

Figure S1. The specific models and detailed collecting protocol by smartphone. **a** Smartphone-specific operation details **b** Patient preparation

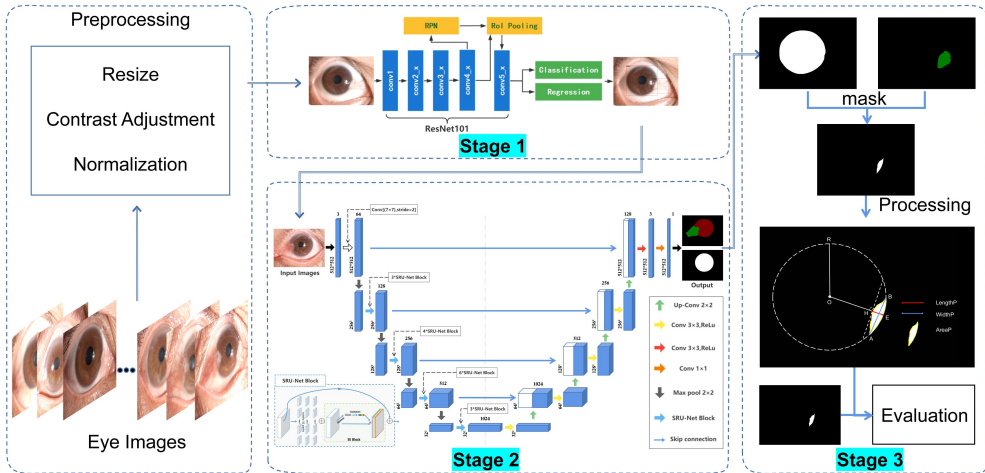


Figure S2. The framework of the automatic detection and grading of the pterygium system.

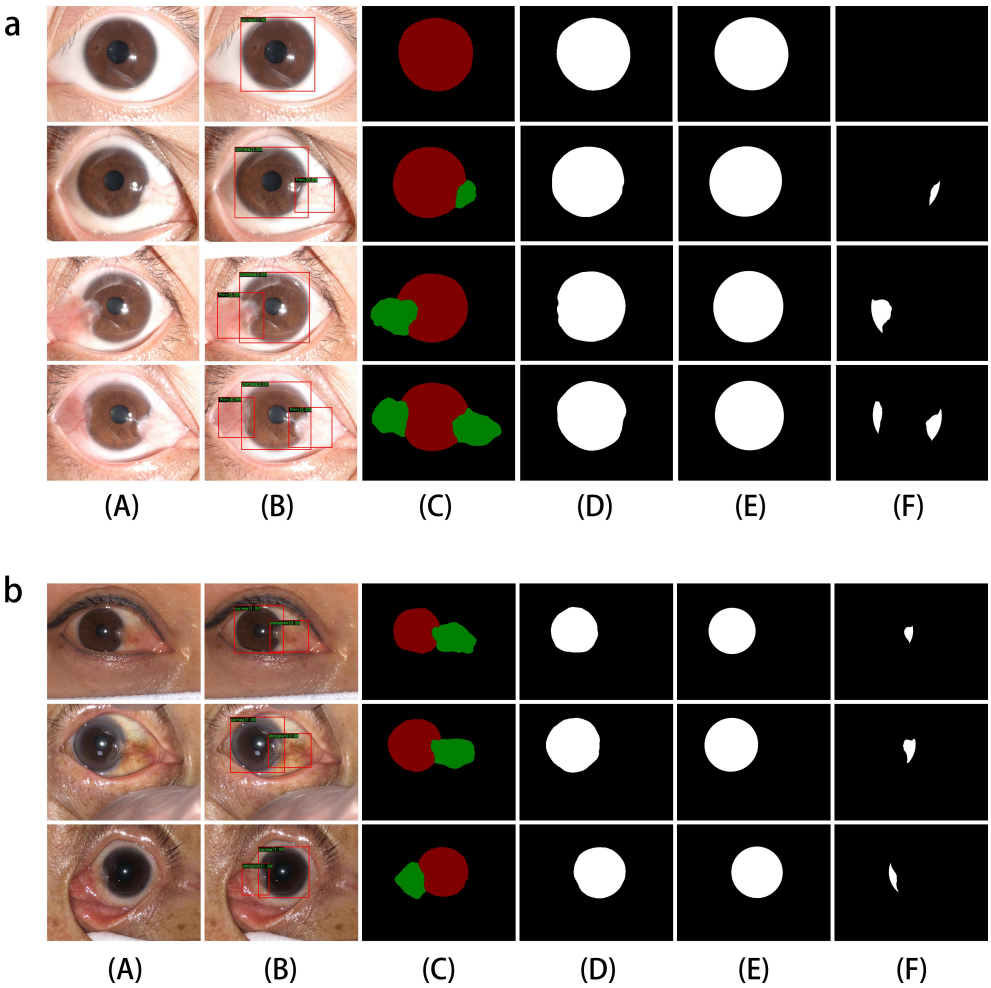


Figure S3. The experimental results of original detection and segmentation images based on SLD and SPB. a The experimental results from SLD. **b** The experimental results from SPB. **A-F** demonstrated the origin images, the detection results, the segmentation results of model Cornea-P, the segmentation results of model Cornea-S, the cornea fitting images, and the images of pterygium invading the cornea.

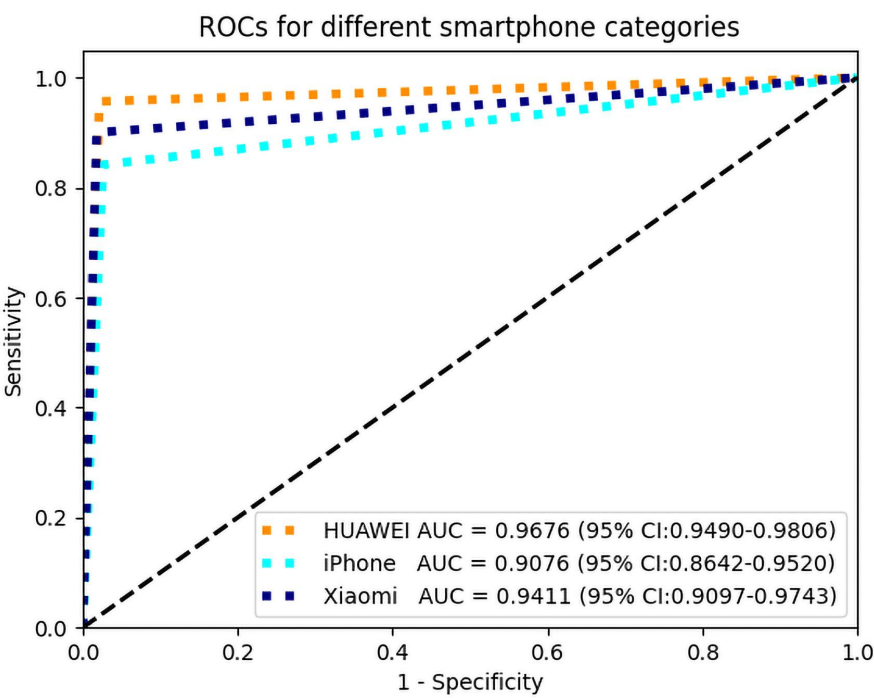


Figure S4. ROC curve and AUC based on images of different smartphone brands in SPB.

Evaluation metrics	SLD	SPB
mAP	0.9881	0.9563
mIoU	0.9788	0.9100
mAcc (%)	96.60	95.24

mAP, mean average precision; mIoU, mean intersection over union; mAcc, mean accuracy of pterygium detection.

Table S1. Performance of DM in detecting pterygium based on SLD and SPB.

CLC	SLD						SPB					
	F ₁	Sensitivity	Specificity	AUC (95% CI)	Kappa	Acc (%)	F ₁	Sensitivity	Specificity	AUC (95% CI)	Kappa	Acc (%)
Grade I	0.9526	0.9377	0.9852	0.9615 (0.9382-0.9819)			0.8837	0.9157	0.9676	0.9416 (0.9119-0.9689)		
Grade II	0.7586	0.9167	0.9592	0.9380 (0.9009-0.9726)			0.8060	0.8182	0.9845	0.9013 (0.8452-0.9533)		
Grade III	0.9359	0.9060	0.9849	0.9455 (0.9234-0.9677)			0.9027	0.8788	0.9483	0.9266 (0.9005-0.9497)		
Micro-Avg	0.9118	0.9201	0.9764	0.9478 (0.9262-0.9718)	0.9193	92.11	0.8981	0.8709	0.9668	0.9295 (0.8952-0.9566)	0.9086	88.31

AUC, area under receiver operating characteristic curve; Acc, accuracy of the system in pterygium grading for classification.

Table S2. Performance of SM3 in pterygium grading for classification in SLD and SPB.

	SLDS						SPBS					
	F ₁	Sensitivity	Specificity	AUC (95% CI)	Kappa	Acc (%)	F ₁	Sensitivity	Specificity	AUC (95% CI)	Kappa	Acc (%)
Grade I	0.9606	0.9531	0.9512	0.9522 (0.9346-0.9712)			0.9683	0.9619	0.9524	0.9603 (0.9443-0.9750)		
Grade II	0.7500	0.8571	0.9694	0.9133 (0.8468-0.9597)			0.8400	1.0000	0.9619	0.9806 (0.9735-0.9886)		
Grade III	0.9538	0.9394	0.9861	0.9628 (0.9458-0.9808)			0.8919	0.8462	0.9697	0.9079 (0.8796-0.9306)		
Micro-Avg	0.9448	0.9165	0.9689	0.9569 (0.9396-0.9764)	0.8972	94.29	0.9313	0.9360	0.9613	0.9426 (0.9175-0.9618)	0.8521	92.38

AUC, area under receiver operating characteristic curve; Acc, accuracy of the system in pterygium grading for classification; SLDS, slit-lamp images of the same group of patients; SPBS, smartphone-based images of the same group of patients.

Table S3. Performance of SM3 in pterygium grading of the same group of patients in SLDS and SPBS.

CLC	F ₁	Sensitivity	Specificity	AUC (95% CI)	Acc (%)
HUAWEI	0.9549	0.8143	0.9729	0.9676 (0.9495-0.9829)	95.65
iPhone	0.9331	0.8474	0.9738	0.9076 (0.8585-0.9485)	84.13
Xiaomi	0.9586	0.8839	0.9823	0.9411 (0.9062-0.9708)	90.00

Table S4. Performance of SM3 in pterygium grading of three different smartphone brands.

CLC	SPBO	Sensitivity	Specificity	AUC (95% CI)	Acc (%)
	F ₁				
Expert1	0.8532	0.8674	0.9168	0.9355 (0.9102-0.9562)	91.40
Expert2	0.9146	0.8619	0.9552	0.9597 (0.9412-0.9752)	94.62
Expert3	0.9235	0.9095	0.9616	0.9677 (0.9510-0.9827)	95.70
Expert-Avg	0.8971	0.8129	0.9445	0.9543 (0.9379-0.9717)	93.91
System-Avg	0.9248	0.7569	0.9624	0.9385 (0.9144-0.9620)	88.52

Table S5. Comparison of the performance of SM3 and experts in grading of pterygium.