger number of layers while reducing errors. The skip connections enable the addition of outputs from stacked layers to the outputs of preceding layers, enabling the training of significantly deeper networks than was previously feasible. However, after the introduction of the skip connection approach, the output has been changed from "H(y)" to Equation (4).

$$H(y)=f(y)+y \tag{4}$$

In addition, whether using a convolutional layer or a pooling layer, the input dimension and the output dimension may be different from one another. As a consequence of this, the issue may be resolved by using the strategies of padding a zero by making use of the skip connection in order to enlarge its dimensions and by attaching 1x1 convolutional layers to the input in order to fulfill the required dimensions. In this particular scenario, the outcome is represented by Equation (5), in which "w1" is an additional parameter that has been included.

$$H(y)=f(y)+w1.y \tag{5}$$

In ResNet models make use of residual connections to help solve the issue of training with vanishing gradients. ResNet-18 is a variant of the Residual Neural Network (ResNet) architecture that consists of 18 layers. ResNet-18 ResNet has gained widespread adoption in computer vision tasks, particularly in areas like image classification and object detection. Its effectiveness in handling complex visual data has made it a popular choice in these domains. Despite its relatively compact size, ResNet-18 exhibits remarkable accuracy and generalization capabilities, making it a popular choice for various applications in the field of computer vision. Resnet-18 is used for the following implementation. This makes it possible to train deeper networks efficiently and raises the performance of the model as a whole. In comparison to deeper variants, ResNet18 is a relatively light network, which speeds up training and uses less computing power.

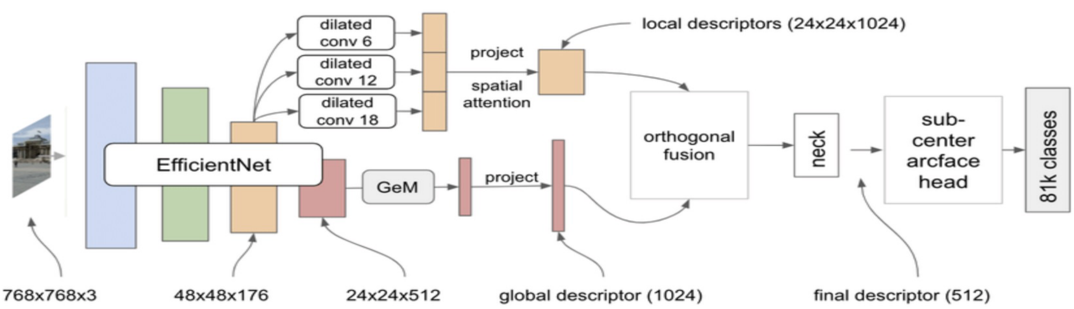

Figure 8 The figure illustrates the basic architecture of Resnet.

3. DenseNet

DenseNet is a CNN architecture which differs from traditional feed-forward CNNs. DenseNet architecture is composed of Dense Blocks and Transition Layers, establishing direct connections between each layer and every other layer in the network (Huang et al., 2017). DenseNet-121 and DenseNet-169 are two popular variants of DenseNet, characterized by the number of layers and parameters.

In a traditional CNN, each layer receives input from the previous layer and generates an output feature map. However, as the network becomes deeper, the issue of vanishing gradients arises, hindering effective training. This occurs when information is lost or 'disappears' as it traverses through numerous layers. The vanishing gradient problem poses limitations on the network's ability to effectively learn and achieve high performance.

DenseNets address the vanishing gradient problem by ensuring dense connections between layers, promoting feature reuse and propagation. Unlike traditional CNNs, DenseNets require fewer parameters as they eliminate the need to learn unnecessary feature mappings. In a DenseNet, every single layer is directly connected to all other layers, resulting in a total of $L(L+1)/2$ direct connections in the network for L number of layers.

The DenseNet architecture improves gradient flow, facilitates information flow across layers, and effectively utilizes network parameters. The specific configurations of DenseNet, including DenseNet-121 and DenseNet-169, are summarized in Table 3.

Table 3 Overall DenseNet architectures for ImageNet (Huang et al., 2017).

Layers	Output Size	DenseNet-121($k = 32$)	DenseNet-169($k = 32$)	DenseNet-201($k = 32$)	DenseNet-161($k = 48$)
Convolution	112×112	7×7 conv, stride 2			
Pooling	56×56	3×3 max pool, stride 2			
Dense Block (1)	56×56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$
Transition Layer (1)	56×56	1×1 conv			
	28×28	2×2 average pool, stride 2			
Dense Block (2)	28×28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$
Transition Layer (2)	28×28	1×1 conv			
	14×14	2×2 average pool, stride 2			
Dense Block (3)	14×14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 36$
Transition Layer (3)	14×14	1×1 conv			
	7×7	2×2 average pool, stride 2			
Dense Block (4)	7×7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$
Classification Layer	1×1	7×7 global average pool			
		1000D fully-connected, softmax			

Connectivity, DenseBlocks, Growth Rate, and Bottleneck layers are the component elements of the DenseNet architecture.

a. Dense Connectivity

DenseNets utilize a unique connectivity pattern where the feature maps of previous layers are concatenated and used as inputs in subsequent layers, resulting in reduced parameters and enabling feature reuse. Each layer in DenseNet receives the concatenation of feature maps [z0, z1, z2...] from all preceding layers as input, denoted as [z0, z1,..., z_{l-1}] in Equation (6). This concatenation simplifies the implementation of H_l, combining multiple inputs into a single tensor for further processing.

$$z_l = H_l\left(\left[z_0, z_1, \ldots, z_{l-1}\right]\right) \tag{6}$$

b. Dense Blocks

To address situations where the size of feature maps varies, DenseNets incorporate Dense Blocks. These blocks maintain a constant size of feature maps while varying the number of filters between each pair of maps. Transition Layers, located between the blocks, play a crucial role in down-sampling by reducing the number of channels by half. Equation (1) defines H_l as a composite function in each layer, encompassing rectified linear unit (ReLU), batch normalization (BN), and convolution (Conv) operations, ensuring efficient information processing within the DenseNet architecture.

c. Growth Rate

The growth rate, denoted by "N", determines the number of new features added by each layer to the existing global state, resulting in an expansion of the feature map size. It governs the amount of information contributed to each layer of the network. Equation (7) illustrates that if each function H_l generates "n" feature maps, the lth layer must include input feature-maps, where "n_0" represents the number of channels in the input layer. DenseNets distinguish themselves from traditional network architectures by accommodating remarkably thin layers while maintaining effective information flow throughout the network.

$$n_l = n_0 + n * (l - 1) \tag{7}$$

d. Bottleneck layers

To address the issue of a large number of inputs in deeper layers, a bottleneck layer with a 1x1 convolution is introduced before each 3x3 convolution. This bottleneck layer helps enhance the efficiency and speed of computations. DenseNets stand out from traditional CNN or ResNet architectures by evaluating fewer parameters, promoting feature reuse, and constructing more compact models. These characteristics contribute to their state-of-the-art performance and efficiency across various datasets. The dense connectivity of DenseNets plays a key role in achieving these results. The schematic layout of DenseNet is depicted in Figure 9.

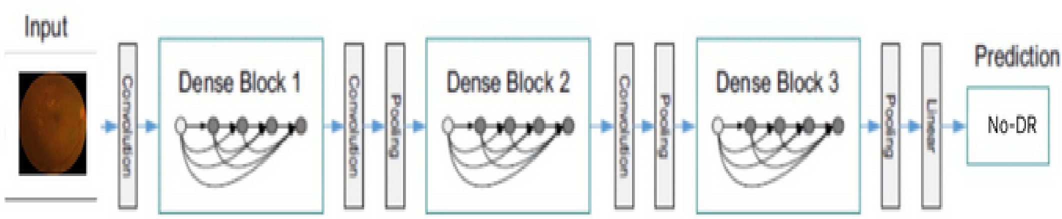

Figure 9 The figure illustrates a deep DenseNet architecture consisting of 3 dense blocks. The layers that connect two adjacent blocks are referred to as transition layers, which employ convolution and pooling operations to update the size of feature maps.

B. Model Ensemble and Stacking

In the task of predicting diabetic retinopathy, an ensemble approach was employed using multiple deep learning models. ResNet18, DenseNet121, and EfficientNet-B5 were among the models selected for the ensemble. These models, originally trained on the ImageNet dataset, were fine-tuned by adapting their final layers to match the classes in the specific diabetic retinopathy dataset used in the study. This customization allowed the models to leverage their pre-trained knowledge and potentially improve their performance in diabetic retinopathy detection.

To further enhance the predictive performance, a stacking approach was implemented. Stacking involves training an ensemble model that combines predictions from multiple base models to make final predictions. In this study, the outputs from ResNet18, DenseNet121, and EfficientNet-B5 were concatenated to form the input for the stacking model. By leveraging the collective knowledge and predictions of these base models, the stacking model aimed to make more accurate and robust predictions for diabetic retinopathy.

The ensemble of multiple deep learning models and the stacking technique offered several advantages. First, it allowed for leveraging the strengths of each base model and compensating for their individual weaknesses. Second, the ensemble approach helped reduce the risk of overfitting and improved the generalization capability of the predictive model. Lastly, the stacking method provided a mechanism to learn higher-level representations and capture complex interactions between features, potentially leading to improved performance in diabetic retinopathy classification. Through this ensemble approach and the utilization of ResNet18, DenseNet121, and EfficientNet-B5 models, the study aimed to create a robust and accurate system for diabetic retinopathy prediction. The results obtained from this ensemble approach could provide valuable insights and contribute to the development of more reliable diagnostic tools for early detection of DR.

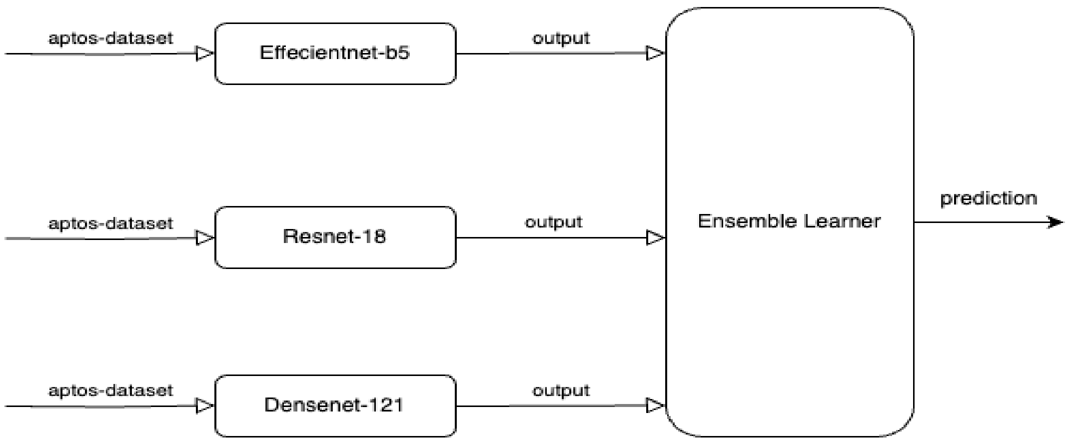

Figure 10 The figure illustrates an ensemble stacking of the base models which are Efficientnet-b5, Resnet-18 and Densenet-121.

The stacking model used in this study was a straightforward linear model. A 15-element vector is created by concatenating the outputs of the base models, each of which generates a prediction vector of length 5 (corresponding to the 5 classes in the dataset). The model is able to learn the best combination of the predictions from the basic model by passing this vector through a linear layer with five outputs. By employing this stacking strategy, the model is able to fully capitalize on the advantages of each base model, possibly leading to a more reliable and precise predictive model. The stacking model is trained using the Adam optimizer and a common loss function (in our instance, Cross Entropy Loss).

C. Dataset Pre-processing

In the preprocessing stage of the APTOS-2019 dataset, a transformation pipeline is applied to the retinal images. This pipeline includes several transformations to prepare the images for analysis. First, we checked for any outliers but couldn't find any duplicate or none pictures. The following images were then resized to fixed 224x224 pixels images. This resizing step ensures that all images have the same dimensions, facilitating consistent processing of the dataset. Next, these images were converted to tensors, which are numerical representations and are best suitable for DL models. This conversion enables efficient computation and integration with deep learning frameworks. Finally, the pixel values of the images are normalized. The normalization process adjusts the pixel values to standardize the data with a mean of (0.485, 0.456, 0.406) and a standard deviation of (0.229, 0.224, 0.225). The following normalization ensures that the pixel values are standardized and helps in mitigating the impact of varying image characteristics. By applying these preprocessing steps, the APTOS-2019 dataset is prepared in a consistent and standardized format, which is essential for training and evaluating deep learning models on the dataset.

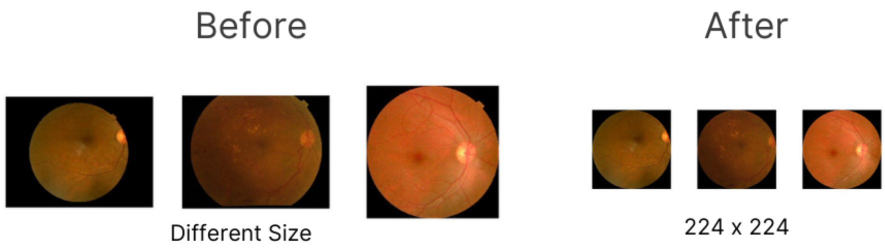

Figure 11 The figure depicts the preprocessing stage of the images in the APTOS-2019 Kaggle dataset.

D. Training and Validation Split

To guarantee that the models could be properly trained and evaluated in this research, the dataset was divided in training subset and validation subset. By dividing the dataset in this form, the models can be trained on one part (i.e training set) of the data, and their performance can then be assessed on a second subset of the data (the validation set), which they were not exposed to during training. In a real-world application, this helps in estimating how well the trained models will respond to new data. This was done using pandas' sample function, which takes a sample of data at random, according to the provided code in which we have chosen the 80% of data for the training set since the parameter was set to 0.8.

The random_state parameter was set to a specific value to ensure the reproducibility of the results. It ensures that each time the code is executed, the random selection of data will be the same. After choosing the training set, the validation set was made by removing its indices from the entire dataset, thereby erasing the 20% of data that wasn't part of the training set. This type of split was used for this study on the detection of diabetic retinopathy since it is frequently used in ML and DL projects to provide a balanced approach to training and testing models.

Table 4 A numerical representation of training and test images of dataset

Images	Training	Validation
Randomly selected Images	80%	20%
Total Images 3662	2930	732

V. RESULTS AND DISCUSSION

A. Training Process

Standard deep learning techniques were used to handle the training of the models, including the base models and the stacking model. The Adam optimization technique was used for optimizing the parameters of the models. This adaptive learning rate method uses little memory and is computationally effective. Due to its effectiveness, it has been widely used in deep learning. A common approach for multi-class problems with classification, the Cross-Entropy Loss, was used to train the models. The variation between the expected probability and the actual labels is measured by the Cross-Entropy Loss. As a result, the model is recommended to output low probability for the incorrect classes and high probabilities for the correct classes.

The main models were initially trained separately. The stacking model was then trained using the Adam optimization algorithm and Cross-Entropy Loss, and its data inputs were concatenated and used in this process. The stacking model learns to correct for errors made by the base models by training it to base its predictions on their outputs, which will enhance the accuracy of all predictions.

B. Base Model Performance

In the case of the individual base models, they were evaluated using the accuracy metric, which is the proportion of correct predictions out of the total predictions. For the EfficientNet-B5 model, the accuracy on the validation set was approximately 82.24%. The second model, ResNet18, achieved an accuracy of about 78.96% on the validation set. The third model, DenseNet121, performed comparably to the first model, with an accuracy of about 81.28% on the validation set. In general, these results suggest that all three base models were able to learn useful representations from the training data and generalize well to unseen data in the validation set. However, there were variations in their performances, and this further justifies the use of model ensemble and stacking techniques to leverage their complementary strengths and improve the overall system's performance.

Table 5 Results from the three base models before ensemble learning

EfficientNet-B5	**ResNet18**	**DenseNet121**

82.24%	78.96%	81.28%

C. Stacking Model

The base models (Efficientnet-b5, Resnet-18, Densenet-121) predictions were used to train the stacking model. The stacking model's input was a combination of the predictions from these models. The typical method of using the original features as input (in this case, the image data) is different from this. The stacking model was trained using a standard supervised learning procedure. Binary Cross-Entropy with Logits Loss, a loss function appropriate for binary classification issues, was employed. Adam was used for training the model as its the widely used optimizer algorithm because of its effectiveness. Every 20 epochs, the loss was recorded during the 500 epochs of training in order to track its progress. The model parameters were preserved for later usage after the training.

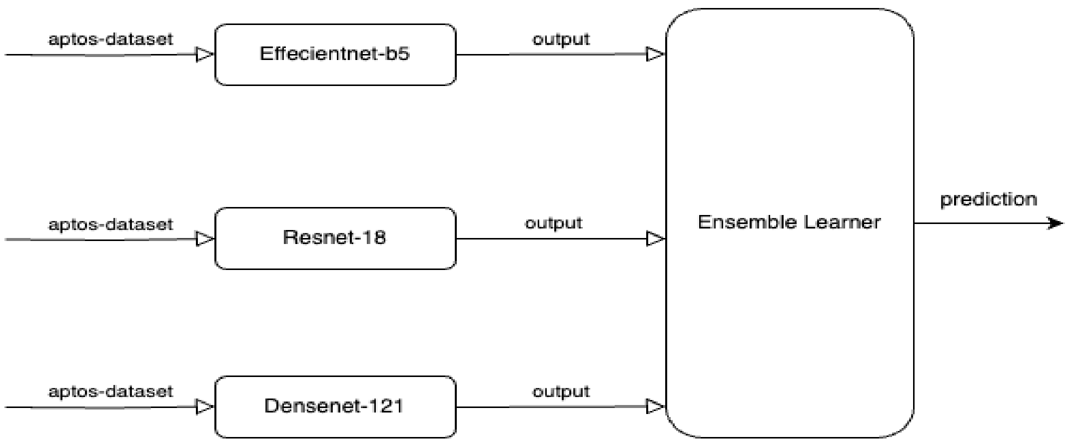

Figure 12 The figure depicts the three base models as inputs for the ensemble model and then the ensemble model will do the prediction based on it.

1. Accuracy

Accuracy can be computed by considering the positive and negative classes. TP (True Positives) represents the number of instances correctly classified as the observed class, TN (True Negatives) represents the number of instances correctly classified as the remaining classes, FP (False Positives) represents the number of misclassified instances as the remaining classes, and FN (False Negatives) represents the number of misclassified instances as the observed class.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

2. Recall/Sensitivity

Recall, also referred to as sensitivity or true positive rate, is a performance metric commonly utilized in classification tasks. It quantifies the model's ability to accurately identify positive instances among all the actual positive instances in a dataset. Mathematically, recall is calculated by dividing the true positives (TP) by the sum of true positives and false negatives (FN).

$$Sensitivity = \frac{TP}{TP+FN} \quad (9)$$

If the value of high recall is big it indicates that the model is successfully capturing a significant proportion of the positive instances and has a low false negative rate. It is a useful metric when the focus is on minimizing false negatives, such as in medical diagnosis or detecting rare events.

3. Specificity

Specificity, also referred to as the true negative rate, is a performance metric commonly employed in binary classification tasks. It assesses the model's capability to accurately identify negative instances among all the actual negative instances in a dataset. Mathematically, specificity is calculated by dividing the true negatives (TN) by the sum of true negatives and false positives (FP):

$$Specificity = \frac{TN}{TN+FP} \quad (10)$$

A high specificity value indicates that the model is successfully capturing a significant proportion of the negative instances and has a low false positive rate. It is a valuable metric when the emphasis is on reducing false positives. Specificity is the complement of the false positive rate (1 - false positive rate) and provides insights into the model's ability to accurately classify negative instances.

4. Precision

Precision, also referred to as positive predictive value, is a performance metric commonly employed in binary classification tasks. It assesses the accuracy of the positive predictions made by a model. Mathematically, precision is calculated by dividing the true positives (TP) by the sum of true positives and false positives (FP).

$$Precision = \frac{TP}{TP + FP} \quad (11)$$

A high precision value indicates that the model is making accurate positive predictions with a low rate of false positives. It is particularly valuable when minimizing false positive errors is a priority, such as in medical diagnostics or credit fraud detection. Precision provides insights into the model's ability to correctly classify positive instances, especially when the cost of false positives is significant. However, precision does not consider false negatives and may not fully capture the model's ability to correctly identify all positive instances in the dataset.

5. F1-Score

The F1 score is a metric that evaluates a model's performance in binary classification tasks by taking into account both precision and recall. It provides a balanced assessment when dealing with imbalanced datasets or when both false positives and false negatives are equally significant. The F1 score is calculated as the harmonic mean of precision and recall, ensuring that both metrics contribute equally to the overall score:

$$F1 \ = 2\ x\ \frac{precision\ x\ recall}{precision + recall} \qquad (12)$$

A high F1 score indicates that the model achieves both high precision and high recall, signifying accurate positive predictions and correct identification of a substantial portion of positive instances in the dataset. The F1 score ranges from 0 to 1, with 1 being the optimal score.

The F1 score is a valuable metric when there is a trade-off between precision and recall, providing a consolidated assessment of a model's overall performance in binary classification tasks. It is particularly useful when considering both false positives and false negatives in the evaluation process.

6. Receiver Operating Curve (ROC)

The Receiver Operating Characteristic (ROC) curve is a graphical representation of a binary classification model's performance. It is constructed by plotting the true positive rate (TPR), also known as sensitivity or recall, against the false positive rate (FPR) at various classification thresholds. The TPR represents the ratio of true positive instances correctly classified as positive, while the FPR represents the ratio of negative instances incorrectly classified as positive. The ROC curve provides valuable insights into the model's discriminatory capabilities between the positive and negative classes.

The ROC curve is a valuable tool for evaluating the trade-off between the true positive rate and the false positive rate. It allows us to assess the overall classification performance of a model by analyzing the shape of the curve and its distance from the diagonal line, which represents random guessing. The area under the ROC curve (AUC-ROC) is a commonly used metric to measure the discriminative power of a model. An AUC-ROC score of 1 indicates a perfect classifier, while a score of 0.5 suggests a model that performs no better than random guessing. The ROC curve facilitates the comparison of different models or parameter settings within the same model, enabling the selection of an appropriate threshold based on the desired balance between sensitivity and specificity.

7. Area Under The Curve (AUC)

The Area Under the Curve (AUC) is an important metric for assessing the performance of a binary classification model. It specifically refers to the area under the Receiver Operating Characteristic (ROC) curve, which plots the true positive rate (TPR) against the false positive rate (FPR) at various classification thresholds. The AUC quantifies the model's overall discriminative power and its ability to differentiate between positive and negative instances.

The AUC score ranges from 0 to 1, where a score of 1 represents a perfect classifier that achieves a TPR of 1 (sensitivity) while maintaining an FPR of 0 (specificity). On the other hand, a random classifier would have an AUC of 0.5, indicating no discriminative power. Generally, a higher AUC score indicates better performance in classifying positive and negative instances.

The AUC is widely used in evaluating classification models, particularly in scenarios involving class imbalance or unequal costs of false positives and false negatives. It provides a comprehensive measure of the model's ability to rank instances correctly and serves as a useful tool for comparing different models or different parameter settings within the same model.

Metrics were used to evaluate the stacking model's performance on the validation set. On the validation images, the stacking model's accuracy was 92.76%. This indicates that for approximately 93% of the photos in the validation set, the stacking model provided accurate predictions. The precision and recall score, known as the F1 score, was approximately 0.91. This shows that a decent balance between accurately detecting positive cases (recall) and reducing false positives (precision) was maintained by the stacking model.

Table 6 Showing the metrics scores for the ensemble stacking model

Matrics	**f1**	**accuracy**	**precision**	**recall**	**roc_auc**
Score	0.92	0.90	0.90	0.91	0.92

Additionally, the precision score was approximately 0.90, indicating that the model correctly predicted positive cases 90% of the time. Recall was approximately 0.92, indicating that 92% of the positive cases in the validation set were properly identified by the model. Finally, the model's ROC-AUC score was close to 0.93, showing that it is very capable of differentiating between the classes. A high score on the ROC-AUC measure denotes a strong model in binary classification problems. These metrics indicate the effectiveness of the chosen ensemble and stacking technique overall, indicating that the stacking model performed well on the validation set.

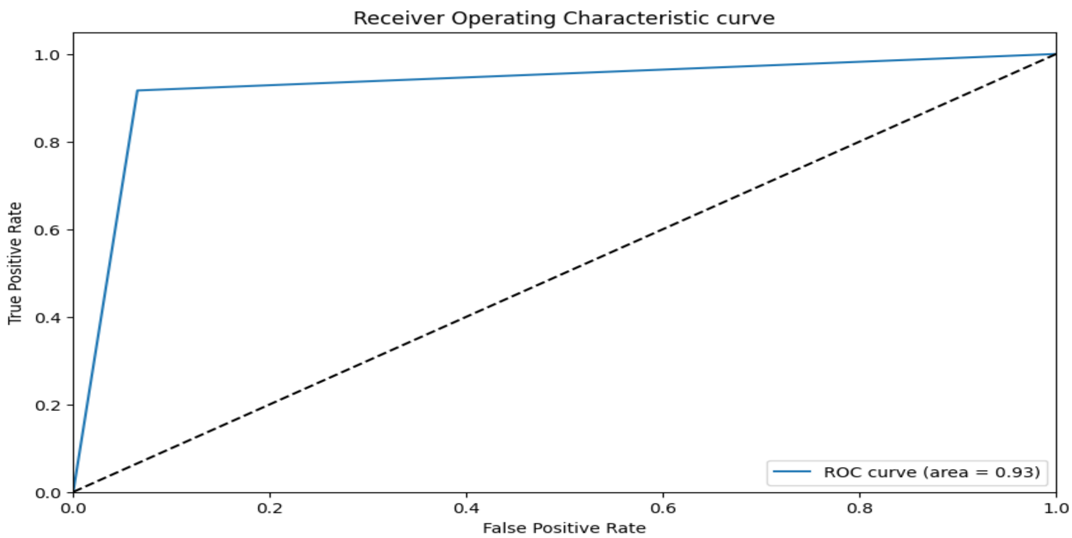

Figure 13 The figure showing the (ROC) curve area of the ensemble stacking model.

D. Few-Shot Learning with APTOS-2019

We used few-shot learning as an experimental strategy with the APTOS-2019 dataset. This method seeks to develop models that can efficiently learn knowledge from a limited amount of data. Due to the APTOS-2019 dataset's unbalanced class distribution, the results, however, were not satisfactory. Even after changing the dataset according to the few shot model requirements by putting all the classes in the different folders and augmenting and adding more pictures to the data with less data still didn't resolve the problem. An unbalanced dataset can have a major impact on the effectiveness of a machine learning model, especially when utilizing a few-shot learning strategy that mainly relies on a tiny sample of data.

E. Using Meta Learning Approach

Moving forward, we employed the meta learning approach on the APTOS-2019 dataset for multi-class classification of diabetic retinopathy. The meta learning approach leverages the ensemble of multiple deep learning models to improve the classification performance. By combining the outputs of different base models, a stacking model was created to make predictions based on the collective knowledge of the base models. This approach allows for enhanced accuracy and robustness in classifying the retinal images into their respective categories. The performance of the stacking model on validation images was as follows:

Table 7 Showing the overall metrics scores for the ensemble stacking model against the classification DR.

Matrics	**f1**	**accuracy**	**precision**	**recall**
Score	0.833	0.833	0.835	0.833

The (ROC) curve analysis, particularly the Area Under the Curve (AUC) metric, provides an evaluation of the overall performance and discriminative power of the meta learning-based approach in diabetic retinopathy classification. You can see from Table 8 that some classes showed better performance while others didn't. Class 0 and 2 which had more images in the dataset showed good performance compared to other classes. The results suggest an acceptable level of performance, but there is room for improvement.

Table 8 The ROC-AUC scores were varying significantly across different classes.

Classes	**O**	**I**	**2**	**3**	**4**
Score	0.978	0.726	0.874	0.709	0.716

After evaluating the multi-class model's performance, we modified the strategy for binary classification. This modification was developed in light of the difficulties that multi-class classification presents when dealing with imbalanced data. As a result, the findings significantly improved and the accuracy of the stacking model exceeded 90%. The experiments have demonstrated that although meta-learning and few-shot learning algorithms can be used to diagnose diabetic retinopathy, careful consideration of the class distribution in the dataset is necessary to ensure the best model performance.

```
Accuracy of the stacking model on the validation images: 0.9275956284153005
F1 score of the stacking model on the validation images: 0.9087779690189328
Precision of the stacking model on the validation images: 0.9010238907849829
Recall of the stacking model on the validation images: 0.9166666666666666
```

Figure 14 The figure showing the Model Performance for binary classification. Accuracy: 92, F1-score: 90, Precision: 90, Recall: 91

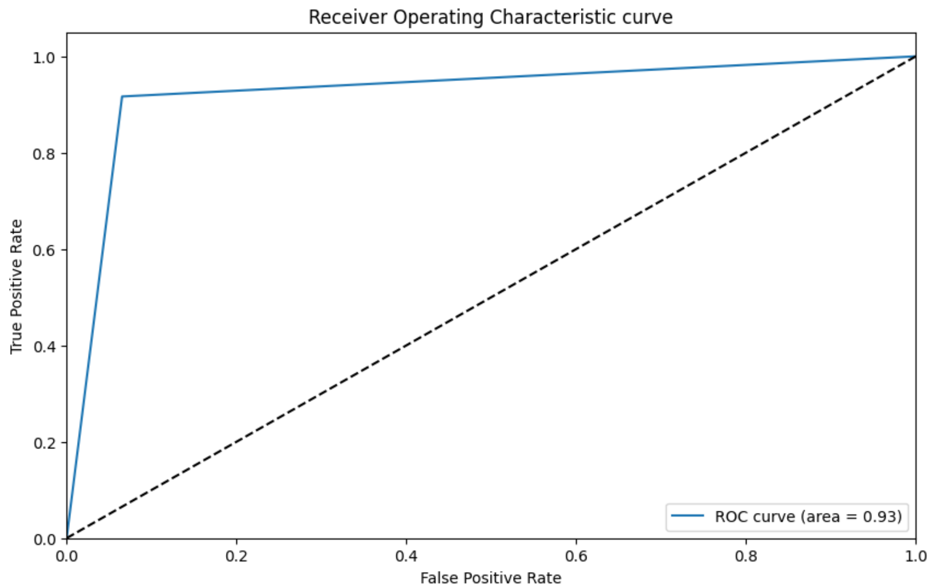

Figure 15 The figure showing the ROC area of Model Performance for binary classification which is 93%.

F. Summary of Findings

This study employed deep learning techniques to classify the severity of Diabetic Retinopathy, a serious eye condition common in diabetic patients. Specifically, the study used an ensemble of pretrained convolutional neural networks (EfficientNet-B5, ResNet18, and DenseNet121) and a stacking approach to improve model performance. The base models achieved individual validation accuracies around 78% to 82%. However, the stacking model achieved a validation accuracy of approximately 93%, showcasing the effectiveness of the ensemble and stacking method. The stacking model also achieved impressive scores on other performance metrics, including precision, recall, F1-score, and ROC-AUC score, indicating its robustness in classifying the DR condition. These findings suggest that ensemble learning and stacking methods can be powerful tools in the context of medical image classification tasks, particularly when the goal is to leverage the strengths of multiple models to improve overall performance.

G. Potential Improvements and Future Work

Despite the encouraging outcomes achieved thus far, several potential areas of improvement and directions for future work. Fine-tuning the base models on the specific task could potentially improve their individual performances and, by extension, the performance of the ensemble model. Additionally, exploring different architectures for the stacking model could lead to improved performance. Hyperparameter tuning, such as adjusting the learning rate or the number of epochs, could also yield better results. Another promising direction for future work could involve expanding the ensemble to include more diverse base models. The current study utilized models based on CNNs; including models based on different architectures, such as transformer models, could potentially increase the robustness of the ensemble.

H. Implications of the Study

This study has several implications, especially for the application of DL in the medical field. First off, it shows how DL models could be used to help with the identification of diseases like diabetic retinopathy. Accurate early detection can significantly enhance a patient's prognosis, resulting in better medical outcomes. Additionally, the application of ensemble and stacking methods demonstrates how we can capitalize on the complementing qualities of diverse models to enhance overall performance, a strategy that can be used in a variety of other healthcare situations. The study additionally emphasizes the value of ongoing model validation and evaluation to make sure the model works effectively on both training and real-world data. This is essential in the medical field because misdiagnosis may be quite expensive and the stakes are so high.

VI. CONCLUSION

In conclusion, this study showcased the effectiveness of using an ensemble of pretrained CNNs and a stacking approach for classifying the severity of Diabetic Retinopathy. The stacking model, consisting of EfficientNet-B5, ResNet18, and DenseNet121, achieved a validation accuracy of approximately 93%, surpassing the individual performances of the base models. The robustness of the stacking model was demonstrated by its high precision, recall, F1-score, and ROC-AUC score. These findings highlight the power of ensemble learning and stacking methods in medical image classification tasks, particularly for leveraging the strengths of multiple models to improve overall performance.

Moving forward, there are several areas for potential improvement and future work. Fine-tuning the base models on the specific task, exploring different architectures for the stacking model, and conducting hyperparameter tuning could further enhance the performance of the ensemble. Additionally, expanding the ensemble to include models based on different architectures could increase its robustness. These avenues of research hold promise for achieving even higher accuracy in classifying diabetic retinopathy.

The implications of this study extend beyond diabetic retinopathy classification. It highlights the potential of DL models in aiding disease identification and improving patient outcomes through accurate early detection. The application of ensemble and stacking methods in the healthcare domain offers a valuable strategy for leveraging the strengths of diverse models. Ongoing model validation and evaluation are crucial in the medical field to ensure effective performance on both training and real-world data. By continually refining and validating these models, we can mitigate the risks associated with misdiagnosis and ultimately improve patient care.

VII. REFERENCES

BOOKS

Dr. med. Florian Lang (Prof.) Diabetic Retinopathy. In: Lang, F. (eds) **Encyclopedia of Molecular Mechanisms of Disease. Springer, Berlin, Heidelberg. (2009)**. https://doi.org/10.1007/978-3-540-29676-8_9136

ARTICLES

H. JIANG et al., "A Multi-Label Deep Learning Model with Interpretable Grad-CAM for Diabetic Retinopathy Classification," **42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 2020**, 1560-1563 https://doi.org/10.1109/EMBC44109.2020.9175884

GAYATHRI S., VARUN P. Gopi, P. Palanisamy, "A lightweight CNN for Diabetic Retinopathy classification from fundus images" , **Biomedical Signal Processing and Control, Volume 62, 2020**, 102115, ISSN 1746-8094 https://doi.org/10.1016/j.bspc.2020.102115

NIKOS Tsiknakis, DIMITRIS Theodoropoulos, GEORGIOS Manikis, EMMANOUIL Ktistakis, OURANIA Boutsora, ALEXA Berto, FABIO Scarpa, ALBERTO Scarpa, DIMITRIOS I. Fotiadis, KOSTAS Marias, "Deep learning for diabetic retinopathy detection and classification based on fundus images". **Computers in Biology and Medicine, Volume 135, 2021**, 104599, ISSN 0010-4825 https://doi.org/10.1016/j.compbiomed.2021.104599

H. Jiang, K. Yang, M. Gao, D. Zhang, H. Ma and W. Qian, "An Interpretable Ensemble Deep Learning Model for Diabetic Retinopathy Disease Classification," **2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany,**

2019, pp. 2045-2048,
https://doi.org/10.1109/EMBC.2019.8857160.

AZIZ, T., Charoenlarpnopparut, C. & Mahapakulchai, S. Deep learning-based hemorrhage detection for diabetic retinopathy screening. **Sci Rep 13, 1479 (2023)**
https://doi.org/10.1038/s41598-023-28680-3

S. Roychowdhury, D. D. Koozekanani and K. K. Parhi, "Automated detection of neovascularization for proliferative diabetic retinopathy screening," **38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Orlando, FL, USA, 2016**, pp. 1300-1303,
https://doi.org/10.1109/EMBC.2016.7590945

S. Mishra, S. Hanchate and Z. Saquib, "Diabetic Retinopathy Detection using Deep Learning," **2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE), Bengaluru, India, 2020**, pp. 515-520
https://doi.org/10.1109/ICSTCEE49637.2020.9277506

A. Chandran, K. K. Nisha and S. Vineetha, "Computer aided approach for proliferative diabetic retinopathy detection in color retinal images," **International Conference on Next Generation Intelligent Systems (ICNGIS), Kottayam, India, 2016**, pp. 1-6,
https://doi.org/10.1109/ICNGIS.2016.7854014

S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson and M. Goldbaum, "Detection of blood vessels in retinal images using two-dimensional matched filters," in IEEE Transactions on Medical Imaging, vol. 8, no. 3, pp. 263-269, Sept. 1989,
https://doi.org/10.1109/42.34715

SARKI R, AHMED K, WANG H, ZHANG Y. Automated detection of mild and multi-class diabetic eye diseases using deep learning. **Health Inf Sci Syst. 2020**
PMID: 33088488
https://doi.org/10.1007/s13755-020-00125-5

T. Nazir, M. Nawaz, J. Rashid, R. Mahum, M. Masood, A. Mehmood, F. Ali, J. Kim, H.-Y. Kwon and A. Hussain, “Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model", **2021; 21(16):5283** https://doi.org/10.3390/s21165283

TAO Li, YINGQI Gao, KAI Wang, SONG Guo, HANRUO Liu, HONG Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening", **Information Sciences, Volume 501, 2019**, Pages 511-522, ISSN 0020-0255
https://doi.org/10.1016/j.ins.2019.06.011

WEI Zhang, JEI Zhong, SHIJUN Yang, ZHENTAO Gao, JUNJIE Hu, YUANYUAN Chen, ZHANG Yi, "Automated identification and grading system of diabetic retinopathy using deep neural networks", **Knowledge-Based Systems, Volume 175, 2019,** Pages 12-25, ISSN 0950-7051
https://doi.org/10.1016/j.knosys.2019.03.016

OKATAN, A., KARLIK, B., & YAGMUR, F. D. (2009). Detection of retinopathy diseases using artificial neural network based on discrete cosine transform. **Neural Network World**, 19(2), 215.

H. CHEN, X. ZENG, Y. LUO and W. Ye, "Detection of Diabetic Retinopathy using Deep Neural Network," **2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), Shanghai, China, 2018**, pp. 1-5
https://doi.org/10.1109/ICDSP.2018.8631882.

I. Ardiyanto, H. A. Nugroho and R. L. B. Buana, "Deep learning-based Diabetic Retinopathy assessment on embedded system," **2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea (South), 2017**, pp. 1760-1763,
https://doi.org/10.1109/EMBC.2017.8037184.

X. Wang, Y. Lu, Y. Wang and W. -B. Chen, "Diabetic Retinopathy Stage Classification Using Convolutional Neural Networks," **2018 IEEE International Conference on Information Reuse and Integration (IRI), Salt Lake City, UT, USA, 2018**, pp. 465-471
https://doi.org/10.1109/IRI.2018.00074

RISHAB Gargeya, THEODORE Leng "Automated identification of diabetic retinopathy using deep learning Ophthalmology", **124 (7) (jul 2017)**, pp.

962-969, https://doi.org/10.1016/j.ophtha.2017.02.008

S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," **in IEEE Access, vol. 7, pp. 150530-150539, 2019,** https://doi.org/10.1109/ACCESS.2019.2947484.

ENKVETCHAKUL, PREM & SURINTA, OLARIK & NOPPITAK, SANGDAOW. (2022). "Effective Data Resampling and Meta-learning Convolutional Neural Networks for Diabetic Retinopathy recognition". **ICIC Express Letters, Part B: Applications. 13**. 939-948 https://doi.org/10.24507/icicelb.13.09.939

TONGFENG Sun, SHIFEI Ding, LILI Guo, "Low-degree term first in ResNet, its variants and the whole neural network family", **Neural Networks, Volume 148, 2022**, Pages 155-165, ISSN 0893-6080. https://doi.org/10.1016/j.neunet.2022.01.012

ELECTRONIC SOURCES

Ophthalmology Physicians & Surgeons, PC , **Eye diseases Service**. https://www.eyeops.com/contents/our-services/eye-diseases/diabetic-retinopathy

Kaggle, Competitions, Aptos-2019 blindness detection dataset. https://www.kaggle.com/c/aptos2019-blindness-detection

Medium, Computer Vision Article, Resnet Architecture Explained. https://medium.com/@siddheshb008/resnet-architecture-explained-47309ea9283d

RESUME

Name Surname: Muhammad Ammar KHAN

EDUCATION:

- **Bachelor**: 2019, COMSATS University, Faculty of Computer Sciences, Department of Software Engineering
- **M.S**: 2023, Istanbul Aydin University, Faculty of Computer Sciences, Department of Software Engineering, Artificial Intelligence and data science (English)(Thesis) Program

PROFESSIONAL EXPERIENCE:

- **Senior Full Stack Developer**
  Lithium Tech | Islamabad, Pakistan Jul, 2O21 – Present

- **Full Stack Developer**
  Xypher | Islamabad, Pakistan Apr, 2O21 – Jul, 2021

- **Backend Developer**
  VisionX | Islamabad, Pakistan Nov, 2O2O – Apr, 2021

- **Backend Developer**
  Iplex | Islamabad, Pakistan Jul, 2O2O – Nov, 2O2O

- **Backend Developer**
  Oudinc | Islamabad, Pakistan Sep, 2019 – Jul, 2O2O

Projects:

- **Physeo**
  A website like udemy for online dental courses.
  Lithium Tech

- **Woo**
  A social media app for short videos competition.
  Lithium Tech

- **LoadEx**
  Keeping track of live location of all the vehicles of the largest cargo company of Pakistan. Xypher

- **Mailroom**
  React Website and mobile app for package delivery and tracking.
  VisionX

- **Beveg**
  BeVeg is the world's vegan certification website.
  Iplex

- **Food Apps**
  An online website for food order from user account and selling from owner account.
  Oudinc

- **Artist Station**
  App where an artist can sell his/her products or services and the user can book for the service and can purchase the products online.
  Oudinc

- **Locum Plus**

Website for doctors, medical equipment, clinics services and job postings.
Oudinc

PUBLICATIONS FROM DISSERTATION, PRESENTATIONS AND PATENTS:

- MUHAMMAD AMMAR KHAN, ALI OKATAN. (2023). **Diabetic Retinopathy Detection Using Meta Learning and Deep Learning Techniques "EURAS Journal of Engineering and Applied Sciences" Vol 3 Issue 2 journal**

OTHER PUBLICATIONS, PRESENTATIONS AND PATENTS:

- NONE