CHAPTER I
INTRODUCTION

1.1 Background

The retina plays a crucial role among ocular structures, including blood vessels, the optic nerve, and photoreceptors. Hence, any disruption within these structures can lead to abnormalities in retinal function. Currently, when we observe the top 3 common diseases that lead to blindness in Thailand, namely cataracts, refractive lens error, and diabetic retinopathy (DR) (Isipradit et al., 2014). Although, cataracts are the most common cause of blindness but they are surgically treatable and often present clear early symptoms. On the other hand, diabetic retinopathy is irreversible and asymptomatic in its early stages. Consequently, DR emerges as the most common disease associated with irreversible vision loss and negligible early detection. As a result, we have to rapidly investigate and determine the solution for addressing this disease. Diabetic retinopathy (DR) is a condition stemming from elevated blood pressure and abnormal glucose levels associated with diabetes. DR can result in leakage and swelling of blood and fluid within the retina, ultimately leading to blindness and posing a significant threat to visual health. Additionally, studies have reported that approximately one-third of diabetes patients are at risk of developing DR (Wong et al., 2018) and once diagnosed with DR, these patients are 2.5 to 4 times more likely to develop sustained blindness (Wykoff et al., 2021). Moreover, statistics from 2010 reveal that 3.7 million individuals experienced visual impairment, and 0.8 million suffered from blindness due to DR (Wong et al., 2018). Interestingly, The prevalence of diabetic retinopathy (DR) is estimated to increase from 103 million in 2020 to 130 million in 2030 and further to 160 million in 2045 (Teo et al., 2021). However, A report from 2015 indicated that there were approximately 230,000 ophthalmologists across 194 countries, including Thailand, with an annual growth rate of 2.6% (Resnikoff et al., 2019). As a result, there is a trend to face a shortage of ophthalmologists in the future, resulting in a higher chance of neglect in diabetic patients, who require regular screening, as well as among DR patients, who need periodic re-evaluation and accurate grading to prevent the onset of severe scenarios. Moreover, research in Thailand revealed a significant disparity in the distribution of ophthalmologists across different regions. While the average ratio stands at approximately one ophthalmologist per fifty thousand individuals nationwide (Estopinal et al., 2013), this ratio is heavily skewed due to the concentration of ophthalmologists in urban areas. For instance, in the capital city, there are approximately 437 ophthalmologists (Royal College of Ophthalmologists of Thailand, 2024) per 5.5 million individuals (Department of Provincial Administration, 2024). Conversely, in larger provinces like Ubon Ratchathani, there are only 9 ophthalmologists (Royal College of Ophthalmologists of Thailand, 2024) serving a population of 1.8 million individuals (Department of Provincial Administration, 2024). This uneven distribution poses challenges in providing timely and accurate diagnoses, as nurses and physicians in hospitals and healthcare centers outside major cities may have to assume responsibilities held by ophthalmologists, increasing the risk of delayed or incorrect diagnoses.

Various object detection algorithms currently are being employed to automatically detect critical structures such as the optic disc and macula. These algorithms aim to replicate the screening methods used by expert ophthalmologists, to address the issue associated with collecting unusable images by nurses or physicians during patient consultations. In (Sinthanayothin et al., 1999), the method utilizes an inverted Gaussian template to detect the macula and employs the brightest pixel to locate the optic disc in the fundus image. Additionally, morphological techniques are deployed to address an exudates and blood vessels in the image, facilitating the locating of the optic disc and macula (Sekhar et al., 2008). In (Welfer et al., 2011), an approach for removing unwanted lesions and noise is proposed, along with optic disc detection using information from the vascular tree in the fundus image. In (Tariq et al., 2012), the method utilizes a Gaussian Mixture Model to locate the macula by aggregating five feature vectors related to the macula. In (Zheng et al., 2014), the approach employs a two-stage detection process, comprising coarse and fine stages, to determine the location of the optic disc. Subsequently, the macula is located through circular scanning around the optic disc. In (Deka et al., 2015), the method extracts blood vessels from the fundus image to utilize the vessel-free area for detecting the macula location. In (Kamble et al., 2017), intensity lines are sampled from the image to create an intensity profile. Then, signal processing techniques, specifically the Daubechies 4 wavelet, are then applied to determine the locations of the optic disc and macula.

Furthermore, to mitigate the risk of incorrect diagnosis, several research studies have employed deep learning for diabetic retinopathy (DR) severity grading. In

(Zhang et al., 2019), the method introduces ensemble backbone networks to enhance feature extraction capabilities and utilizes a customized fully connected neural network, termed SDNN, as a classifier. Additionally, Bayesian optimization for hyperparameter tuning is proposed to enhance the performance of the Inception-V4 network in (Shankar et al., 2020). In (A. He et al., 2020), the DR grading network is enhanced by introducing a novel attention block called the category attention block (CAB), specifically designed to tackle the imbalance issue present in various datasets. Moreover, in (Sun et al., 2021), the method introduces the bio-marker establishing from explainable attention map by leveraging the attention mechanism inherent in the customized Vision Transformer (ViT) network. In (Li et al., 2022), a pyramid network is introduced, capable of processing retinal images at various resolutions, thereby enhancing the network's ability to understand fine-to-coarse details present in the images. Additionally, the method proposes the utilization of attention maps as guidelines for training the networks. In their study, (Tusfiqur et al., 2022) introduced a comprehensive training approach aimed at developing a robust diabetic retinopathy (DR) grading network. This approach comprises three learning approaches: adversarial learning, supervised learning, and expert feedback learning. Adversarial learning is employed to train the lesion segmentation network, utilizing lesion map predictions to refine the grading network. The supervised learning approach trains the grading network using retinal images alongside lesion maps. Additionally, expert feedback learning involves leveraging expert ophthalmologist feedback to validate predictions and tune the grading network.

In clinical practice, the grading of diabetic retinopathy (DR) is conventionally conducted by trained ophthalmologists or retinal specialists, who evaluate retinal fundus images according to standardized clinical protocols, such as the International Clinical Diabetic Retinopathy (ICDR) scale. This grading process necessitates the identification of pathological features, including microaneurysms, hemorrhages, hard exudates, and neovascularization, which are indicative of disease progression. Each image is examined independently and categorized into one of five DR severity levels (0–4). In many hospitals, especially in Thailand, this process is manual, time-consuming, and subject to inter-observer variability, particularly in borderline cases. Due to limited specialist availability, it also causes delays in diagnosis and treatment in rural or resource-limited regions, thereby timely and accurate automated DR grading is crucial to address these issue. Nevertheless, building such systems requires large, diverse, and balanced datasets, criteria that are at odds with the characteristics of our current dataset, which is relatively small and highly imbalanced. While this limitation can be addressed by relying on ophthalmologists to generate new labeled data, the process is hindered by another practical challenge: a substantial portion of the image database consists of medically unsuitable images. Consequently, ophthalmologists must expend significant effort in screening out medically unsuitable images before proceeding with diagnostic labeling, which diminishes efficiency and slows the development of reliable DR grading models. Consequently, we are interested in establishing an end-to-end automated expert system that can perform the screening and grading processes on the retinal image. Additionally, to achieve this both tasks, the expert system must comprise two sub-systems. Firstly, in image screening system, we implement the template matching technique and anatomical knowledge of ocular structures to extract relevant features. Then, these features are utilized to locate the optic disc and macula, fulfilling the criteria of expert requirements. Lastly, we utilize the model or rule to classify image as medically unsuitable and medically suitable retinal image. The reason to use handcrafted feature extraction for ocular composition detection instead of deep learning is driven by the limited size of our available dataset, which comprises approximately 500 images. Secondly, for DR grading, we turn to deep learning to address this task. This decision is motivated by the availability of various public datasets and the outstanding performance of deep learning methods in this task.

1.2 Research Objective

- Implementing the correlation filtering technique to screen the retinal fundus image
- Implementing the deep learning to classify the severity level of DR based on retinal fundus image

1.3 Scope and Limitations

This research project investigates the object detection algorithm for fundus image quality assessment that is utilized to screen the fundus image and establish our own dataset, as well as the deep learning for grading the DR severity level from the fundus image. Moreover, the dataset in this work consists of IDRiD for ocular object detection, APTOS2019 used to train the DR grading network. In object detection, we use the template matching technique to detect the optic disc and macula. The reference template is created by averaging the N number of optic disc or macula images, and we will increase the macula detection by utilizing the ROI of macula. In deep learning, we trained five pretrained networks, including ResNet 50, VGG 19, Inception V3, DenseNet 161, Swin Transformer, or the other networks that can outperform the baseline score, based on DR grading. Then, we select the optimal network or ensemble for improving and tuning. Eventually, we aim to achieve an false discovery rate (FDR) of 0.05 and a recall score of 0.90 in image screening (Coyner et al., 2018; Fleming et al., 2006). In the DR grading network, we expect to exceed human performance, which is typically 0.894 in accuracy, 0.714 in F1 macro score and 0.871 in QWK (Krause et al., 2018).