CHAPTER I
INTRODUCTION

1.1 Research Background

Diabetes mellitus is a chronic metabolic disorder characterized by elevated blood glucose levels. Over time, this condition leads to retinal microvascular complications, resulting in blood vessel leakage known as Diabetic Retinopathy (DR) (Sarki et al., 2020). As a leading cause of vision loss in adults aged 20–74, early screening is vital to prevent irreversible blindness. However, traditional screening is hindered by a severe shortage of specialists; it is estimated that a functional system requires 50 ophthalmologists per million people (Resnikoff et al., 2020), creating a massive diagnostic burden that human grading alone cannot sustain.

To address this, the Virtual Vision Image and Pattern Research Group (VVIP-RG) developed a web-based DR detection system using the ResNet50 architecture, trained on the IDRiD (Porwal et al., 2018). While ResNet50 is a standard for handling complex features, its implementation in the VVIP-RG Retina Project revealed significant drawbacks. A previous study by (Alviantara, 2023) achieved a classification accuracy of only 53.84% and reported long processing times. Furthermore, the high computational requirements of ResNet50 make it fundamentally "heavy," consuming excessive energy and resulting in high inference latency (Dolz et al., 2023; Hernández et al., 2024). This prevents the system from being deployed on mobile platforms or in resource-constrained environments where preliminary screening is critical.

The necessity for a lighter model is further emphasized by the global transition toward smartphone-based teleophthalmology. Research indicates that smartphone-acquired images can achieve near-perfect diagnostic accuracy (AUC of 0.99) compared to traditional tabletop cameras (Han et al., 2021). Moreover, modern screening is moving toward ultra-low-cost, portable hardware, such as the GSoP adapter, which costs less than $10 to produce (Abay et al., 2025). Because these portable systems, like the PanOptic iExaminer, are increasingly used by non-specialists for rapid retinal analysis (Hu et al., 2020), the underlying AI must provide instantaneous, on-device feedback without relying on expensive, high-latency cloud infrastructure.

As the field moves toward large-scale in-home testing to alleviate the burden on healthcare systems (Pieczynski et al., 2021), there is an urgent need for models that balance high sensitivity with low latency. While non-mydriatic tabletop cameras suffer from a high technical failure rate of 12.1%, efficient architectures can mitigate these issues through real-time, on-device quality checks (Kanclerz et al., 2021).

Recent studies have shown that MobileNetV3 is significantly more efficient than its predecessors and heavy CNNs, achieving DR detection accuracies as high as 96.72% to 98.56% (Prajapati, Baliarsingh, & Hota, 2023; Rimaru et al., 2024). By utilizing MobileNetV3 as a successor to the ResNet50 model used by Alviantara, this study aims to develop a lighter yet more accurate classification scheme. This optimization is the key to enabling the integration of DR screening into mobile applications, expanding accessibility in areas with limited computing infrastructure, and ultimately supporting earlier preliminary severity classification.

1.2 Problem Identification

Although the potential for using AI in the screening of Diabetic Retinopathy (DR) is immense, its implementation on mobile devices faces significant technical challenges, which form the core of this study. First, the IDRiD dataset used exhibits a severe class imbalance, where the distribution of data across severity levels is uneven (Porwal et al., 2018). This imbalance is the primary cause of the model's poor performance in recognizing minority classes, as shown by (Alviantara, 2023), whose study achieved accuracy 53.84% because the model struggled with generalization due to the dominance of certain classes.

Second, there is a fundamental challenge in using the MobileNetV3 architecture, which is trained by default on the ImageNet dataset. The visual characteristics of ImageNet, which contains common everyday objects, are very different from those of retinal fundus images, which are highly specific and medical in nature (Howard et al., 2019). This domain gap raises the question of whether conventional transfer learning techniques are sufficiently effective, or whether a hybrid approach and additional optimization schemes are needed to adapt the model's weights to the textures and microvascular lesions in the retina (Alyoubi et al., 2020; Zhou et al., 2023).

Therefore, this research is designed as an experimental study to explore various improvement schemes aimed at addressing these shortcomings. The primary focus of this experiment is to identify the best approach for improving accuracy while maintaining computational efficiency, whether through data balancing techniques, specific data augmentation methods, or potential hybrid architectures. This study aims to demonstrate whether MobileNetV3 can be optimized to overcome the limitations of the IDRiD dataset and the differences in ImageNet characteristics, thereby producing a model that is not only lightweight for mobile devices but also can support preliminary screening (Prajapati, Baliarsingh, & Hota, 2023; Rimaru et al., 2024).

1.3 Research Questions

Based on those problems, the study formulations raised in this study are

1. How is the scheme design in classifying the severity level of diabetic retinopathy disease using MobileNetV3 architecture?
2. What is the performance level of the designed scheme for classifying the severity level of diabetic retinopathy using the MobileNetV3 architecture?

1.4 Research Objectives

Based on the research questions above, the research objectives are

1. To design a scheme for classifying the severity level of diabetic retinopathy using the MobileNetV3 architecture.
2. To obtain the performance of the designed scheme for classifying the severity level of diabetic retinopathy using MobileNetV3.

1.5 Research Limitations

The scope and limitations of the problems in this study are listed as follows:

1. The primary dataset used in this research is the Indian Diabetic Retinopathy Image Dataset (IDRiD). Selected experiments also use controlled additions of retinal images from DiaRetDB and APTOS. The official IDRiD validation and test subsets remain unchanged during the external-data experiments.
2. The mobile application is an initial Flutter-based prototype designed to demonstrate the feasibility of running a local TensorFlow Lite model for diabetic retinopathy (DR) severity classification using raw RGB fundus images.
3. The prototype is developed and tested exclusively as an Android APK. An iOS version is not included in this scope.
4. A supplementary printout-based input trial was conducted to approximate a mobile acquisition scenario. In the main prototype flow, retinal fundus images can still be selected from the gallery or camera. The printout-based trial does not replace validation using actual clinical hardware, such as fundus cameras or smartphone retinal adapters.
5. The prototype is not a production-ready clinical system. Usability testing, interface evaluation, and clinical validation by medical professionals are outside the scope of this study.
6. Public release, application-store deployment, maintenance, privacy policy development, and production-level security assessments are not covered.

1.6 Research Significances

The expected significance of this research is described as follows:

1. Theoretical Benefits
   a) For researchers, this research is expected to add insight into the implementation of deep learning, especially MobileNetV3, for severity-based diabetic retinopathy classification. This research also provides an experimental reference regarding preprocessing, data-balancing, external-data addition, and deployment-oriented model selection.

2. Practical Benefits
   a) For health-related implementation, the publication of this research is expected to provide an initial reference for the future development of mobile-based preliminary diabetic-retinopathy severity classification systems. However, this research does not directly evaluate the use of the system by health workers or integrate it into real clinical workflows. Therefore, the practical benefit for health workers is indirect and should be understood as a potential contribution after further clinical validation, usability testing, and deployment assessment.
   b) For the community, the results of this study can help introduce the role of artificial intelligence in supporting early awareness of diabetic retinopathy. The prototype demonstrates how retinal fundus images can be processed locally through a mobile application, although the output should not be interpreted as a clinical diagnosis.
   c) For researchers, the results of this study provide a reference in determining suitable schemes for diabetic-retinopathy severity classification, especially in comparing experimental-best models and deployment-oriented models. This research can also provide a description of the implementation of digital image processing, deep learning, model conversion, and mobile prototype integration.
   d) For application developers, this research provides an initial example of integrating a TensorFlow Lite diabetic-retinopathy classification model into a Flutter-based Android prototype. The prototype can be used as a starting point for further development, including broader device benchmarking, iOS implementation, batch image processing, and hybrid cloud-local mechanisms.