CHAPTER V
CONCLUSION and OUTLOOK

This thesis introduces a comprehensive end-to-end diabetic retinopathy (DR) classification framework based on retinal fundus imaging, comprising two primary modules: an image screening module and a DR grading module. The screening module employs a template-based correlation filtering approach to detect crucial ocular structures, namely the optic disc and macula, followed by either a rule-based or machine learning (ML)-based decision mechanism to classify images as medically suitable or unsuitable. Extensive evaluation on private datasets confirms the effectiveness of the screening method, which achieves a high recall of 0.906, ensuring the retention of diagnostically valuable images, and a low false discovery rate (FDR = 0.065), indicating a high chance of excluding clinically relevant images. The performance is further enhanced by the use of anatomically tailored template designs and optimal sampling strategies, with the normalized correlation coefficient (CCORR_NORMED) demonstrating robust matching performance under varying illumination conditions. Additionally, the integration of a region-of-interest (ROI) framework further improves the reliability of macula detection. Ultimately, out-of-distribution testing confirms that the proposed method maintains competitive performance when generalized to unseen datasets. In the ML context, generalized models such as Random Forest and Logistic Regression demonstrate promising robustness to distributional shifts, highlighting their potential utility in real-world screening applications. As a result, they reveal the opportunity to improve this work with advanced ML in the future.

Our diabetic retinopathy (DR) grading module employs the Swin Transformer (S) network as the backbone, coupled with a fully connected neural network for classification. In the context of imbalanced data, we address it by using SMOTE to synthesize data along with various data augmentation strategies. The model is further enhanced through fine-tuning on the APTOS 2019 dataset. Moreover, comprehensive evaluation demonstrates that the proposed model achieves an F1 macro score of 0.693 and a quadratic weighted kappa (QWK) score of 0.903, surpassing previously reported methods in both metrics. Moreover, the classification report indicates strong performance in detecting early DR stages, suggesting that the model is well-suited for early-stage

DR screening applications. Nevertheless, detailed analysis via the confusion matrix and classification report reveals underperformance in minority classes, especially severe and proliferative DR, due to overlapping lesion characteristics and limited training data. Additionally, a performance gap remains between the model and expert ophthalmologists, with an F1 macro score of 0.714. To further improve the performance of deep learning-based algorithms, three critical cores must be considered: data quality and quantity, model capacity (i.e., number of parameters), and computational resources. Given the current limitations, our primary focus is on improving the dataset, as the existing data are both imbalanced and limited in size. This constraint has led to suboptimal performance, particularly in minority classes such as severe NPDR and PDR. To address this, we plan to expand the dataset by collecting more diverse samples and by leveraging generative models, such as generative adversarial networks (GANs), variational autoencoders (VAEs), or diffusion models, to synthesize images of advanced DR stages. Once the data-related issues are mitigated, we will focus on increasing model capacity by incorporating architectures with a larger number of parameters, followed by scaling computational resources to support training on larger models and datasets.

In the context of innovation, our work focuses on the development of a comprehensive end-to-end system that integrates both DR grading and image screening modules. The DR grading module is designed to assess the severity level of diabetic retinopathy solely based on a patient's fundus image, thereby offering a user-friendly and accessible diagnostic tool. In parallel, the image screening module evaluates the quality of the fundus images, which is a critical component. By automatically filtering out medically unsuitable images prior to manual labeling, the system reduces the burden on ophthalmologists, enabling them to concentrate on their primary task of annotating images. As a result of this module, we obtain a more informative and higher-quality dataset, which in turn enhances the performance of the DR grading model. Ultimately, by integrating these modules with existing resources, including ophthalmologists, a data lake, and a diagnostic web application, we establish a functional end-to-end system. This platform supports automated DR severity grading through telemedicine while enabling continuous model improvement through feedback and data accumulation. From a market perspective, the global artificial intelligence market for retinal image analysis is valued at approximately USD 147.8 million in 2024 (Transpire, 2024), with a projected compound annual growth rate (CAGR) of 13.5%. Of this, clinical diagnosis and early detection screening account for an estimated 45.6%, or USD 67.4 million

(market.us, 2024). Major players in the field, such as Topcon Corporation, RetinAI Medical AG, and Eyenuk Inc., leverage advantages like 24/7 ophthalmologist-backed systems and advanced AI technologies. However, these solutions are frequently associated with high costs and limited scalability to local or rural areas. Our approach distinguishes itself by targeting affordability and accessibility. We aim to develop both hardware and software solutions that are affordable cost for deployment in local areas, including local healthcare centers, medium-sized hospitals, and small hospitals, estimated at approximately 890 facilities nationwide. Notably, there are currently no domestic competitors in Thailand offering a fully integrated solution comprising both diagnostic instruments and software. The present situation presents a great opportunity for market leadership within the country, with the potential to scale further into the Southeast Asian region.

In the context of practice, The system is currently at the proof-of-concept stage but has been developed with practical deployment. The image screening module has been evaluated on both public and private datasets, including those collected in Thailand, indicating its readiness for real-world implementation. In contrast, while the DR grading module demonstrates reliable performance, surpassing related works and closing the level of ophthalmologists, it has only been validated on a single dataset. As a result, further evaluation across multiple datasets from local hospitals is necessary to assess generalizability and strengthen clinical reliability. At present, the project comprises a fully integrated pipeline, including a data lake, labeller, diagnostic web application, and an image screening system. Thus, the DR grading component remains the final element requiring refinement before full deployment as an automated clinical decision-making system. Nevertheless, if the goal is to establish a decision support system rather than full automation, the current DR grading model is sufficiently reliable, particularly for early-stage detection, and can already serve as a valuable tool for assisting ophthalmologists and physicians. This is especially beneficial in local or rural healthcare settings where access to retinal specialists is limited. Finally, the system can act as a front-line screening tool to flag potentially abnormal cases and support clinicians in making timely and accurate decisions, ultimately improving early detection and accelerating referrals for further examination. In the long-term strategic plan, the proposed system is designed for continuous improvement. Upon integration of new datasets from additional hospitals, the image screening module will automatically filter medically suitable images and then store them into a cleaned database, and the diagnostic outcomes provided by ophthalmologists will be used to continuously retrain the

DR grading model. Once the model achieves a reliable and clinically acceptable level of performance, we plan to deploy the system in partner hospitals such as Suranaree University of Technology Hospital (SUT) and Maharat Nakhon Ratchasima Hospital to collect real-world feedback and further optimize the system. In parallel, we will make efforts directed toward expanding collaborations with other hospitals to increase data diversity and model robustness. Simultaneously, we aim to secure small-scale research funding to support the development of intellectual property, including the pursuit of patents. These patents can serve as collateral for acquiring larger-scale funding to further scale development and deployment. Moreover, as there is not a current domestic competitor offering both integrated software and hardware solutions, this positions us with a unique opportunity to lead the national market. If the implementation proceeds successfully, we plan to expand regionally, targeting the Southeast Asian healthcare market, where similar needs and infrastructure gaps exist.

In summary, this work introduces an end-to-end system for diabetic retinopathy classification, integrating image screening and DR grading modules. The screening component demonstrates readiness for deployment with high recall and low FDR, while the grading model, based on a fine-tuned Swin Transformer, achieves competitive performance with a macro F1-score of 0.693 and QWK of 0.903 but performance on minority classes remains a challenge. Therfore, future work will emphasize on dataset expansion in both diversity and size, and model scaling. As a result of existing infrastructure, including a data lake, web application, and screening system, this work is positioned for real-world decision support application, particularly in local and rural healthcare settings. Fianlly, the system holds strong potential for clinical integration, continuous improvement, and national deployment as a AI-driven diagnostic solution.