CHAPTER 1

INTRODUCTION

Diabetes Mellitus (DM), an important health concern affecting the general population marked by persistent hyperglycemia, is increasing rapidly in both developed and developing countries. It leads to DR, which is a critical complication with the potential for causing permanent blindness in its later stages. DR is most often due to blood vessel dysfunction of the retina, which is responsible for providing necessary nutrients to the retinal light-sensitive tissue [1].

According to the World Health Organization (WHO), 87.6 million patients were diagnosed with DR in 2019, which is expected to be 115.1 million by 2030 [2]. In addition, the International Diabetes Federation's report in 2021 shows that the prevalence of diabetes is 537 million, and it is predicted to rise to 643 million by 2030 and 783 million by 2045 [3]. It is also particularly common in specific patient populations, affecting up to 40% of people with type 2 diabetes and 86% of people with type 1 diabetes. Several clinical trials have demonstrated that early treatment and effective control of blood sugar levels are pivotal in lowering the risk of DR progression [4]. Hence, in this research to prevent vision loss, DR stage classification is done using a hybrid model. Chapter 1 introduces DR, emphasizing its clinical importance and the critical role of early detection. This chapter outlines the motivation for applying DL to DR classification. It presents the specific objectives of this thesis, which focus on improving classification accuracy, minimizing overfitting, and enhancing computational efficiency.

1.1 DIABETIC RETINOPATHY OVERVIEW

Lesions in DR represent abnormal changes in retinal blood vessels and the adjacent tissues, which occur under persistent high blood sugar, and the lesions are important indices of the development of the disease. Knowledge of lesion types is essential to characterise and diagnose the severity of DR, as the three central lesions related to DR are Microaneurysms (MAs), Haemorrhages (HEMs), and Exudates (EXs), and each sheds light on different aspects of retinal damage.

MAs are the earliest visible DR lesions, interpreted as focalised ballooning of the capillaries in the retina. These MAs are observed when the vessel walls weaken and blood leaks, appearing as small, circular red spots on the retina. While MAs themselves are not toxic, their appearance warns of the development of DR, acting as an early alarm of their subsequent threat to vision.

HEMs are another typical lesion identified in DR, resulting from a breach of fragile blood vessels. There are two varieties of HEMs: dot-and-blot HEMs and flame-shaped HEMs. The HEM dots or blots are small and round, whereas the HEM flame-shaped are larger, elongated, and located in the nerve fibre layer. Both forms of HEMs are a sign of retinal vessel injury and cause vision loss if not addressed.

EXs are the deposits of proteins and lipids in the retina due to leakage from blood vessels. Hard EXs appear as yellowish-white homeostatic material, typically found in the macula, which is essential for acute vision. Soft EXs (cotton wool spots) are caused by retinal ischaemia and appear as white, fluffy patches on the retina. EXs commonly indicate more progressed DR stages and visual damage, if not treated correctly [5].

These lesions aid in diagnosing DR and help inform the clinician about the type of treatment (laser therapy, Anti-Vascular Endothelial Growth Factor (VEGF) injections) needed to prevent further damage to the retina. Recently, the DL based models for automated DR detection have become popular, and they distinguish and categorise these lesions accurately, allowing for relatively timely and more accurate diagnoses. The most fundamental task in the classification of DR is to classify and grade the disease for severity according to the International Clinical DR (ICDR) Disease Severity Scale.

DR is generally categorized into Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR). NPDR has three stages of progression (mild, moderate, and severe), with PDR representing the most advanced stage based on the degree of retinal blood vessel damage. Patients may remain asymptomatic or relatively free of symptoms with MAs in the mild disease stage.

The moderate stage involves retinal HEMs and hard EXs, causing slight vision issues. The severe stage is characterised by extensive blood vessel blockage, leading to significant retinal ischemia. In the very severe stage, the risk of transitioning to PDR becomes high. PDR, the most advanced stage, is marked by abnormal new blood vessel growth, i.e., Neovascularisation (NV), which bleeds into the eye and leads to substantial vision loss or complete blindness.

Early signs of NPDR often involve blurred vision, floaters, or night vision difficulties, yet many individuals remain asymptomatic until significant retinal damage occurs. Therefore, the research concentrates on automated image classification using DL, which plays a key role in the early determination of the stages of NPDR and facilitating early intervention and prevention of irretrievable loss of vision due to PDR. Figure 1.1 illustrates the NPDR stages as normal, mild, moderate, and severe.

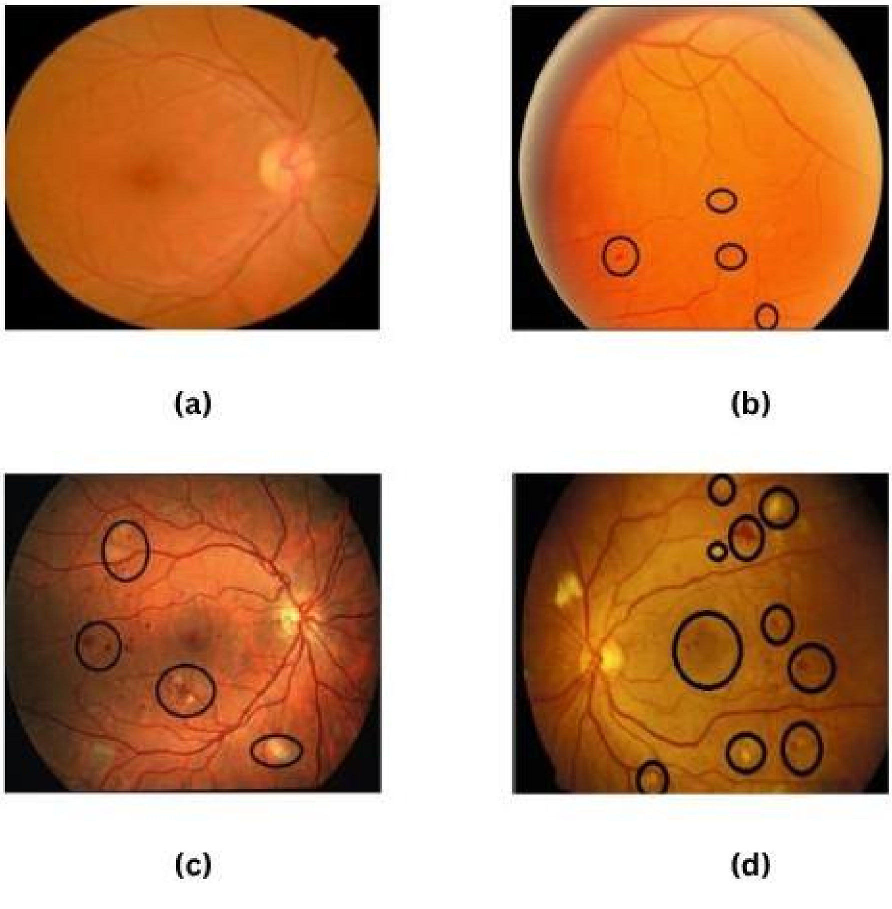

Figure 1.1 Lesion Classification [6,7] a) Normal b) Mild c) Moderate d) Severe

In the past decade, Machine Learning (ML) and DL algorithms have become widely recognised in automating the process of DR screening, offering enhanced diagnostic accuracy and efficiency. CNN and other Artificial Intelligence (AI) techniques show significant promise in identifying DR lesions such as MAs, HEMs and cotton wool spots, which are vital indicators of disease severity. By utilizing these technologies, automated systems assist clinicians in identifying DR at an early stage, even before clinical symptoms become evident. Such systems also offer remote screening, especially in underserved populations where the availability of specialized eye care professionals is scarce [8]. AI-powered DR grading systems use large datasets of labelled fundus images for training and are employed across different clinical settings, encouraging timely and quality provision of care.

To further enhance the accuracy of DR detection and grading, hybrid models that combine multiple DL architectures are being explored. The standard grading protocol for DR severity is essential for guiding treatment decisions and monitoring disease progression, comprising four key stages: No DR (Class 1), Mild DR (Class 2), Moderate DR (Class 3), and Severe DR (Class 4). This classification system provides the critical framework for clinicians to assess the condition and determine appropriate management strategies. Continuous advancements in screening technology promise to improve real-time detection capabilities, which could significantly impact the management of DR and contribute to reducing the global burden of diabetes-related vision loss.

1.2 COMPLICATIONS DUE TO DR

DR is the leading cause of diabetic blindness and results from years of elevated blood sugar that harms the small blood vessels of the retina. As the disease advances from non-proliferative to proliferative stages, blocked vessels trigger retinal ischemia, prompting the dangerous growth of abnormal, fragile new blood vessels. These vessels bleed into the vitreous (vitreous HEM), causing scar tissue formation and retinal detachment (often tractional). They develop in the eye's drainage angle, inhibit outflow from the region, and lead to neovascular glaucoma of the eye. All of these separate complications, vitreous HEM, what used to be called tractional retinal detachment, and neovascular glaucoma, pose significant clinical challenges and result in severe and irreversible loss of vision if not managed with intensity.

In PDR, aneurysmal retinal vessels form on the surface of the retina or extend into the vitreous. This fragile scar tissue tends to flatten as it grows. This type of scar tissue causes considerable traction that pulls on the retina's surface. If the traction is powerful, it will pull the retina off its supporting structure, resulting in a tractional retinal detachment. The warning signs for patients could include flashes of light or a striking increase. After detaching, the retina no longer works, resulting in severe and often irreversible vision loss in the affected area. Timely surgical repair, usually involving the removal of scar tissue and vitreous gel from the eye (vitrectomy) and techniques to reattach the retina, is essential to repair the detachment and prevent vision loss.

Vitreous HEM is another frequent complication of DR, which arises when bleeding from leaking retinal vessels enters the vitreous body, which is the gel substance that fills the interior of the eye. The leaked blood then pools nearby in the middle of the eye, resulting in the visible dark spots or floaters in the visual field. In some patients, vitreous HEM can also resolve spontaneously as the blood is cleared from the eye. However in more severe or moderate cases, vitrectomy or laser surgery are usually needed to remove the pooling of blood and restore vision. Even when restored, vision is not instantaneous, but rather takes a few months, and sometimes the patient still have temporary issues with their vision for some amount of time. Another potentially blinding complication in DR is glaucoma, which develops from uncontrolled vascular proliferation that blocks the normal outflow pathways of the aqueous humour. This fluid fills the anterior segment of the eye. This pressure is called Intraocular Pressure (IOP); if it goes untreated, it causes injury to the optic nerve.

High IOP slowly damages eyesight, causing the typical peripheral vision loss that comes with glaucoma. This pressure-related damage causes partial or complete loss of vision in the later stages. Control of IOP by medications, laser treatment, and surgery is necessary to prevent irreversible damage to the optic nerve.

Blindness is the most serious consequence of DR. Usually, it comes due to macular oedema when fluid accumulation causes the macula, the central section of the retina for high-resolution vision, to swell and distort. This oedema is commonly due to the exudation of blood and fluid from the injured blood vessels. Uncontrolled disease leads to long-standing foveal oedema and irreversible visual loss. Even through medical treatments, vision cannot be restored when the damage is extensive. The prevention of blindness depends on the early diagnosis or treatment of the disease through routine eye examination in the early stages of the disease or the rapid use of laser photocoagulation and/or intravitreal anti-VEGF injections (to reduce swelling, preventing additional vision losses).

Hence the complications from DR, such as retinal detachment, vitreous HEM, glaucoma, and blindness, are directly correlated with the abnormal proliferation of blood vessels into the retina. If left untreated, the complications ultimately result in severe and permanent loss of vision [9].

1.3 CURRENT DIAGNOSTIC AND SCREENING METHODS

With the advancement of imaging technology, imaging techniques to diagnose DR, including Fluorescein Angiography (FA), Optical Coherence Tomography (OCT), OCT Angiography (OCTA), and Color Fundus Photography (FP) in ophthalmology, are needed. With the increasing size of imaging datasets, manual analysis becomes less practical, and automatic computational methods for an efficient data workflow arise [10].

1.3.1 Fluorescein Angiography

FA continues to be a fundamental imaging method for evaluating DR, yielding vital information about the retinal vascular changes. It is beneficial for assessing capillary nonperfusion, ischemia, and retinal NV, which are the major indicator of the disease progression. FA is performed using the intravenous injection of a dye (fluorescein) that fluoresces under a light source, displaying the retinal blood vessels with details. This imaging technique is highly sensitive, facilitating the detection of early vascular abnormalities that may not be obvious during standard clinical examinations, such as FP. One of FA's primary strengths is its ability to detect retinal NV, a sign of proliferative DR. As the disease progresses, the retina responds to ischemia (reduced blood supply) by forming new, abnormal blood vessels to supply oxygen to the affected tissues [11].

These neovessels are often fragile and prone to leakage, which drives further adverse effects such as vitreous HEM or retinal detachment. FA is sensitive in demonstrating areas of NV that are not visible by standard clinical examination and in providing necessary information for treatment planning. This imaging modality aids the direction of interventions such as laser photocoagulation or intravitreal anti-VEGF injections by visualizing vascular abnormalities otherwise not visible.

1.3.2 Optical Coherence Tomography

OCT, a standard method for evaluating DR, generates high-resolution cross-sectional views of the retinal layers [12]. Using near-infrared light, OCT captures the retinal structure, thickness, and reflectivity in detail, which helps diagnose and monitor the treatment of DR, especially Diabetic Macular Edema (DME). The non-contact technique identifies retinal thickening from fluid accumulating in the macula, a DME symptom that results in vision loss if untreated. OCT allows accurate measurement of Central Retinal Thickness (CRT), an essential parameter for establishing the severity and determining the timely treatment of DME.

1.3.3 OCT Angiography

It is a non-invasive image enhancement from previous OCTs. It allows noninvasive visualization of the retinal blood flow instead of the contrast dye. It produces sharp high resolution angiography images and can examine retinal microvasculature in detail in DR, and because of its fast scanning feature, OCTA imaging is helpful in gathering information on microvasculature problems, as well as for early diagnosis and follow-up of DR progress [13].

The foveal capillary network, a central region for detecting retinal changes in diabetic patients, is the strength of OCTA. Since the fovea is responsible for clear central vision, the microvascular changes of the fovea are typically indicative of early DR, which is effectively monitored using OCTA, which images both the superficial and deep retinal capillaries and offer a comprehensive view of retinal blood flow. Such capability is beneficial for detecting ischemia and NV, the most critical indicator for the progression of DR, without invasive dye injection. Therefore, it is used to make the early diagnosis and monitor the progression of DR.

1.3.4 Colour Fundus Photography

Fundus photographs can be performed using a number of different methods, such as Single-Field, Two-Field, Seven Standard Fields (SSF), Widefield, Ultra-Widefield, Stereo, and Montage imaging; among this SSF approach is a valid method in FP to assess DR severity that captures detailed images of the macula, optic disc, and the periphery, providing a generalized picture of retinal function. This standardized approach improves the diagnostic precision for specific stages of DR and treatment planning. Comparatively, regarding detecting the early indications of DR, such as MAs, HEMs, and neovascularization, FP via SSFs is more sensitive than conventional ophthalmoscopy. Such high sensitivity makes FP a helpful screening tool for DR, particularly for detecting early diabetic changes that are not easily detectable by clinical examination.

However, the SSF method also has its disadvantages. Its biggest challenge is that it is resource-intensive because it relies on the skill of photographers and processors who have to be capable of getting clear pictures and analyzing the results [14].

1.4 LIMITATIONS OF CURRENT DIAGNOSTIC APPROACHES

Current DR diagnostic techniques have several limitations that hinder their widespread implementation and usability. Although ophthalmoscopy is widely used in outpatient-based screening programs for DR, it is highly dependent on skilled professionals. This dependence on skilled labour limits its scale-up, particularly in areas with few ophthalmologists. Thus, ocular examination for DR is not sufficient in these lower resource settings where specialized care is restricted.

FA is an essential diagnostic procedure for DR, enabling detailed retinal blood flow images and identifies diseases such as NV. Nonetheless, it is invasive, and the dye must be injected, making the method time-consuming and uncomfortable for patients. However, the lack of trained personnel to interpret the results complicates its use, especially in rural or disadvantaged regions lacking specialized health care.

OCT and OCTA systems offer high-resolution images of the retina, all owing to the accurate assessment of the thickness of the layers and vasculature in DR. However, their efficacy is based on subjective interpretation, which results in interclinician variability. These are expensive technologies with limited availability, especially in resource-poor healthcare systems. Their high cost and special demands preclude routine application; therefore, using them for large-scale Primary Health Care (PHC) based DR screening in underserved or poor areas is not feasible.

Colour FP is excellent at detecting DR but requires trained staff to obtain and interpret good-quality images. It is a time-consuming process, and specialized equipment makes it difficult to undertake at scale. These limitations limit its utility for daily use, especially in resource-poor regions lacking a skilled workforce. Hence, FP is less practical for large-scale DR screening in resource-limited areas because of its accessibility and affordability problems.

DR screening structurally faces barriers such as restricted access to specialized care, high cost, and dependence on professional labour. Further obstacles to wider use include the lengthy test and expensive device requirements, particularly in low-resource regions. These issues result in delayed early detection and treatment and contribute to the high numbers of people with undiagnosed DR at risk of severe vision loss in resource-poor settings. These challenges underscore the need for more effective, affordable, and universally available diagnostic modalities that will guarantee the prompt detection and management of DR and in turn decrease the global incidence of vision loss caused by diabetes.

1.5 ROLE OF AI IN DR

AI technology is modernizing the field of ophthalmology. By providing primary ophthalmologists with valuable diagnoses and treatments, AI technology will reshape the entire field of ophthalmology, making the industry more effective and better serving patients.

By using large clinical data, AI offers novel solutions to the problem of screening and managing ophthalmic disorders, particularly when subspecialist care is not readily available in primary eye centers. For conditions such as DR, where prevention and early intervention are crucial in preventing vision loss, patient-care platforms that utilize AI that analyze large quantities of patient data provide a significant edge [15].

One remarkable key advantage of AI applications in ophthalmic care is their potential to overcome the problems associated with the enormous patient burden of fundus diseases. AI-based algorithms optimize the workflow by automating the reader evaluation of retinal images. This is particularly helpful in scenarios with a high patient prevalence and a scarce availability of experts. AI automates the tedious work of image feature analysis and interpretation. This allows ophthalmologists to concentrate on quality patient care and treatment planning, improving overall clinical efficiency and decreasing misdiagnosis rates.

In traditional ophthalmologic image analysis, the mission of detecting features is heavily based on the insight of researchers and a priori knowledge, thus limiting both the scope and resolution of detection. These hand methods are naturally slow, have a heavy mechanical workload, and are sensitive to operator differences. Contemporary computer image processing proposes new solutions for assessing complex patterns in fundus images based on the use of large datasets to find and validate these patterns automatically. A data-driven strategy for recovering the vast range of features improves diagnostic accuracy, better adaptation to variability, and decreased reliance on manual non-semantic features as in traditional feature engineering methods [16].

The high-tech computer analysis software becomes clearer when viewing larger datasets, more capable at catching subtle pathological details that may not be caught when one is visually scanning the microscope. This skill is extremely valuable at the time of DR diagnosis because early detection of MAs and vessel changes really enhances the quality of treatment. It assists health care professionals in sharpening clinical judgment, which results in improved patient care and overall health.

1.6 MOTIVATION

Modern co**m**putational vision systems, concentrating on hierarchical pattern recognition models, have made considerable progress in recent years and are increasingly competing with traditional analysis techniques in various application domains. These models learn rich visual representations over the group of pixels. The model capability bypasses the constraints of former feature extraction approaches and further benefits applications such as image categorization and object localization.

Computer-aided DR detection has significantly progressed due to the capability of extracting detailed information automatically from medical images. However, precise diagnostic classification systems for this disorder have proven remarkably challenging to develop. Conventional DL models may easily fall into overfitting, require high computing costs, and have poor detection performance for all DR stages [17]. Problems with small special databases result in loss of generalization capacity and diagnostic reliability.

TL is a potential approach to address these challenges. It facilitates the transfer of learned knowledge from large-scale general-purpose datasets. Such a strategy improves generalization and performance, particularly when annotated medical data is scarce.

EfficientNetV2L is a cutting-edge DL model with enormous potential for the DR classification task because of its high efficiency and accuracy. Combined with a custom CNN model using the hybrid strategy, EfficientNetV2L takes advantage of pre-training feature extraction, which alleviates overfitting, reduces the computation complexity, and enhances the detection of all four early DR stages, significantly improving the overall diagnostic performance.

The hybrid model, which merges TL with custom CNN layers, verifies that generalization of the model on different datasets is possible and achieves high accuracy when classifying DR stages.

Using the generalization capacity of EfficientNetV2L and fine-tuning it with task-specific layers, the model is aimed at precise detection of fine retinal abnormalities, the key to the early and accurate detection of DR. These systems are immensely beneficial in the context of diagnostic workflows. It supplements the ophthalmologist's diagnosis and aids in better treatment decisions, patient wellness, and healthfulness.

1.7 OBJECTIVES

The main aim of this research is to classify DR stages. To achieve this goal, the main objectives are as follows:

1. To fine-tune and evaluate the pretrained models for DR stages classification.
2. To design and evaluate a custom CNN architectures for DR stages classification.
3. To develop the hybrid DL model by incorporating the fine-tuned pretrained model and the custom CNN architecture that predicts all four stages (Normal, Mild, Moderate, and Severe) of DR with better accuracy.

1.8 METHODOLOGY

This research aims to design a reliable and efficient DR classification system capable of accurately detecting all four DR stages while addressing the key challenge of computational complexity.

A systematic methodology is employed structured into three key phases, aligning with the research objectives to improve model’s effectiveness in DR stages classification. Figure 1.2 presents a flowchart depicting the developed methodology's step-by-step framework.

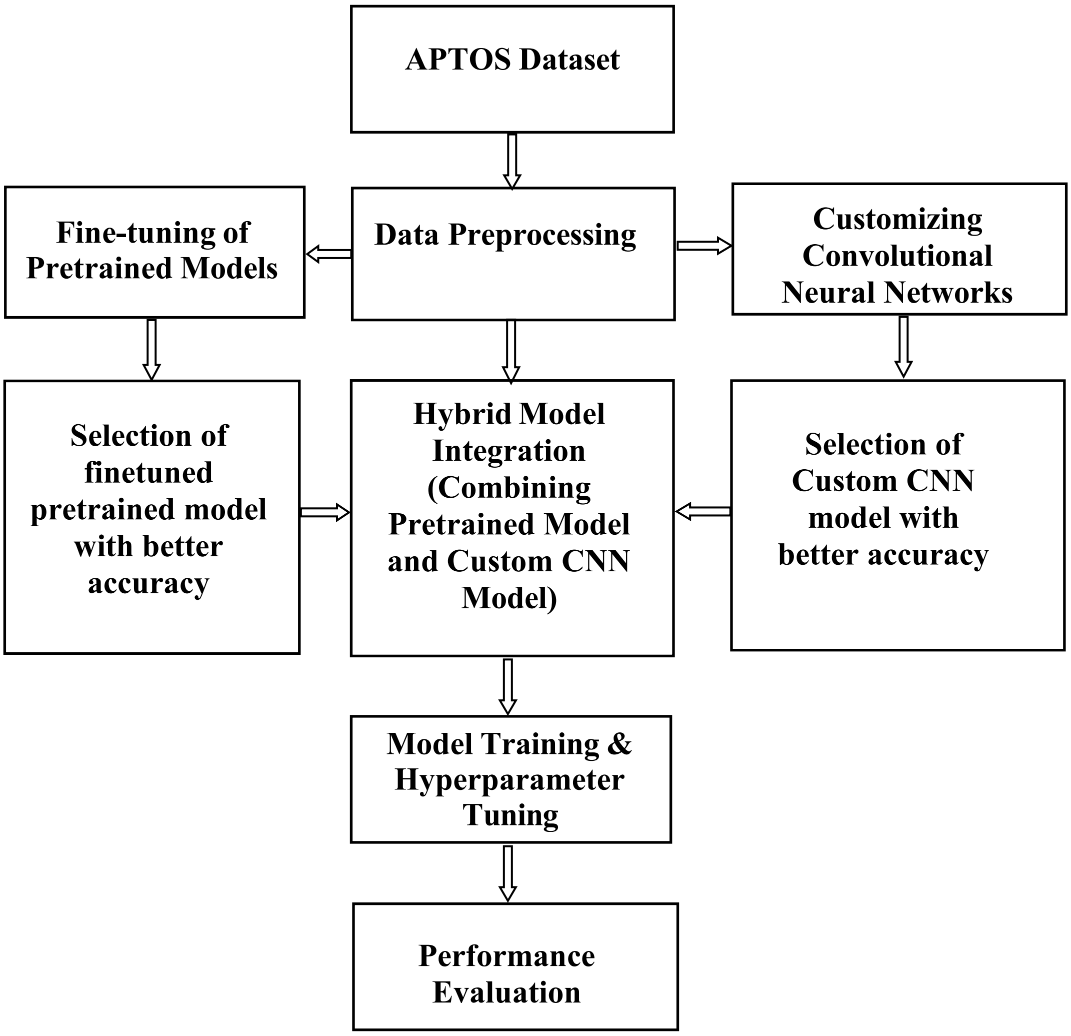

Figure 1.2 Methodology

Phase 1: Fine-tuning of the pretrained models

In the first phase, TL is employed to enhance the accuracy of high-performance models in DR classification, such as VGG16, ResNet-50, InceptionV3, DenseNet121 and EfficientNetV2L finetuned with the help of the APTOS dataset, utilizing its deep feature extraction capabilities. The fine-tuning models mitigated overfitting and enhanced model adaptability to real-world clinical scenarios.

Phase 2: Developing a Custom CNN Model

The second phase focuses on designing a custom CNN architecture to further strengthen the model's robustness and optimize computational efficiency. This model contains suitably designed convolutional layers, Batch Normalization (BN), and dropout operations to learn the necessary retinal features and avoid overfitting.

This architecture enriches the pre-trained models by generating more refined feature representations for DR classification. The lightweight nature of this CNN ensures computational feasibility and can be applied in a clinical setting, particularly in resource-poor regions unable to access high-end GPUs.

Phase 3: Developing a Hybrid Model

The third phase combines the pre-trained model with specialized CNN layers to create a hybrid approach where the layers efficiently extract features and grasp high-level representations from retinal images, leading to high classification accuracy. The hybrid architecture is optimized through hyperparameter tuning, including learning rate scheduling, dropout regularization, batch size optimization, and adaptive optimization techniques (such as Adam), ensuring efficient training and convergence.

A comprehensive performance evaluation is carried out to validate the system using four metrics: accuracy, precision, recall and F1-score. The system is trained using the APTOS dataset and tested using real-time clinical data, ensuring clinical relevance and reliability. A comparative analysis is performed to assess improvements in classification accuracy and computational efficiency and compared with an existing hybrid model. By implementing a hybrid DL framework with fine-tuned architectures, this work intends to deliver an automated and scalable system for detecting DR at an early stage, ultimately supporting early intervention, and improving patient outcomes, particularly in resource-limited healthcare settings.

1.9 ORGANIZATION OF THESIS

The organization of the chapters is designed as follows:

Chapter 1 introduces DR, emphasizing its clinical importance and the key role of early detection. This chapter outlines the motivation for applying DL to DR classification. It presents the specific objectives of this thesis, which focus on improving classification accuracy, minimizing overfitting, and enhancing computational efficiency.

Chapter 2 reviews the literature on DL techniques for medical image classification emphasizing DR detection. It examines the strengths and limitations of traditional models, highlighting research gaps, particularly in detecting the four stages of DR and the need for models with lower computational requirements.

Chapter 3 focuses on fine-tuning various pretrained models to evaluate their efficacy in accurately identifying the four stages of DR. It covers the dataset overview, methodology, training procedures, and performance metrics for model evaluation, with the results highlighting the best-performing model.

Chapter 4 discusses about the custom CNN models for DR classification. It also reports on the architectural choices, data augmentation strategies, training schedules, and performance measures, highlighting the model's benefits and the extent to which it outperforms the fine-tuned pre-trained models.

Chapter 5 introduces a hybrid model to improve the classification of DR which merges the pretrained model and custom architecture. The results are compared with the existing hybrid model. This chapter describes the custom dataset overview, preprocessing steps, methodology, training phase, and evaluation metrics.

Chapter 6 summarizes the findings and contributions of this research and discusses the development of reliable models for DR classification. It also details future research directions, such as enlarging data samples, exploring novel architectures, and conducting clinical validation to ensure the feasibility of hybrid model in practice.