



# Prediction of Early Visual Outcome of Small-Incision Lenticule Extraction (SMILE) Based on Deep Learning

Qi Wan · Shali Yue · Jing Tang · Ran Wei · Jing Tang ·  
Ke Ma · Hongbo Yin · Ying-ping Deng

Received: January 11, 2023 / Accepted: February 7, 2023 / Published online: February 24, 2023  
© The Author(s) 2023

## ABSTRACT

**Introduction:** Deep learning (DL) has been widely used to estimate clinical images. The objective of this project was to create DL models to predict the early postoperative visual acuity after small-incision lenticule extraction (SMILE) surgery.

**Methods:** We enrolled three independent patient cohorts (a retrospective cohort and two prospective SMILE cohorts) who underwent the SMILE refractive correction procedure at two different refractive surgery centers from July to September 2022. The medical records and surgical videos were collected for further analysis. Based on the uncorrected visual acuity (UCVA)

at 24 h postsurgery, the eyes were divided into two groups: those showing good recovery and those showing poor recovery. We then trained a DL model (Resnet50) to predict eyes with early postoperative visual acuity of patients in the retrospective cohort who had undergone SMILE surgery from surgical videos and subsequently validated the model's performance in the two prospective cohorts. Finally, Gradient-weighted Class Activation Mapping (Grad-CAM) was performed for interpretation of the model.

**Results:** Among the 318 eyes (159 patients) enrolled in the study, 10,176 good quality femtosecond laser scanning images were obtained from the surgical videos. We observed that the developed DL model achieved a high accuracy of 96% for image prediction. The area under the curve (AUC) value of the DL model in the retrospective cohort was 0.962 and 0.998 in the training and validation datasets, respectively. The AUC values in two prospective cohorts were 0.959 and 0.936. At the video level, the trained machine learning (ML) model (XGBoost) also accurately distinguished patients with good or poor recovery. The AUC value of the ML model was 0.998 and 0.889 in the retrospective cohort (training and test datasets, respectively) and 1.000 and 0.984 in the two prospective cohorts. We also trained a DL model which can accurately distinguish suction loss (100%), black spots (85%), and opaque bubble layer (96%). The Grad-CAM heatmap indicated that our models can

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s40123-023-00680-6>.

Q. Wan · S. Yue · R. Wei · J. Tang · K. Ma (✉) ·  
H. Yin · Y. Deng (✉)  
Department of Ophthalmology, West China  
Hospital of Sichuan University, Chengdu City,  
Sichuan Province, China  
e-mail: 280463919@qq.com

Y. Deng  
e-mail: dyp\_wch@163.com

J. Tang  
Department of Ophthalmology, The People's  
Hospital of Leshan, Leshan City, Sichuan Province,  
China

recognize the area of scanning and precisely identify intraoperative complications.

**Conclusions:** Our findings suggest that artificial intelligence (DL and ML model) can accurately predict the early postoperative visual acuity and intraoperative complications after SMILE surgery just using surgical videos or images, which may display a great importance for artificial intelligence in application of refractive surgeries.

**Keywords:** Deep learning; Machine learning; Prediction; Scanning images; SMILE

### Key Summary Points

#### *Why carry out this study?*

With the application of deep learning and the availability of massive numbers of clinical images, there is a new opportunity to evaluate old techniques for predicting patient diagnosis and prognosis.

#### *What was learned from the study?*

We created an artificial intelligence framework to predict early postoperative visual acuity after SMILE surgery based on surgical videos.

It is now feasible to accurately predict the early postoperative visual acuity and intraoperative complications after SMILE surgery just using surgical videos or images.

## INTRODUCTION

Refractive error, the most prevalent cause of correctable vision impairment, is expected to affect over 6 billion individuals by 2050. Laser correction of refractive error will become the popular elective surgery performed worldwide [1]. Small manual incision lenticule extraction (SMILE) is a novel corneal refractive surgery that is recently been developed for use in patients with myopia and myopic astigmatism to correct

vision, becoming the first choice of treatment for an increasing number of patients and ophthalmologists. As such, SMILE has become the most mainstream refractive surgery, with the advantages of minimal invasion, flapless, and preservation of intact corneal morphology, and provides faster healing of corneal nerve fiber, better biomechanical strength, and a lower incidence of dry eye [2]. Currently, over 2 million myopic patients have undergone SMILE surgery worldwide [3]. Although numerous previous clinical studies have proven its' safety, effectiveness, predictability, and stability in correcting refractive errors, a number of intraoperative and postoperative complications have been reported in different clinical outcomes [4]. Our group has observed that in the clinical setting about 5% of SMILE patients experience poor visual recovery on the first postoperative day (< 20/25). This is similar to the findings of Chansue et al. [5] and Ganesh et al. [6], who reported that about 90%-95% of patients' uncorrected distance visual acuity was 20/20 on the first day after SMILE. However, postoperative visual acuity recovery is still an important indicator for evaluation of laser vision correction, which is as important as the advantages of the surgery itself, and is also closely associated with patients' satisfaction.

The surgical steps in SMILE divided into three main steps: (1) femtosecond laser lenticule construction; (2) lenticule separation; and (3) lenticule extraction. Intraoperative complications may occur in each step. The lenticule separation and extraction steps mainly depend on the surgeon's operative experience and surgical skills. Consequently, there is a potential for various intraoperative complications, such as corneal cap perforation or incisional tear, lenticule dissection difficulties, lenticule remnant, bleeding, and partial centering, when the surgeon is at the initial phase of the surgical learning curve [7, 8]. With improvement in surgical skills and the popularization of SMILE surgical technology, such intraoperative complications correlated with lenticule separation and extraction will be largely avoided [9]. However, lenticule construction is completely dependent on femtosecond laser scanning. Femtosecond laser-related complications, such

as suction loss, black spots, and opaque bubble layer, inevitably affect the quality of intraoperative femtosecond laser scanning. Consequently, the poor quality of femtosecond laser scanning will directly determine the level of difficulty of lenticule separation and extraction, as well as delaying postoperative visual recovery [10].

Recent breakthroughs in artificial intelligence (AI), particularly deep learning (DL), have shown considerable promise for diagnosing a number of prevalent diseases using clinical images [11]. For example, machine and DL have been widely conducted in image processing for pathomics, radiomics and genomics [12]. With the application of DL and the availability of massive numbers of clinical images, there is a new opportunity to evaluate old techniques for predicting patient diagnosis and prognosis. Therefore, in this study, we used a DL model to identify SMILE scanning images and predict the early postoperative visual acuity through supervised learning.

## METHODS

### Selection Cohorts for SMILE Study

This is a retrospective–prospective paired eye-control study that included a retrospective SMILE cohort and two independent prospective SMILE cohorts. The retrospective SMILE cohort was consecutively recruited by the Refractive Surgery Center of West China Hospital, Chengdu, China from July 2022 to August 2022. Two independent prospective SMILE cohorts were respectively collected by the Refractive Surgery Center of West China Hospital and Ophthalmic center, People's Hospital of Leshan, Leshan, China, from August 2022 to September 2022.

In this study, patients with myopia and myopic astigmatism underwent a SMILE refractive correction procedure performed by two surgeons (Y-pD at Chengdu and JT at Leshan). Both surgeons are the director of the Refractive Surgery Center at their respective hospitals and have > 5 years of experience with SMILE surgery, having completed > 2000

procedures per year during this time period. Preoperative baseline examinations of the anterior and posterior segments were performed on all patients. The inclusion criteria were: (1) age > 18 years but < 45 years; (2) spherical refraction < − 10.00 D and myopic astigmatism ≤ − 3.00 D; (3) corrected distance visual acuity (CDVA) ≤ 0 (LogMAR); (4) refractive errors stable for at least 1 year and without any corneal diseases before surgery; (5) central corneal thickness (CCT) ≥ 480 μm and residual stromal thickness (RST) ≥ 280 μm; and (6) postoperative uncorrected visual acuity (UCVA) after 24 h ≤ 0.1 in one eye and > 0.2 (logMAR) in the other eye. The exclusion criteria were: (1) patients who had anisometropia ≥ 1.5 D; (2) existing corneal epithelial damage after surgery; and (3) patients retained myopia postoperatively.

The study was carried out in accordance with the principles of the 1964 Helsinki Declaration and its later amendments and was approved by the Ethics Committee of West China Hospital. Before enrolling in the study, each subject provided written informed consent.

### Surgical Procedure

Oxybuprocaine hydrochloride eye drop was used as a topical anesthesia. The eye had been sterilized and docked. For the SMILE process, a 500-kHz VisuMax femtosecond laser system (Carl Zeiss Meditec, Jena, Germany) with an energy of 130 nJ was employed. The following scans were performed on the lenticule: spiral in for the posterior plane, border, spiral out for the anterior plane, and side cutting. The diameters of the optical zone and corneal cap were 6.0–6.5 mm and 7.0–7.5 mm, respectively. The corneal cap thickness was 120–130 μm. The eye was undocked after the suction was released. The lenticule was first separated at the anterior surface, then split at the posterior surface with a blunt spatula. Next, the lenticule was extracted from the corneal stroma via a small incision. The postoperative care included antibiotic and topical steroid eye drops (0.1% tobramycin dexamethasone [Alcon China Ophthalmic Product Co., Ltd., Beijing, China]; 0.5%

levofloxacin [Santen Pharmaceutical, Ikoma, Nara, Japan]), which were prescribed 4 times a day for 1 week. Also, artificial tears (0.1% sodium hyaluronate) with no preservatives were applied 6 times per day.

### Scanning Images Acquired and Processing

Videos of SMILE procedures were acquired from the VisuMax laser and then divided into a ‘good visual outcome’ category ( $UCVA \leq 0.1$ ) and a ‘poor visual outcome’ category ( $UCVA > 0.2$ ) based on postoperative UCVA after 24 h. The SMILE scanning images were then extracted from the videos (1 image per 0.5 s). Taking into account the large number of images acquired from SMILE videos (usually 80–160 images per video), we selected those images obtained from the end of the posterior plane scanning to side cutting for subsequent analysis. The number of scanning images per video generally ranged from 30 to 34. The eyes of the retrospective SMILE cohort were divided into training and validation datasets at random at a 7:3 ratio. The training dataset was used for model building and hyperparameter tuning, while the validation dataset was used to evaluate generalization performance. Both data augmentation and normalization were employed for the training images, however only normalization was used for validated images. In our investigation, we used random affine modification and horizontal patch flipping to enhance the data. After Z-score normalization on RGB channels, the upgraded images were center cropped to  $224 \times 224$  pixels. This simplified procedure is shown in Fig. 1.

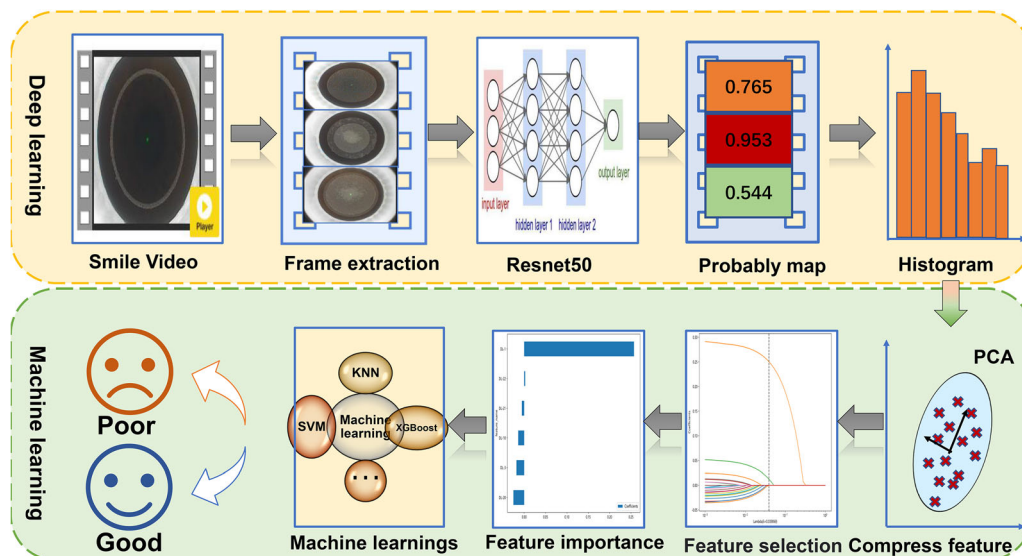
### DL: Feature Extraction and Screening

The videos for the retrospective SMILE cohort were first cropped into scanning images and trained as a DL Resnet50 model. We use a batch size of 32 and default weight initialization. The default optimizer was SGD with a learning rate of  $10^{-2}$  and L2 regularization of  $10^{-5}$ . We trained the Resnet50 model for 50 epochs until the validation loss failed to improve. During image prediction, the Resnet50 model was used to compute scanning image probability with

the video label. Because the videos consisted of numerous scanning images, we aggregated the likely scanning images into a probability map of the video, which was then used to calculate the features based on the image likelihood histogram. Moreover, we conducted principal component analysis (PCA) to compress the likelihood histogram into 24 DL features. Pearson correlation analysis was initially used to eliminate redundant DL features. If the coefficient of two features was  $> 0.9$ , one of the two features was deleted. After that, LASSO-penalized feature selection was used to identify the most significant features. Then, seven classic machine learning classifiers (Decision-Tree, Extra-Tree, KNN, LightGBM, Random-Forest, SVM, and XGBoost) combined with tenfold cross-validation were applied to train models for predicting each video’s classifications [13–15]. Two independent prospective SMILE cohorts were processed in the same way to establish the accuracy and robustness of the model in clinical application.

### Statistical Analysis

All statistical analyses were performed using R (v 4.0.3) or Python (v 3.8.0) with installed packages. Pytorch (v 1.10.1) in Python was used to implement all DL frameworks. The machine learning algorithms were run using Python’s “sklearn” package. Receiver operating characteristic (ROC) curves and area under the curve (AUC) values were generated using the “pROC” package in R. The logarithm of the minimum angle of resolution (logMAR) units were used in the statistical analysis for visual acuity. Continuous variables were described using the mean  $\pm$  standard deviation (SD) or median with interquartile ranges (IQR), and the categorical variables were described using frequencies. The correlation test was evaluated using Pearson coefficients. The Wilcoxon test was used to compare two groups, while the Kruskal–Wallis test was used to compare more than two groups. The Chi-square test was performed to evaluate the associations between cohorts and clinicopathological traits.



**Fig. 1** Analytic workflow in our current investigation. *KNN* K-Nearest Neighbor, machine learning algorithm, *PCA* principal component analysis, *Resnet50* 50-layer convolutional neural network, *SVM* support vector

machine, *XGBoost* Extreme Gradient Boosting machine learning library

## RESULTS

### Study Cohorts

A total of 216 eyes from 113 patients (84 male and 137 female) met our selected criteria and were included in the retrospective SMILE cohort. The median age of the subjects was 27.5 (IQR 22–32) years. Preoperative sphere was – 5.00 D (IQR – 6.00 to – 4.00 D) and the cylinder was – 0.50 D (IQR – 1.00 to – 0.25 D). Preoperative corneal thickness was 539 mm (IQR 519–554 mm), and keratometric 1 (K1) power was 42.85 D (IQR 42.00–43.58 D). The two prospective SMILE cohorts consisted of 48 eyes from 24 individuals (19 male and 29 female) and 54 eyes from 27 individuals (30 male and 24 female), respectively. The baseline characteristics of all patients according to cohort are given in Table 1.

### Performance of Scanning Image Classifier

The scanning image classifier was constructed in the training dataset and validated in the validation dataset of the retrospective SMILE

cohort (training:validation datasets ratio: 7:3) (Table 2). Construction of this image classifier consisted of two steps: image prediction and video prediction. To summarize, the SMILE procedure video was first limited from the completed posterior plane scanning to side cutting, following which the SMILE video was cropped into scanning images, which were then sent into a DL model (Resnet50) to predict postoperative visual acuity status at the level of the images. Secondly, a histogram of scanning image probability was utilized to merge many likely images into a probability matrix of the video. To unify the scanning features of the SMILE video, we performed PCA to compress the probability matrix into 24 DL features. Finally, based on the DL features, we used various machine learning methods to determine the visual acuity of the patient post-surgery.

The performance of the scanning image classifier was evaluated by using the validation dataset in the retrospective SMILE cohort. We found that with increasing number of training iterations, the training accuracy converged near 90% at the first 4000 iterations (Fig. 2a). The confusion matrix illustrated that the Resnet50 model achieved a high accuracy of 96%

**Table 1** Preoperative and intraoperative clinical information of eyes by cohort

Baseline clinical and demographic patient data	Study cohorts			<i>p</i>	Statistical test
	Retrospective cohort	Prospective cohort 1	Prospective cohort 2		
Eyes, <i>n</i>	226	48	54		
Age, years	27.50 [22.00, 32.00]	28.00 [20.00, 34.00]	19.00 [18.00, 21.00]	0	Kruskal–Wallis test
Sex, <i>n</i> (%)				0.09	Chi-square test
<i>Female</i>	137 (60.6)	29 (60.4)	24 (44.4)		
<i>Male</i>	89 (39.4)	19 (39.6)	30 (55.6)		
Sphere, D	− 5.00 [− 6.00, − 4.00]	− 4.75 [− 5.56, − 3.94]	− 4.69 [− 5.75, − 2.75]	0.18	Kruskal–Wallis test
Cylinder, D	− 0.50 [− 1.00, − 0.25]	− 0.62 [− 1.00, − 0.25]	− 1.25 [− 1.69, − 0.75]	0	Kruskal–Wallis test
K1 power, D	42.85 [42.00, 43.58]	43.05 [41.60, 44.00]	42.30 [41.12, 43.00]	0.01	Kruskal–Wallis test
K2 power, D	43.90 [42.90, 44.70]	44.10 [42.98, 45.32]	43.90 [42.47, 45.08]	0.55	Kruskal–Wallis test
Axial length	25.19 [24.73, 25.82]	25.13 [24.55, 25.83]	25.70 [25.05, 26.17]	0.02	Kruskal–Wallis test
Corneal thickness, $\mu\text{m}$	539.00 [519.00, 554.00]	544.50 [522.75, 560.50]	546.50 [537.00, 566.50]	0.01	Kruskal–Wallis test
Cornea diameter, mm	11.60 [11.40, 11.80]	11.65 [11.40, 11.90]	11.60 [11.43, 11.80]	0.36	Kruskal–Wallis test
IOP, mmHg	15.50 [13.80, 17.28]	15.70 [14.55, 17.60]	17.00 [15.00, 19.00]	0.01	Kruskal–Wallis test
Suction time, mm	29.00 [28.00, 29.00]	29.00 [28.75, 29.00]	29.00 [28.00, 29.00]	0.11	Kruskal–Wallis test
Cap thickness, $\mu\text{m}$	120.00 [120.00, 120.00]	120.00 [120.00, 120.00]	120.00 [120.00, 127.50]	0	Kruskal–Wallis test
Cap diameter, mm	7.50 [7.40, 7.50]	7.50 [7.50, 7.50]	7.50 [7.50, 7.50]	0.1	Kruskal–Wallis test
Optical zone, mm	6.50 [6.40, 6.50]	6.50 [6.50, 6.50]	6.50 [6.50, 6.50]	0.05	Kruskal–Wallis test
RST	307.00 [289.00, 324.00]	307.50 [289.75, 325.00]	329.00 [297.50, 346.50]	0.01	Kruskal–Wallis test
Suction loss, <i>n</i> (%)				0.54	Chi-square test

**Table 1** continued

Baseline clinical and demographic patient data	Study cohorts			<i>p</i>	Statistical test
	Retrospective cohort	Prospective cohort 1	Prospective cohort 2		
<i>No</i>	224 (99.1)	47 (97.9)	54 (100.0)		
<i>Yes</i>	2 (0.9)	1 (2.1)	0 (0.0)		
OBL, <i>n</i> (%)				0.19	Chi-square test
<i>No</i>	202 (89.4)	40 (83.3)	51 (94.4)		
<i>Yes</i>	24 (10.6)	8 (16.7)	3 (5.6)		
BA, <i>n</i> (%)				0.25	Chi-square test
<i>No</i>	220 (97.3)	48 (100.0)	54 (100.0)		
<i>Yes</i>	6 (2.7)	0 (0.0)	0 (0.0)		
Postoperative sphere	0.04 [ − 0.50, 0.50]	− 0.15 [− 0.50, 0.25]	− 0.02 [− 0.50, 0.37]	0.37	Kruskal–Wallis test
Postoperative cylinder	− 0.16 [ − 0.25, 0.75]	− 0.29 [− 0.50, − 0.25]	− 0.25 [− 0.50, − 0.14]	0.31	Kruskal–Wallis test

Values in table are presented as the median with the interquartile range (IQR) in square brackets, or as a number with the percentage in parentheses

*BA* Black areas, *D* dioptre, *IOP* intra-ocular pressure, *IQR* interquartile range, *K* keratometric, *OBL* opaque bubble layer, *RST* residual stromal thickness

(Fig. 2b). The ROC curve and precision-recall curve showed that our model manifested an extraordinary performance in the training dataset (ROC-AUC [RAUC] = 0.962; precision-recall AUC [PAUC] = 0.964; Fig. 2c) and validation dataset (RAUC = 0.998; PAUC = 0.999, Fig. 2d). We used the two prospective SMILE cohorts as external datasets to test the generalization ability and robustness of the model. The ROC curve and precision-recall curve suggested that our model also performed well in the two prospective SMILE cohorts (RAUC and PAUC > 0.9; Fig. 2e, f). Subsequently, we compressed the histogram of image likelihood into 24 DL features (Electronic Supplementary Material [ESM] File 2). The Pearson correlation analysis resulted in the retention of 24 DL features for the LASSO-penalized feature selection. We discovered that the LASSO model had the

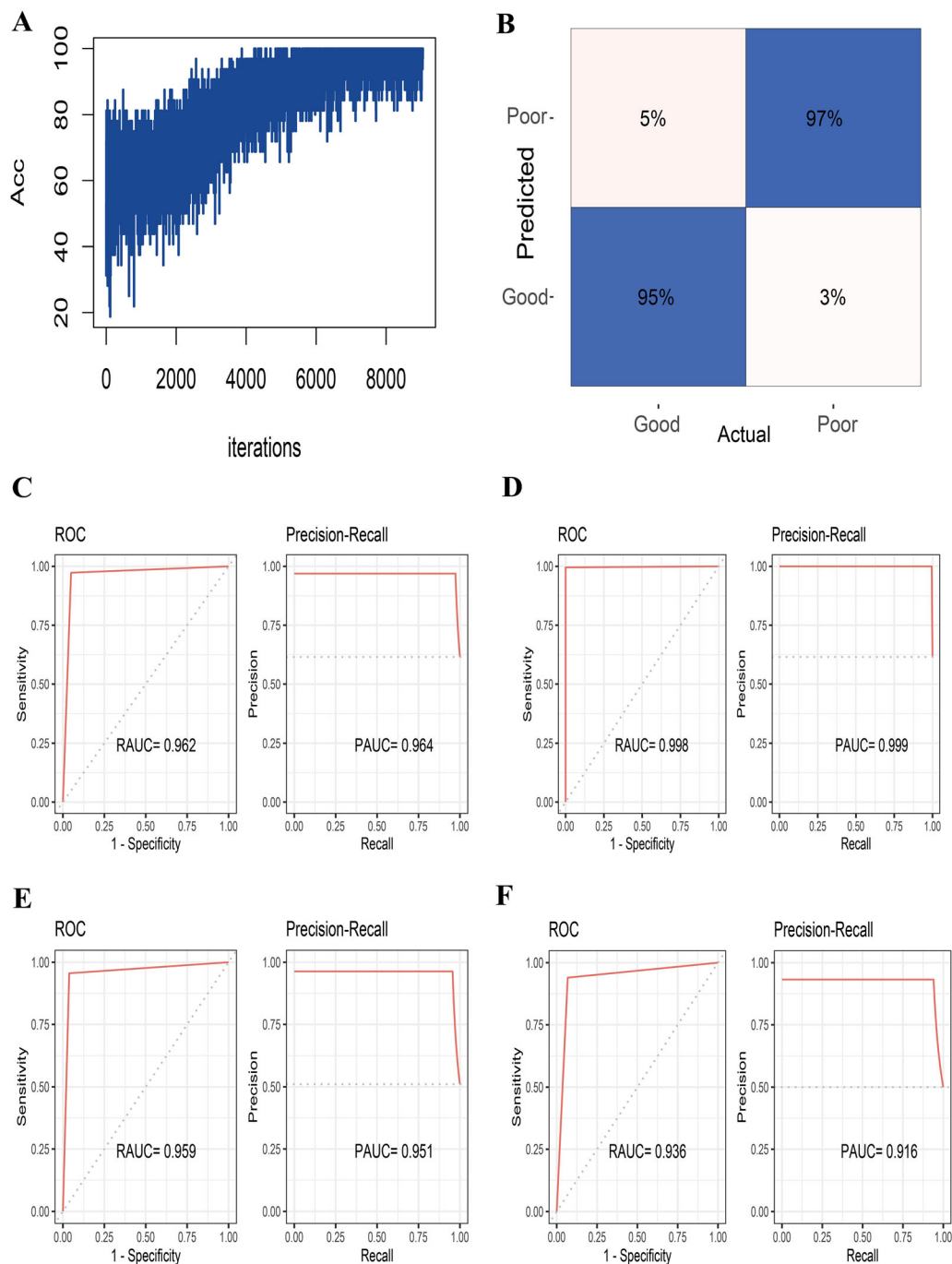
lowest mean squared error (MSE) when the penalization lambda was 0.039 (Fig. 3a). There were six DL features with coefficients greater than zero based on the lambda criteria (Fig. 3b). Finally, the LASSO-penalized model revealed six DL features, and their relative weight is shown in Fig. 3c. The six DL features were then transferred to seven machine learning models and evaluated using tenfold cross-validation. Based on the AUC distribution of these seven machine learning models, the SVM, XGBoost, and LightGBM methods were observed to have the highest values of AUC (Fig. 3d). The accuracy distribution further indicated that XGBoost manifested the best accuracy in the training and test datasets among the seven machine learning methods (Fig. 3e). Therefore, we selected the XGBoost model for video prediction. The predict probabilities for samples in the training and

**Table 2** Clinical characteristics of the training and validation dataset in the retrospective cohort

Clinical characteristics	Datasets		<i>p</i>	Statistical test
	Training	Validation		
<i>n</i>	159	67		
Age, years	29.00 [22.50, 32.00]	27.00 [21.50, 31.50]	0.52	Wilcox test
Sex, <i>n</i> (%)			0.08	Chi-square test
Female	90 (56.6)	47 (70.1)		
Male	69 (43.4)	20 (29.9)		
Sphere, D	− 5.00 [− 5.75, − 4.00]	− 5.25 [− 6.00, − 3.88]	0.33	Wilcox test
Cylinder, D	− 0.50 [− 1.00, − 0.25]	− 0.50 [− 0.75, − 0.25]	0.3	Wilcox test
K1, D	42.80 [42.00, 43.45]	43.00 [42.00, 43.70]	0.46	Wilcox test
K2, D	43.90 [42.90, 44.65]	43.80 [43.00, 44.95]	0.89	Wilcox test
Axial length	25.17 [24.74, 25.86]	25.21 [24.73, 25.74]	0.97	Wilcox test
Corneal thickness, $\mu\text{m}$	540.00 [519.50, 555.00]	535.00 [518.00, 551.50]	0.5	Wilcox test
Cornea diameter, mm	11.60 [11.40, 11.85]	11.60 [11.30, 11.75]	0.37	Wilcox test
IOP, mmHg	15.40 [14.10, 17.25]	15.50 [13.20, 17.20]	0.47	Wilcox test
Suction time, min	29.00 [28.00, 29.00]	29.00 [28.00, 29.00]	0.07	Wilcox test
Cap thickness, $\mu\text{m}$	120.00 [120.00, 120.00]	120.00 [120.00, 120.00]	0.05	Wilcox test
Cap diameter, mm	7.50 [7.40, 7.50]	7.50 [7.30, 7.50]	0.17	Wilcox test
Optical zone, mm	6.50 [6.40, 6.50]	6.50 [6.30, 6.50]	0.12	Wilcox test
RST	307.00 [290.00, 325.00]	307.00 [286.00, 321.00]	0.43	Wilcox test
Suction loss, <i>n</i> (%)			0.89	Chi-square test
<i>No</i>	157 (98.7)	67 (100.0)		
<i>Yes</i>	2 (1.3)	0 (0.0)		
OBL, <i>n</i> (%)			0.09	Chi-square test
<i>No</i>	138 (86.8)	64 (95.5)		
<i>Yes</i>	21 (13.2)	3 (4.5)		
BA, <i>n</i> (%)			0.8	Chi-square test
<i>No</i>	154 (96.9)	66 (98.5)		
<i>Yes</i>	5 (3.1)	1 (1.5)		

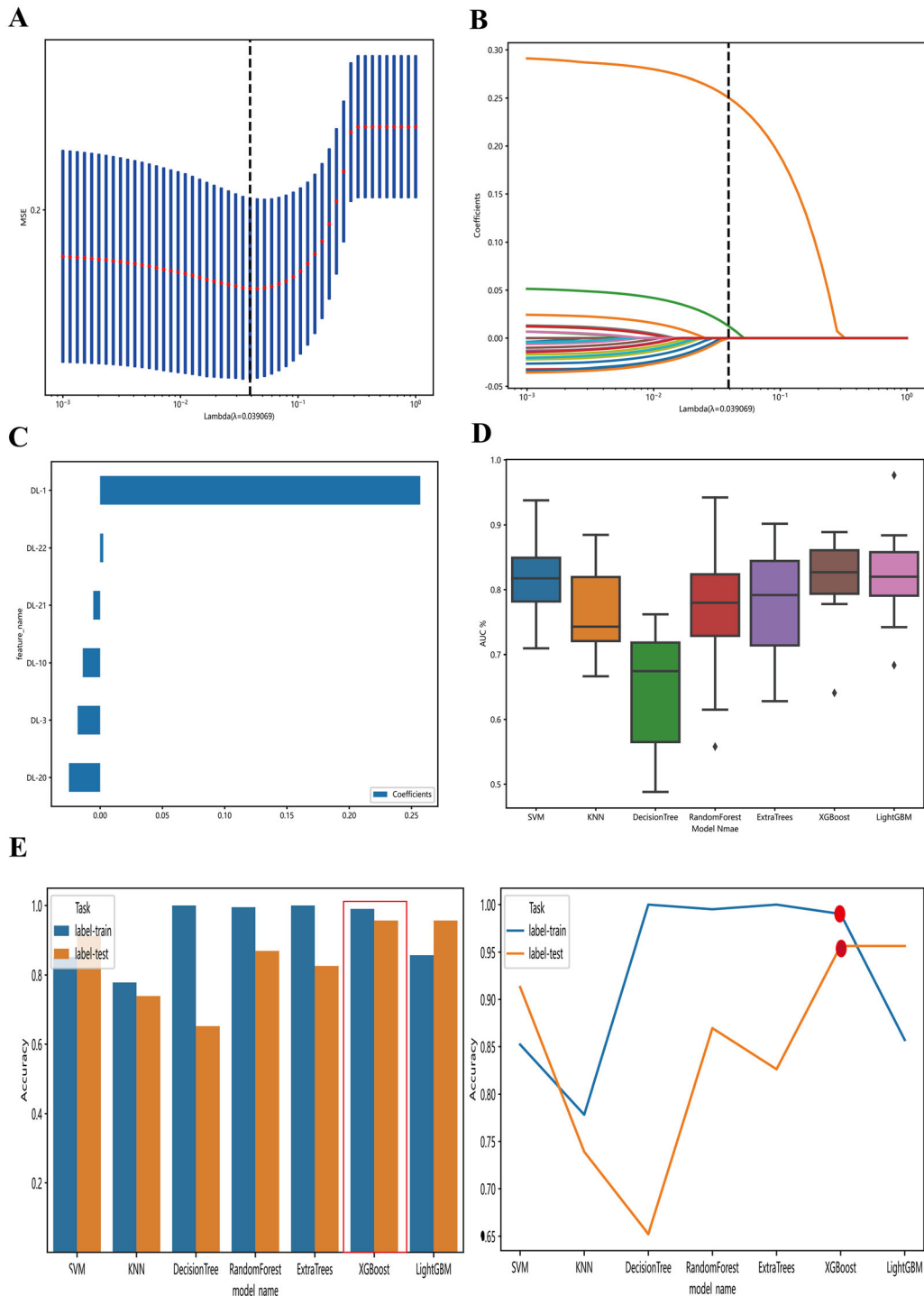
Values in table are presented as the median with the interquartile range (IQR) in square brackets, or as a number with the percentage in parentheses





**Fig. 2** The deep learning (DL) model (Resnet50) for image prediction. **a** Curve for accuracy distribution of model training. **b** Confusion matrix of DL model. **c** The receiver operating characteristic (ROC) curve and precision-recall curve in the training dataset of the retrospective cohort. **d** The ROC curve and precision-recall curve in the validation dataset of the retrospective cohort. **e** The ROC

curve and precision-recall curve for estimating the performance of the model in one of the prospective cohorts. **f** The ROC curve and precision-recall curve for estimating the performance of the model in the second prospective cohort. *RAUC* ROC-area under the curve, *PAUC* precision-recall AUC



**Fig. 3** Selection of the DL features. **a** Distribution of the lowest mean squared error (MSE) with the associated penalization lambda ( $\lambda$ ) value in the LASSO-penalized model. **b** The selected criterion of  $\lambda$  determine the LASSO coefficient profiles of all DL features. **c** Bar plot of coefficient profiles for the selected features in the LASSO-

penalized model. **d** Distribution of the AUCs of 7 machine learning classifiers (SVM, KNN, Decision-Tree, Random-Forest, Extra-Tree, XGBoost, and LightGBM) with tenfold cross-validation. **e** A accuracy distribution of 7 machine learning classifiers in the retrospective cohort

test datasets were uncovered in Fig. 4a and b, respectively. The AUC value of the XGBoost model for the training dataset was 0.998 (95% confidence interval [CI] 0.988–0.999) (Fig. 4c). The AUC value of the XGBoost model for the test dataset was 0.889 (95% CI 0.667–0.889) (Fig. 4d). The AUC values (1.000 and 0.984) in the two prospective SMILE cohorts (Fig. 4e, f) suggested that our XGBoost model also performed very well in terms of predicting in the prospective cohorts.

### Interpretation of Scanning Images

Gradient-weighted Class Activation Mapping (Grad-CAM) was able to calculate the key regions predicted by the model on the SMILE scanning images and therefore was utilized to visualize the heatmap of the model's final convolutional layer, overlaying it on the original images [16]. The red section that leads inwards to the blue section is active, suggesting that the model paid special attention to this region. We observed that our model can precisely concentrate on areas of scanning (Fig. 5). Regarding intraoperative complications, the model mainly focused on the edge of the femtosecond laser scanning for the opaque bubble layer (OBL), whereas for black spots (BS) our model played particular attention to the central area of the femtosecond laser scanning (Fig. 6). Furthermore, we used Image-J software to calculate the percentage of BS area in the femtosecond laser scanning. The region of interest BS is visualized in Fig. 6b. The Wilcoxon test determined that the proportion of region of BS in the poor visual group was larger than that in the good visual group (Fig. 6b).

### Distinguishing SMILE Intraoperative Complications

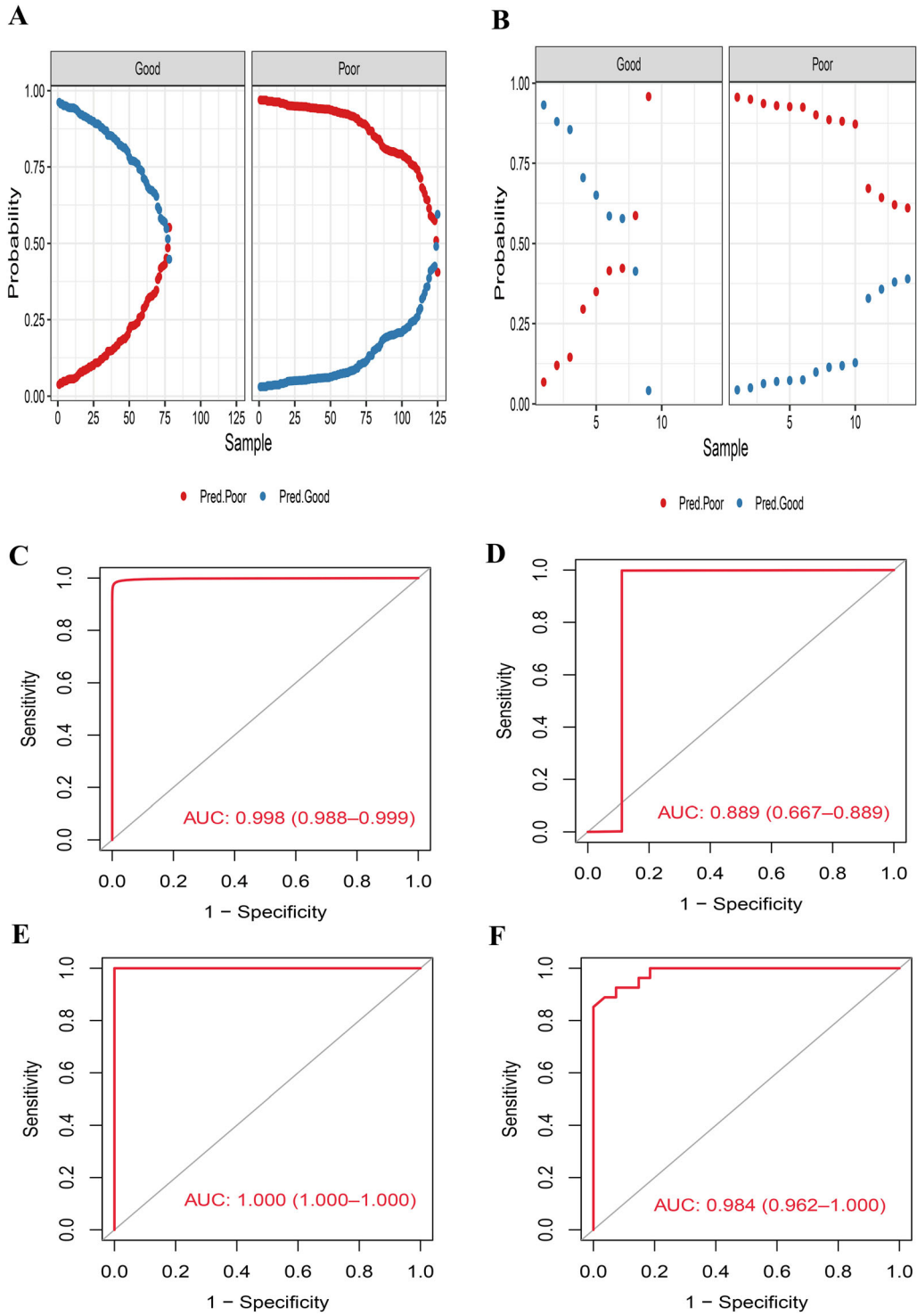
The intraoperative complications which occurred during SMILE surgery include suction loss (LS), OBL and BS. We first collected intraoperative images that were associated with LS ( $n = 120$ ), OBL ( $n = 90$ ), BS ( $n = 150$ ), and normal scanning (Norm:  $n = 150$ ). Subsequently, we randomly split these images into training

and test sets at a 7:3 ratio and trained a Resnet50 model for 50 epochs. Details on the parameters (accuracy, AUC, sensitivity, specificity, positive predictive values, negative predictive values, [recision, recall) for assessment of model are listed in ESM File 3. Grad-CAM was applied to visualize the heatmap and superimposed on the four categories (Fig. 6c). The confusion matrix heatmap illustrated that our model was able to accurately distinguish LS (100%), OBL (96%), BS (85%), and Norm (97%) (Fig. 6c).

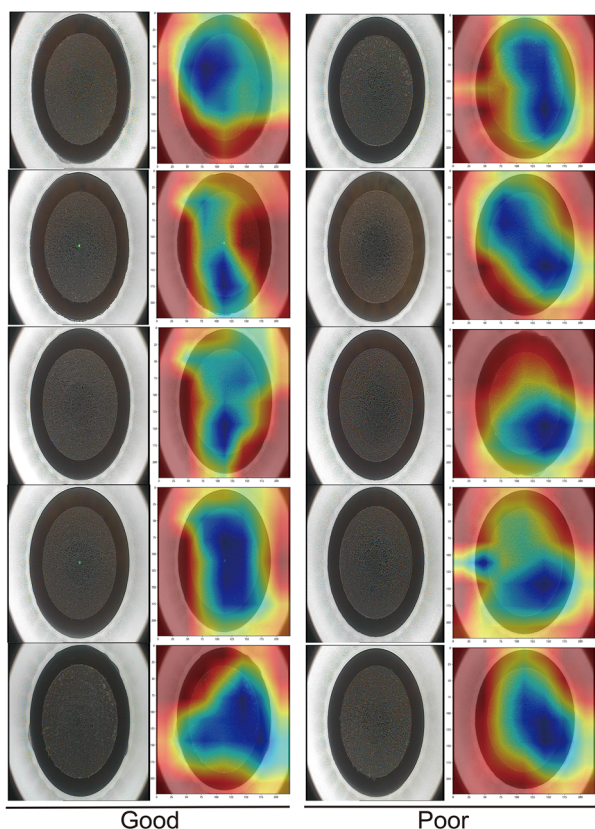
## DISCUSSION

Following the encouraging outcomes of several prospective trials on SMILE surgery and the appearance of recent publications revealing that the visual and refractive corrections achieved with SMILE are comparable to those achieved with femtosecond laser-assisted in situ keratomileusis (FS-LASIK), SMILE refractive surgery has grown in popularity [6, 17]. SMILE surgery also achieves better biomechanical strength and stability characteristics than FS-LASIK [18, 19]. However, a small proportion of patients who have undergone SMILE surgery report delayed visual acuity recovery postoperatively. This delay appears to be particularly evident in patients with similar refractive errors in eyes that underwent SMILE refractive surgery, while the postoperative visual acuity recovery in each eye was totally different. To our knowledge, the present study is the first to develop a DL model for predicting the outcome of SMILE postoperative visual acuity. The results of this work reveal that the developed DL model can predict images and videos with a high accuracy. It is extensively used in clinical practice, allowing every surgeon with a femtosecond laser scanning image or video to acquire an estimated prognosis of each patient.

The formation of the intrastromal lenticule is essential for the safety and predictability of the SMILE procedure [20]. Because the lenticule is exclusively manufactured by the femtosecond laser, femtosecond laser-related intraoperative problems, such as BS and OBL, are unavoidable. BS are defined as several scattered little black



**Fig. 4** The machine learning model (XGBoost) for video prediction. **a, b** XGBoost model was used to calculate the probabilities of patients in the training (**a**) and test (**b**) dataset that was derived from the retrospective cohort. Colored dots indicate each patient, with the difference colors representing the different probability belonging to the subclass. Red dots represent the probability of being in the poor recovery group; blue dots represent the probability of being in the good recovery group. **c** ROC curve and corresponding AUC value of the model in the training dataset. **d** ROC curve and corresponding AUC value of the model in the test dataset. **e** ROC curve for estimating the performance of the model in 1 of the prospective cohorts. **f** ROC curve for estimating the performance of the model in the second prospective cohort

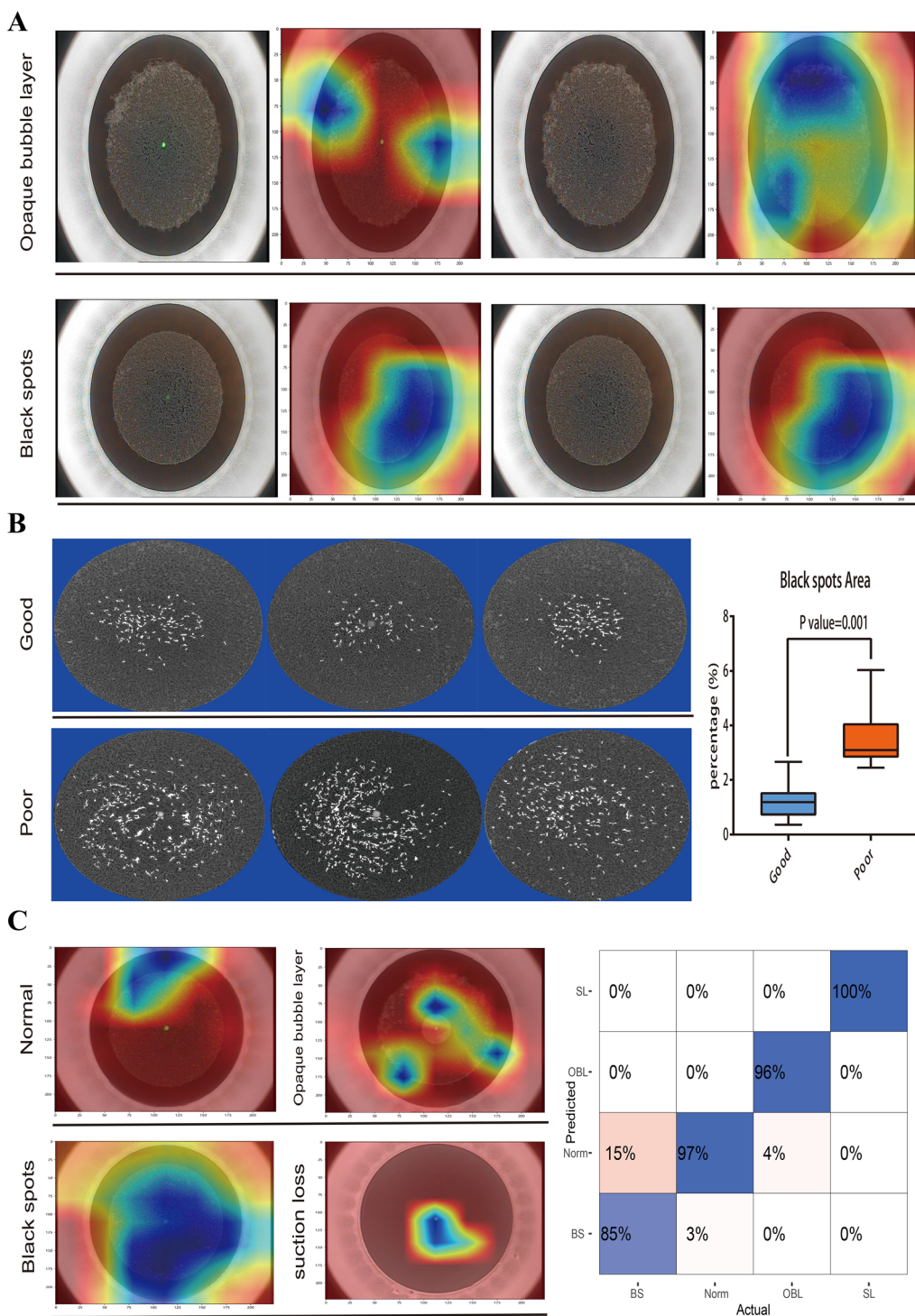


**Fig. 5** The original femtosecond laser images and the corresponding Gradient-weighted Class Activation Mapping (Grad-CAM) heatmaps highlight important regions for the model prediction

dots in the stroma for complete photo disruption following femtosecond laser; in contrast, black areas or black islands occur as patchy or

strips formations. Due to the block of debris, such as foreign bodies, secretions of meibomian, and mucus of the conjunctival at the interface, the black areas or black islands generally format at both the anterior and posterior lenticule for incomplete photo disruption [21–23]. In our work, we observed that BS occurred in almost all SMILE videos, mainly in the posterior lenticule. However, the incidence of black areas in our study eyes was 2.7% (ESM File 1). The incidence of BS determined in our study differs greatly from that reported in earlier studies, but the incidence of black areas is consistent with earlier studies, ranging from 0.33% to 11% [9, 24, 25]. Therefore, we believe that the terms ‘black spots,’ ‘black areas,’ or ‘black islands’ may not have been used consistently in previous publications [26, 27].

Using the DL model, we can easily assess the quality of femtosecond laser images and predict the postoperative visual acuity at 24 h. To visualize the attention mechanisms of DL, we used the Grad-CAM heatmap to calculate the key regions of the SMILE scanning images. The heatmaps indicated that our model pays special attention to the region of femtosecond laser scanning (Fig. 5) and, in particular, to the active region (blue color) where it pays more attention to BS. Therefore, we calculated the percentage of area occupied by black spots using Image-J software and found that the size of BS ranged from 20 to 100 pixels. Moreover, the area percentage of BS in the poor visual acuity group ( $3.49 \pm 1.08\%$ ) was larger than that in the good visual acuity group ( $1.23 \pm 0.68\%$ ). None of the BS in our study were associated with difficulties of intrastromal lenticule separation and extraction and did not lead to impairment of visual outcomes. Hence, we recognized that the BS are the tissue bridges between laser spots. This phenomenon was indirectly confirmed by Lin et al. who estimated the black areas of four levels of low laser energy for SMILE surgery and found that the lowest energy used had the largest area of BS [24]. The surface quality of scanning may be mainly determined by two factors: (1) the distance of laser spots and the delivered energy of per pulse; and (2) tissue reaction for photo disruption [28, 29]. The surface becomes smoother when the laser spots



**Fig. 6** Interpretation of scanning images. **a** Example images for opaque bubble layer (OBL), black spots (BS) and corresponding Grad-CAM heatmaps. **b** Quantitative measurement for percentage area of BS. **c** Grad-CAM

heatmaps for interoperative complications, such as OBL, BS and suction loss (SL), and the confusion matrix heatmap of the model for distinguishing interoperative complications. *Norm* Normal scanning

adjoin more closer. The same spot distance of 4.5 mm and pulse frequency of 500 kHz were utilized in our study. A lower femtosecond laser energy may be inadequate to disturb the deep corneal stroma and result in BS in the posterior lenticule.

We also found that the poor scanning group had a high incidence of OBL (12.6%) compared to good scanning group (8.0%). The OBL all occurred at the peripheral. Although OBL at the edge of lenticule does not influence the final visual outcome, their presence may make it difficult to separate and extract the lenticule, cause transient corneal edema and delay early recovery of postoperative visual acuity [30–33]. The Grad-CAM heatmap suggested that our developed DL model can accurately recognize the position of OBL and make correct predictions (Fig. 6a). The good performance of our model for distinguishing OBL and BS led us to speculate whether a Resnet50 model could be trained to identify SMILE intraoperative complications. The results shown here indicate that the Resnet50 DL model also has a high accuracy to distinguish intraoperative complications. However, despite SMILE providing a promising performance for the correction of myopia and myopic astigmatism, intraoperative problems are unavoidable. Therefore, our DL model can timely and correctly remind the surgeon of the correct management strategies for dealing with intraoperative complications during the operation. The Grad-CAM heatmap also revealed various features to be alert for regarding these intraoperative complications, which may display a great importance for artificial intelligence in the application of refractive surgeries.

Our study's main limitation is the small sample size and short follow-up period. However, in addition, this study only utilized SMILE videos and did not make use of support from other aspects, such as clinical records and other physical examinations. As a result, we were unable to perform multi-dimensional evaluation, thus limiting the accuracy of the results in real-world settings. Therefore, future research

should focus on expanding cohort size, prolonging the follow-up duration, and including multimode data to increase the accuracy of the DL models.

## CONCLUSIONS

Overall, we created a DL model for predicting the early visual outcomes based on SMILE scanning images. Using these images, it is now feasible to distinguish the intraoperative complications of the SMILE procedure than has been previously reported.

## ACKNOWLEDGEMENTS

We thank the Onekey AI platform for code supporting some of the experiments in this study.

**Funding.** This work was supported by the Science & Technology Department of Sichuan Province (China) funding project (No. 2021YFS0221) and 1.3.5 project for disciplines of excellence, West China Hospital, Sichuan University (No. 2022HXFH032, ZYJC21058, 2021-023,2022-014).

**Author Contributions.** QW: implemented the analysis and wrote the original manuscript. RW, SLY, JT: reviewed the articles of interest. HBY, JT, KM, YPD: conceived the idea for this paper. All the authors commented and approved the text.

**Disclosures.** Qi Wan, Shali Yue, Jing Tang, Ran Wei, Jing Tang, Ke Ma, Hongbo Yin, and Ying-ping Deng have nothing to disclose.

**Compliance with Ethics Guidelines.** The study was carried out in accordance with the principles of the 1964 Helsinki Declaration and its later amendments and was approved by the Ethics Committee of West China Hospital.

Before enrolling in the study, each subject provided written informed consent.

**Data Availability.** The datasets used in the current study are available from the corresponding author on reasonable request.

**Open Access.** This article is licensed under a Creative Commons Attribution-Non-Commercial 4.0 International License, which permits any non-commercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc/4.0/>.

## REFERENCES

- Huang G, Melki S. Small incision lenticule extraction (smile): myths and realities. *Semin Ophthalmol*. 2021;36:140–8.
- Ahmed AA, Hatch KM. Advantages of small incision lenticule extraction (SMILE) for mass eye and ear special issue. *Semin Ophthalmol*. 2020;5:224–31.
- Blum M, Lauer AS, Kunert KS, Sekundo W. 10-Year results of small incision lenticule extraction. *J Refract Surg*. 2019;35:618–23.
- Chansue E, Tanesakdi M, Swasdibutra S, McAlinden C. Efficacy, predictability and safety of small incision lenticule extraction (SMILE). *Eye Vis (Lond)*. 2015;2:14. <https://doi.org/10.1186/s40662-015-0024-4>.
- Chansue E, Tanesakdi M, Swasdibutra S, McAlinden C. Efficacy, predictability and safety of small incision lenticule extraction (smile). *Eye Vis (Lond)*. 2015;2:14.
- Ganesh S, Gupta R. Comparison of visual and refractive outcomes following femtosecond laser-assisted Lasik with SMILE in patients with myopia or myopic astigmatism. *J Refract Surg*. 2014;30:590–6.
- Krueger RR, Meister CS. A review of small incision lenticule extraction complications. *Curr Opin Ophthalmol*. 2018;29:292–8.
- Hamed AM, Abdelwahab S, Soliman TT. Intraoperative complications of refractive small incision lenticule extraction in the early learning curve. *Clin Ophthalmol*. 2018;12:665–8.
- Titiyal JS, Kaur M, Rathi A, et al. Learning curve of small incision lenticule extraction: challenges and complications. *Cornea*. 2017;36:1377–82.
- Hamed AM, Heikal MA, Soliman TT, et al. Smile intraoperative complications: incidence and management. *Int J Ophthalmol*. 2019;12(2):280–3.
- Gehrunge M, Berman AG, O'Donovan M, Fitzgerald RC, Markowitz F. Triage-driven diagnosis of Barrett's esophagus for early detection of esophageal adenocarcinoma using deep learning. *Nat Med*. 2021;27(5):833–41.
- Pan Y, Lei X, Zhang Y. Association predictions of genomics, proteomics, transcriptomics, microbiome, metabolomics, pathomics, radiomics, drug, symptoms, environment factor, and disease networks: a comprehensive approach. *Med Res Rev*. 2022;42(1):441–61.
- Ayyaz MS, Lali MIU, Hussain M, et al. Hybrid deep learning model for endoscopic lesion detection and classification using endoscopy videos. *Diagnostics (Basel)*. 2021;12(1):43.
- Zhang C, Liu D, Huang L, et al. Classification of thyroid nodules by using deep learning radiomics based on ultrasound dynamic video. *J Ultrasound Med*. 2022;41:2993–3002.
- Cao R, Yang F, Ma SC, et al. Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer. *Theranostics*. 2020;10:11080–91.



16. Selvaraju RR, Cogswell M, Das A, et al. Grad-CAM: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis*. 2020;128(2):336–59.
17. Li M, Zhou Z, Shen Y, et al. Comparison of corneal sensation between small incision lenticule extraction (SMILE) and femtosecond laser-assisted Lasik for myopia. *J Refract Surg*. 2014;30(2):94–100.
18. Wang D, Liu M, Chen Y, et al. Differences in the corneal biomechanical changes after smile and Lasik. *J Refract Surg*. 2014;30(10):702–7.
19. Wu D, Wang Y, Zhang L, et al. Corneal biomechanical effects: small-incision lenticule extraction versus femtosecond laser-assisted laser in situ keratomileusis. *J Cataract Refract Surg*. 2014;40(6):954–62.
20. Barrio JLA, Vargas V, Al-Shymali O, Ali JL. Small incision lenticule extraction (SMILE) in the correction of myopic astigmatism: outcomes and limitations—an update. *Eye Vis (Lond)*. 2017;4:26.
21. Ma J, Wang Y, Chan T. Possible risk factors and clinical outcomes of black areas in small-incision lenticule extraction. *Cornea*. 2018;37(8):1035–41.
22. Krueger RR, Meister CS. A review of small incision lenticule extraction complications. *Curr Opin Ophthalmol*. 2018;29(4):292–8.
23. Wang M, Chen Y, Zhang Y. Laser-assisted subepithelial keratomileusis substituted for an aborted small-incision lenticule extraction due to large black area formation. *J Cataract Refract Surg*. 2020;46(6):913–7.
24. Lin L, Weng S, Liu F, et al. Development of low laser energy levels in small-incision lenticule extraction: clinical results, black area, and ultrastructural evaluation. *J Cataract Refract Surg*. 2020;46(3):410–8.
25. Asif MI, Mehta JS, Reddy J, Titiyal JS, Maharana PK, Sharma N. Complications of small incision lenticule extraction. *Indian J Ophthalmol*. 2020;68(12):2711–22.
26. Qiu P, Yang Y. Analysis and management of intraoperative complications during small-incision lenticule extraction. *Int J Ophthalmol*. 2016;9(11):1697–700.
27. Wang Y, Ma J, Zhang L, et al. Postoperative corneal complications in small incision lenticule extraction: long-term study. *J Refract Surg*. 2019;35(3):146–52.
28. Kunert KS, Blum M, Duncker GIW, et al. Surface quality of human corneal lenticules after femtosecond laser surgery for myopia comparing different laser parameters. *Graefes Arch Clin Exp Ophthalmol*. 2011;249(9):1417–24.
29. Kymionis GD, Kankariya VP, Plaka AD, Reinstein DZ. Femtosecond laser technology in corneal refractive surgery: a review. *J Refract Surg*. 2012;28(12):912–20.
30. Ma JN, et al. The effect of corneal biomechanical properties on opaque bubble layer in small incision lenticule extraction (SMILE). *Chinese J Ophthalmol*. 2019;55(2):115–21.
31. Jung H-G, Kim J, Lim T-H. Possible risk factors and clinical effects of an opaque bubble layer created with femtosecond laser-assisted laser in situ keratomileusis. *J Cataract Refract Surg*. 2015;41(7):1393–9.
32. Liu C, Sun C, Hui-Kang Ma D, et al. Opaque bubble layer: incidence, risk factors, and clinical relevance. *J Cataract Refract Surg*. 2014;40:435–40.
33. Son G, Lee J, Jang C, et al. Possible risk factors and clinical effects of opaque bubble layer in small incision lenticule extraction (SMILE). *J Refract Surg*. 2017;33:24–9.