# CHAPTER I
# INTRODUCTION

## 1.1 Background

Diabetic Retinopathy (DR) is a microvascular complication of diabetes mellitus that causes progressive damage to the blood vessels of the retina [1], [2]. This condition can develop into visual impairment and permanent blindness if not detected and treated early [3]. Based on global-scale research, the prevalence of DR in individuals with diabetes reaches 22.27%, with 6.17% classified as vision-threatening Diabetic Retinopathy (VTDR), and this figure is projected to increase to 160.5 million cases by 2045 [4]. Other studies also indicate that the number of DR patients has continuously increased over the past 20 years, from 10.9% in 2007 to 20.8% in 2021, demonstrating the growing burden of this disease in society [5]. Based on observed clinical characteristics, Diabetic Retinopathy is classified into several severity levels, ranging from mild to advanced proliferative stages, where each level requires a different management approach [6]. Given that the progression of this disease can be prevented through early intervention, detection and classification of Diabetic Retinopathy is a crucial aspect of eye health management in the diabetic population.

Conventional detection of Diabetic Retinopathy is performed through evaluation of retinal fundus images by ophthalmology specialists. In this process, specialist physicians identify various pathological indicators on the retina, such as microaneurysms, hemorrhages, exudates, and neovascularization, to determine the severity of the disease [7]. Although this method can produce accurate results, the examination process relies on the availability of expert personnel and may vary in terms of inter-evaluator interpretation consistency [8]. These limitations have driven the development of computer-based systems to support the medical diagnostic process, known as computer-aided diagnosis (CAD). CAD systems aim to provide objective and consistent analysis of medical images, thereby assisting medical personnel in the diagnostic decision-making process [9].

In the field of computer-aided diagnosis, deep learning methods have been widely used due to their ability to automatically learn feature representations from data without requiring manual feature engineering [10]. Convolutional Neural Network (CNN) is one of the most widely used deep learning architectures, designed to process grid-shaped data such as images, through a convolution mechanism that can effectively recognize local spatial patterns in visual data [11]. One study showed that CNN models can classify the severity of Diabetic Retinopathy with an accuracy of up to 90.6% [12].

Along with the development of deep learning architectures, various modern CNN variants have been developed with a focus on improving model efficiency and performance. Several architectures that have been widely used in computer vision applications include ResNet, DenseNet, and EfficientNet [13]. Among these various architectures, EfficientNet has demonstrated competitive performance in various computer vision applications using a *compound scaling* approach to balance three dimensions of the neural network, namely depth, width, and input resolution [14]. This approach allows proportional scaling of model capacity without inefficient complexity additions. EfficientNet has several variants with different complexity levels, ranging from B0 to B7, where each variant represents a *trade-off* between accuracy and computational efficiency [14].

The EfficientNet-B0 variant is the baseline version of the EfficientNet family developed through neural architecture search [14]. EfficientNet-B0 has a relatively smaller number of parameters compared to other variants [14]. This architecture uses mobile inverted bottleneck convolution (MBConv) as its main building block, enabling efficient feature extraction with lower memory usage and computational operations [14], [15]. The *compound scaling* characteristics of EfficientNet-B0, which balances network depth, width, and resolution, make it suitable as a feature extraction *backbone* in medical image classification applications, particularly in retinal fundus images that require the ability to detect fine details such as microaneurysms, hemorrhages, and exudates [16].

One of the challenges in developing Diabetic Retinopathy classification models is the imbalanced data distribution across classes. Public datasets used in research, such as *APTOS 2019*, show significant variation in sample counts at each severity level, creating representational imbalances between mild, moderate, severe, and proliferative classes [17], [18], [19]. Class imbalance in medical datasets can cause deep learning models to be biased toward classes with more samples, resulting in decreased performance on underrepresented classes [20].

One approach developed to address the limitation of annotated data in medical images is *few-shot learning*, a method that can reduce data scarcity problems and improve the speed and robustness of medical image analysis using a limited number of training examples per class [21]. *Few-shot learning* differs from conventional supervised learning that requires abundant data for each class, as it focuses on learning generalizable representations from only a few examples [21]. In the context of *few-shot learning*, various *meta-learning* approaches have been developed, generally categorized into three main groups: metric-based, model-based, and optimization-based learning [22]. Among these, the metric-based learning approach operates by learning a distance or *similarity function* in the *embedding space*, enabling the model to compare new samples with class representations learned during the meta-training process.

Prototypical Network is one of the metric-based *few-shot learning* methods that constructs prototype representations for each class in a high-dimensional *embedding space* [23]. This method works by computing the centroid or average of *support set* sample embeddings for each class as that class's prototype [23], [24]. *Query sample* classification is then performed based on the Euclidean distance between the *query sample*'s embedding and each class prototype, where samples are classified into the class with the nearest prototype [23]. This approach allows the model to recognize classes with limited sample counts through discriminative representation learning in the *embedding space* [23], [25]. Prototypical Networks have been applied in various classification domains, including computer vision and medical imaging, with results demonstrating the ability to handle imbalanced data conditions [26].

Previous research has explored the use of EfficientNet as a *backbone* for feature extraction in Diabetic Retinopathy classification, and the application of Prototypical Networks to handle *few-shot learning* conditions in various application domains [16], [17], [18], [19], [24], [25], [26], [27]. However, based on literature review, no research has been found that specifically develops a hybrid EfficientNet-B0 and Prototypical Network model for multi-class Diabetic Retinopathy classification on the *APTOS 2019 dataset*. The hybrid model in this context refers to an *end-to-end* integration of EfficientNet-B0 as a feature extractor with a Prototypical Network as a classifier, where both components are trained jointly through backpropagation to simultaneously optimize embedding representation and prototype formation. This hybrid approach differs from conventional approaches that use EfficientNet and classification methods separately, as it enables more adaptive feature representation learning toward the characteristics of *few-shot learning* and class distribution imbalance.

Based on the problems and opportunities that have been described, this study aims to develop a hybrid model integrating EfficientNet-B0 as a feature extraction *backbone* with a Prototypical Network as a classification mechanism to handle multi-class Diabetic Retinopathy classification with five severity levels (0–4). This hybrid model is trained *end-to-end*, enabling joint optimization of feature extraction and class prototype formation. This study is expected to contribute to the development of Diabetic Retinopathy classification methods, particularly in the context of handling class distribution imbalance through the integration of *few-shot learning* with an efficient CNN architecture.

## 1.2 Problem Formulation

Based on the background described, this study explores the integration of advanced deep learning for medical diagnostics. The following research questions have been formulated to guide this investigation:

1. How is the hybrid EfficientNet-B0 and Prototypical Network model applied for multi-class Diabetic Retinopathy classification?

2. How does the integration of Prototypical Network affect classification performance compared to baseline models?

### 1.3 Research Objectives

Based on the problem formulation that has been established, this study has the following objectives:

1. To apply the hybrid EfficientNet-B0 and Prototypical Network model for multi-class classification of Diabetic Retinopathy images.
2. To analyze the effect of Prototypical Network integration as a classifier on classification performance compared to baseline models.

### 1.4 Research Benefits

Based on the research objectives stated above, the following research benefits are identified:

1. For the Researcher, this study provides an opportunity to implement knowledge gained during studies, particularly in the fields of deep learning, *few-shot learning*, and computer-aided diagnosis in medical image classification. This study also trains analytical and problem-solving skills in developing hybrid models and strategies for handling class imbalance in medical datasets.
2. For the Healthcare Sector, the results of this study can contribute to the development of computer-aided diagnosis systems for objective and consistent detection and classification of Diabetic Retinopathy. The developed model has the potential to support medical personnel in the Diabetic Retinopathy screening process.
3. For Academics and Further Researchers, this study can serve as a reference for students, researchers, or practitioners interested in developing *few-shot learning*-based systems and integrating EfficientNet-B0 architecture for medical image classification with class imbalance. Additionally, this study enriches the literature on the application of Prototypical Networks in the medical imaging domain and can form the basis for developing similar methods in other diseases or medical imaging modalities.

## 1.5 Research Limitations

Based on the problem formulation described above, several research limitations are established to ensure that this study does not become too broad and remains focused on the problems to be addressed:

1. The dataset used is *APTOS 2019*, so the research results are only tested on the distribution and characteristics of that dataset. External data from other sources is not included in the scope of training and evaluation.
2. The research focus is limited to retinal fundus images for multi-class Diabetic Retinopathy classification with five severity levels (0: No DR, 1: Mild, 2: Moderate, 3: Severe, 4: Proliferative DR).
3. The research uses a hybrid model with EfficientNet-B0 as the *backbone* feature extractor and Prototypical Network as the classifier trained *end-to-end*. Other *backbone* architectures such as ResNet, DenseNet, or transformer-based models are not within the scope of this research.
4. Image preprocessing is limited to Circle Crop and Enhanced Green techniques. Synthetic data augmentation techniques or other preprocessing methods are not used to maintain the integrity of medical information in fundus images.
5. Resampling techniques (oversampling/undersampling), class weighting, and synthetic data generation are not used in this study in order to directly observe the characteristics and *trade-offs* that arise. Analysis of the impact of class imbalance on per-class performance is part of the model evaluation.
6. Model performance evaluation uses classification metrics including accuracy, macro-precision, macro-recall, macro-F1-score, confusion matrix, and quadratic weighted kappa (QWK) to assess overall and per-class model performance, as well as to analyze performance *trade-offs* between majority and minority classes in the metric-based learning approach.
7. Evaluation is performed using a training and test set split on the *APTOS 2019 dataset*, without testing on external datasets.