

CONTENTS

	Page
ABSTRACT IN THAI	I
ABSTRACT IN ENGLISH	II
ACKNOWLEDGEMENTS	III
CONTENTS	VII
LIST OF TABLES	X
LIST OF FIGURES	XII
CHAPTER	
I INTRODUCTION	1
1.1 Background	1
1.2 Research Objective	4
1.3 Scope and Limitations	4
II LITERATURE REVIEW	6
2.1 Diabetic Retinopathy	6
2.2 Macula and Optic Disc Detection	12
2.2.1 Conventional detection approaches	12
2.2.2 Novel detection approaches	15
2.3 Diabetic Retinopathy Classification	17
2.3.1 Multi-classification	17
2.4 Evaluation Matrix	21
2.4.1 Confusion matrix	21
2.4.2 Accuracy, Precision, and Recall	22
2.4.3 F1 score	22
2.4.4 Receiver Operating Characteristic (ROC) curve	22
2.4.5 Area Under the Curve (AUC)	23
2.4.6 Quadratic Weighted Kappa (QWK) coefficient	24
2.4.7 Euclidean Distance (ED)	26
2.4.8 Intersection over Union (IoU)	26
2.4.9 Average Precision (AP) and Mean Average Precision (mAP)	26

CONTENTS (Continued)

	Page
III RESEARCH METHODOLOGY	29
3.1 Dataset	29
3.2 Image Screening	31
3.2.1 Optic disc and macula detection	31
3.2.2 Screening algorithm by rulebased	34
3.2.3 Screening algorithm by Machine learning (ML)	35
3.3 Evaluation of Image Screening	36
3.4 DR Grading	38
3.4.1 Data preparation	38
3.4.2 Data augmentation and balancing	38
3.4.3 Architecture	38
Swin Transformer	39
3.4.4 Training setting and strategy	40
3.5 Evaluation of DR Grading	41
3.6 Computational Resources	41
IV RESULTS AND DISCUSSION	42
4.1 Results of Image Screening	42
4.1.1 Optical disc and macula detection	42
4.1.2 Screening algorithm	45
4.2 Ablation Study of Image Screening	48
4.2.1 Template size and sampling amount	48
4.2.2 Matching functions	50
4.2.3 Region of interest (ROI)	52
4.2.4 The generalization of the proposed method	53
4.3 Results of DR Grading	54
4.4 Ablation Study of DR Grading	61
4.4.1 Backbone model selection	61
4.4.2 Data sampler	62
4.4.3 Loss functions	64
4.4.4 The impact of SMOTE	65
4.4.5 The impact of fine-tuning	69

CONTENTS (Continued)

	Page
V CONCLUSION AND OUTLOOK	73
REFERENCES	78
APPENDIX	87
CURRICULUM VITAE	89

LIST OF TABLES

Table		Page
3.1	The tabular data features for ML model.	35
4.1	The general performance of proposed method on IDRiD dataset.	43
4.2	The macula detection on Messidor dataset.	43
4.3	The performace of proposed image screening	46
4.4	The performace of machine learning in image screening. Bold represents the best score and <u>underline</u> represents the second best score.	47
4.5	CCOEFF_NORMED achieves the highest AP ₅₀ score for both de- tection task.	52
4.6	The quantitative result of ROI technique.	52
4.7	The performace of machine learning in image screening. Bold represents the best score and <u>underline</u> represents the second best score.	54
4.8	The classification report of proposed model on the APTOS 2019 dataset.	56
4.9	The classification report of proposed model on the APTOS 2019 dataset.	59
4.10	The general performance of proposed model compared the hu- man performance. Bold represents the best score.	60
4.11	The general performance of each backbone model. Bold repre- sents the best score and <u>underline</u> represents the second best score.	62
4.12	The influence of imbalance data sampler to DenseNet 161 and Swin Transformer. Bold indicates the best score.	63
4.13	The influence of loss functions on DenseNet 161 and Swin Trans- former. Bold represents the best score.	65
4.14	The classification report of Swin Transformer model with and with- out SMOTE.	66

LIST OF TABLES (Continued)

Table		Page
4.15	The classification report of DenseNet161 model with and without SMOTE.	67
4.16	The classification performance of KNN model with different pre-trained models. Notably, we use macro score for comparison .	67
4.17	The classification report of Swin Transformer model with and without SMOTE.	68
4.18	The classification report of DenseNet161 model with and without SMOTE.	68
4.19	The classification performance of KNN model with different pre-trained models. Notably, we use textbfmacro score for comparison	68
4.20	The classification performance of DenseNet161 and Swin s models before and after fine-tuning. Pretrained models are trained on ImageNet-1K dataset. Fine-tuned models are pretrained model and then, tuned on APTOS 2019 dataset. Notably, we use only macro scores for comparison.	69
4.21	The classification performance of KNN model with different image types and pretrained models. Notably, we use only macro score for comparison	70
4.22	The classification performance of DenseNet161 and Swin s models before and after fine-tuning. Pretrained models are trained on ImageNet-1K dataset. Fine-tuned models are pretrained model and then, tuned on APTOS 2019 dataset. Notably, we use only macro scores for comparison.	71
4.23	The classification performance of KNN model with different image types and pretrained models. Notably, we use only macro score for comparison	72

LIST OF FIGURES

Figure		Page
2.1	The retinal image comparison of normal and DR.	9
2.2	Diabetic retinopathy (DR) lesions are categorized based on the progression of lesion development; ranging from mild to severe stages. a) Microaneurysms (MA); b) Retinal Hemorrhage (HM); c) Hard Exudates (HE); d) Soft Exudates (SE); e) Intraretinal Microvascular Anomalies (IRMA) and Venous Beading (VB); f) and g) Neovascularization (NV); h) Fibrous Proliferation (FP); i) Preretinal and Vitreous Hemorrhage (PRH, VH) Barbara Davis Center for Diabetes School of Medicine, 2024.	9
2.3	The figure illustrates the various severity levels of DR in the patient's retina: a) No DR; b) Mild NPDR; c) Moderate NPDR; d) Severe NPDR; and e) PDR Karthik, 2019.	11
2.4	The confusion matrix.	21
2.5	The ROC curve illustrate the trending of FPR and TPR as a function of threshold.	23
2.6	ROC curve interpretation by curve reading.	24
2.7	ROC curve interpretation by utilizing AUC.	24
2.8	The figure illustrate the two PR curve including raw PR curve and interpolated PR curve, which is obtained by utilizing the 11-point interpolation approach.	28
3.1	The class distribution of these datasets reveals that class 0 (no diabetic retinopathy) is the majority class, indicating an imbalance issue. The total number of images in the training sets of APTOS 2019 is 3,662.	31
3.2	the cropped dark field images, a) images without cropping, b) images with cropping.	32
3.3	The figure shows the result of ROI cropping from both eyes, a) the left eye and b) the right eye. The macula and optic disc's ground-truth locations are represented by the red and green dots.	33

LIST OF FIGURES (Continued)

Figure		Page
3.4	The acceptance region of R1 and R2.	35
3.5	The figure illustrates components of the image that are used to calculate the $Dy_{intercept}$ and $Dr_{distance}$ features. (left) The $Dy_{intercept}$ is the deviation between OD-M line and reference line on the y-axis at $x = 0$. (right) The $Dr_{distance}$ is the distance from the center of the image to the center of the macula.	36
3.6	The example of background images which generate by covering the optic disc and macula by the mean value of fundus image. .	37
3.7	The distribution of the synthetic samples generated by SMOTE in the prior and posterior stage.	39
3.8	The architecture and attention mechanism of the Swin Transformer. a) The window shifting mechanism allow the model to capture global and nearest neighbour context. b) Swin Transformer architecture and blocks	40
4.1	The agreement area of each R-criterion on Messidor dataset. . .	43
4.2	The detection results of the proposed method on the IDRiD dataset. The three images above (a–c) showcase good quality predictions, whereas the images below (d–f) exhibit poor quality predictions. Where, cross sign (X) is a predicted location, the dot sign (●) is a ground-truth location, the green color represents the optic disc, and the red color represents the macula.	44
4.3	The detection results of the proposed method on the Messidor dataset. The three images above (a–c) showcase good quality predictions, whereas the images below (d–f) exhibit poor quality predictions. Where, cross sign (X) is a predicted location, the dot sign (●) is a ground-truth location, and the red color represents the macula.	45
4.4	The confusion matrix of proposed image screening method. . .	46
4.5	Example of image screening results: (a) True positive; (b–c) False positive; (d) True negative; and (e–f) False negative.	47

LIST OF FIGURES (Continued)

Figure		Page
4.6	The template parameter of optic disc. Heatmap depicts the sensitivity of template dimensions via the AP score. The line plot shows the impact of sampling.	48
4.7	The template parameter of macula. Heatmap depicts the sensitivity of template dimensions via the AP score. The line plot shows the impact of sampling.	49
4.8	The optimal template for optic disc and macula.	50
4.9	The qualitative result of macula detection, a) without ROI, b) with ROI. Where, cross sign (X) is a predicted location, the dot sign (●) is a ground-truth location, the green color represents the optic disc, and the red color represents the macula.	53
4.10	The confusion matrix of proposed model.	55
4.11	The figure illustrates the ROC curve and AUC for each class, indicating that the proposed model outperforms in classifying class 1 across a variety of threshold values. However, the other classes exhibit constraints related to the threshold value, impacting their lower classification performance.	57
4.12	The PR curve and AUC for each class, which is more informative than the ROC curve in context of imbalance dataset.	58
4.13	The confusion matrix of proposed model.	59
4.14	The figure illustrates the ROC curve and AUC for each class, indicating that the proposed model outperforms in classifying class 1 across a variety of threshold values. However, the other classes exhibit constraints related to the threshold value, impacting their lower classification performance.	60
4.15	The PR curve and AUC for each class, which is more informative than the ROC curve in context of imbalance dataset.	61

LIST OF FIGURES (Continued)

Figure		Page
4.16	The figure illustrates the data sampler strategies: a) the entire training dataset; b) the training data in a batch with a sequential sampler strategy, which sequentially selects the data from beginning to end; c) the training data in a batch with a random sampler strategy, which randomly selects the data from the entire dataset; and d) the training data in a batch with an imbalance sampler strategy, which attempts to select the data from each class. Each color represents data from a different class.	63
4.17	The t-SNE visualization of the DenseNet 161 model's feature space. The colors represent different classes; cyan is class 0; red is class 1; green is class 2; orange is class 3; and yellow is class 4, respectively. a) and b) illustrate the feature space of model without tuning while c) and d) illustrate the feature space of model with tuning. Moreover, a) and c) are the feature space of RGB images while b) and d) are the feature space of grayscale images.	71
4.18	The t-SNE visualization of the Swin s model's feature space. The colors represent different classes; cyan is class 0; red is class 1; green is class 2; orange is class 3; and yellow is class 4, respectively. a) and b) illustrate the feature space of model without tuning while c) and d) illustrate the feature space of model with tuning. Moreover, a) and c) are the feature space of RGB images while b) and d) are the feature space of grayscale images.	72
A.1	The Human Subject Protection (HSP) certificate from Suranaree University of Technology.	88
A.2	The Ethics of AI certificate from University of Helsinki.	89