

Supplementary Tables

Table S2: Demographic data of individuals in SEED external test set

	Number of individuals	Number of eyes/images
Aged 60 years old and above (%)	1,594 (55.8)	2,982 (54.6)
Male gender (%)	1,449 (50.7)	2,784 (51.0)
Malay (%)	993 (34.8)	1,870 (34.3)
Chinese (%)	577 (20.2)	1,118 (20.5)
Indians (%)	1,287 (45.0)	2,467 (45.2)

Table S3: Number of eyes by diabetic retinopathy severity in both training and test datasets

Kaggle Training Dataset		
Non-Referable	Normal	30280
	Mild	3131
Referable	Moderate	5078
	Severe	759
	Proliferate	283
Kaggle Internal Validation Dataset		
Non-Referable	Normal	803
	Mild	79
Referable	Moderate	127
	Severe	22
	Proliferate	7
Kaggle Internal Test Dataset		
Non-Referable	Normal	802
	Mild	83
Referable	Moderate	132
	Severe	19
	Proliferate	9
SEED Validation Dataset		
Non-Referable	No Retinopathy	4382
	Minimal	157
	Mild	371
Referable	Moderate	313
	Severe	25
	Proliferate	207

Messidor-1 Validation Dataset		
Non-Referable	Class 0	546
	Class 1	153
Referable	Class 2	247
	Class 3	254

Table S5: Performance of the models in the internal Kaggle dataset after adding 200 poor quality retinal images

<i>Model</i>	<i>Sensitivity analysis: internal test set (1,045 images + 200 poor quality images)</i>	
	AUC (%)	Sensitivity @ 80% Specificity
VGG19	83.3	72.8
ResNet50	88.5	82.3
DenseNet201	89.1	77.6
InceptionV3	87.2	79.5
EfficientNetV25	87.0	80.0
VAN_Small	88.3	80.4
SWIN_Tiny	95.2	93.3
CrossViT_Small	92.4	89.0
ViT_Small	92.9	85.2

Table S6: Performance of models in detecting referable DR among eyes of individuals aged below and above 60 years old in SEED external test.

Model	Individuals aged below 60 years old	Individuals aged 60 years and above
	AUC (%)	AUC (%)
VGG19	96.5	91.5
ResNet50	95.3	89.5
DenseNet201	96.0	89.7
InceptionV3	95.6	87.6
EfficientNetV25	96.4	91.7
VAN_Small	97.6	92.6
SWIN_Tiny	98.7	95.8
CrossViT_Small	96.5	92.8
ViT_Small	97.0	94.0

Table S7: Performance of models in detecting referable DR among eyes of female and male individuals in the SEED external test.

Model	Among female individuals	Among male individuals
	AUC (%)	AUC (%)
VGG19	93.5	94.3
ResNet50	92.0	92.7
DenseNet201	92.8	92.8
InceptionV3	92.0	91.3
EfficientNetV25	94.0	94.0
VAN_Small	95.1	95.2
SWIN_Tiny	97.0	97.5
CrossViT_Small	94.5	94.4
ViT_Small	95.5	95.5

Table S8: Performance of models in detecting referable DR by ethnic subgroups in SEED external test.

Model	Among Malay individuals	Among Chinese individuals	Among Indian individuals
	AUC (%)	AUC (%)	AUC (%)
VGG19	95.4	88.5	94.8
ResNet50	94.1	85.6	93.8
DenseNet201	94.8	87.0	93.8
InceptionV3	92.0	89.1	92.5
EfficientNetV25	96.2	88.9	94.6
VAN_Small	95.9	91.3	96.2
SWIN_Tiny	98.0	97.5	97.0
CrossViT_Small	97.1	92.0	94.7
ViT_Small	97.1	93.9	96.0