Special Issue

Keratoconus Screening Based on Deep Learning Approach of Corneal Topography

Bo-I Kuo[1,2], Wen-Yi Chang[3], Tai-Shan Liao[4], Fang-Yu Liu[1,5], Hsin-Yu Liu[1], Hsiao-Sang Chu[1], Wei-Li Chen[1,6], Fung-Rong Hu[1], Jia-Yush Yen[7], and I-Jong Wang[1,6]

[1] Department of Ophthalmology, National Taiwan University Hospital, Taipei, Taiwan
[2] Department of Ophthalmology, Taipei City Hospital, Renai branch, Taipei, Taiwan
[3] National Center for High-Performance Computing, National Applied Research Laboratories, Hsinchu, Taiwan
[4] Taiwan Instrument Research Institute, National Applied Research Laboratories, Hsinchu, Taiwan
[5] Department of Ophthalmology, Tri-Service General Hospital, National Defense Medical Center, Taipei, Taiwan
[6] Graduate Institute of Clinical Medicine, College of Medicine, National Taiwan University, Taipei, Taiwan
[7] Department of Mechanical Engineering, National Taiwan University, Taipei, Taiwan

Purpose: To develop and compare deep learning (DL) algorithms to detect keratoconus on the basis of corneal topography and validate with visualization methods.

Methods: We retrospectively collected corneal topographies of the study group with clinically manifested keratoconus and the control group with regular astigmatism. All images were divided into training and test datasets. We adopted three convolutional neural network (CNN) models for learning. The test dataset was applied to analyze the performance of the three models. In addition, for better discrimination and understanding, we displayed the pixel-wise discriminative features and class-discriminative heat map of diopter images for visualization.

Results: Overall, 170 keratoconus, 28 subclinical keratoconus and 156 normal topographic pictures were collected. The convergence of accuracy and loss for the training and test datasets after training revealed no overfitting in all three CNN models. The sensitivity and specificity of all CNN models were over 0.90, and the area under the receiver operating characteristic curve reached 0.995 in the ResNet152 model. The pixel-wise discriminative features and the heat map of the prediction layer in the VGG16 model both revealed it focused on the largest gradient difference of topographic maps, which was corresponding to the diagnostic clues of ophthalmologists. The subclinical keratoconus was positively predicted with our model and also correlated with topographic indexes.

Conclusions: The DL models had fair accuracy for keratoconus screening based on corneal topographic images. The visualization mentioned in the current study revealed that the model focused on the appropriate region for diagnosis and rendered clinical explainability of deep learning more acceptable.

Translational Relevance: These high accuracy CNN models can aid ophthalmologists in keratoconus screening with color-coded corneal topography maps.

Introduction

Keratoconus is a noninflammatory corneal disease characterized by stromal thinning, anterior protrusion, and irregular astigmatism.[1] It typically manifests at puberty and progresses until the fourth decade of life, resulting in irreversible vision loss.[2] Unfortunately, the etiology and causes of keratoconus have still not been elucidated even though several theories and genetic factors have been proposed.[3] Furthermore, no effective treatment modalities are yet available to prevent the progression of early keratoconus or forme fruste keratoconus.[4,5] Therefore it is crucial to detect the

keratoconus or forme fruste keratoconus (subclinical keratoconus) earlier and subsequently prevent possible risk factors from its progression.[5] In addition, early detection may help us understand the natural course of keratoconus.[5,6]

Notably, advanced keratoconus can be diagnosed based on classic clinical signs, such as Munson's sign, Vogt's striae, and Fleischer's ring under slit-lamp examination; however, these signs are not present in the early stage of keratoconus.[7] Nonetheless, emerging technology might aid in the early diagnosis of keratoconus.[8] Among these, computerized corneal videokeratography (topography) is the most often used tool to detect the patterns of corneal topography and keratometric parameters of keratoconus, is useful in the early detection, and enables to discern its natural and treatment courses.[9] Several indexes of corneal topography have been developed to differentiate between keratoconus and normal eyes, such as the zone of increasing corneal power, inferior-superior (I-S) corneal power asymmetry, and skewing of the steepest radial axes.[10] However, the large numbers and complexity of indexes provided by videokeratographs pose a clinical challenge to ophthalmologists.[11]

Recently, compared with conventional techniques, deep learning (DL) has been shown to achieve significantly higher accuracies in several domains, including natural language processing,[12] computer vision,[13] and voice recognition.[14] The use of artificial intelligence (AI; machine intelligence) or DL has been primarily applied in medical imaging analysis, wherein DL systems have exhibited robust diagnostic performance in detecting various ocular imaging, principally fundus photographs and optical coherence tomography (OCT),[15] in diagnosing or screening diabetic retinopathy (DR),[16] glaucoma,[17] age-related macular degeneration (AMD),[18] and retinopathy of prematurity (ROP).[19] Notably, several methods have been used for automatic diagnosis of ectatic corneal disorders by using corneal topography, such as the discriminant analysis,[6] and then a neural network approach.[20] These approaches achieved a global sensitivity of 94.1% and a global specificity of 97.6% (98.6% for keratoconus alone) in the test set.[21] Hwang et al.[22] used multivariate logistic regression analysis to tease out those variables available from slit-scan tomography and spectral-domain OCT for clinically normal fellow eyes of highly asymmetric keratoconus patients. Similarly, AI has been used to improve the detection of corneal ectasia susceptibility by using tomographic data with the random forest (RF) that provided the highest accuracy among AI models in this sample with 100% sensitivity for clinical ectasia and was named the Pentacam Random Forest Index (PRFI).[23] The PRFI had an area under the curve (AUC) of 0.992 (sensitivity = 94.2%, specificity = 98.8%, and cut-off = 0.216), statistically higher than the Belin-Ambrósio deviation (BAD-D; AUC = 0.960, sensitivity = 87.3%, and specificity = 97.5%).[23]

Automatic analysis with these topographic parameters through the use of machine learning could provide reasonably accurate clinical diagnosis.[6,20] However, the interpretation of color-coded corneal maps and the patterns of subtle features in the image might prove challenging to apply to classical machine learning owing to a large amount of information required for processing. In contrast to classical machine learning, the convolutional neural network (CNN) exhibited its ability in recognizing images without the need for indices input for training, thereby exhibiting suitability for pattern recognition of color-coded corneal topography maps. In the current study, we sought to take advantage of DL in pattern recognition to detect corneal topographic pattern differences between the normal population and patients with keratoconus who present a variety of topographic patterns. We suggest using DL with a CNN model to diagnose keratoconus with the pattern of color-coded corneal topography, which has subtle features and large amount of information, instead of complex topographic indexes.

Methods

This retrospective study was conducted at the Department of Ophthalmology, National Taiwan University Hospital (NTUH), and was approved by the Institutional Review Board of NTUH as per the tenets of the Declaration of Helsinki.

Study Population and Examinations

The patients were selected from the corneal clinic of the Department of Ophthalmology in NTUH. There were four corneal specialists involved with the diagnosis and evaluation of the keratoconus and the candidates for refractive surgery from 2007 to 2020. Initially, total 299 patients were included. Patients with previous eye surgery, ocular trauma, contact lens discontinuation less than three weeks and younger than 20 years old were excluded. In the end, 206 patients were enrolled, and a total of 354 images were used in this study. All patients underwent a comprehensive ophthalmologic examination. The study group comprised clinically manifested keratoconus eyes and subclinical keratoconus eyes. Criteria for diagnosing manifested keratoconus were defined per clinical findings and

topographic criteria previously described.[8,9] Clinical signs of keratoconus were the existence of central protrusion of the cornea, Fleischer's ring, Vogt's striae, and focal corneal thinning on slit-lamp examination. Topographic criteria were central K value >47 diopter,[24] I-S value >1.4 diopter,[25] keratoconus percentage index (KISA%) >100%,[26] and asymmetric bowtie presentation.[27] The main criteria of subclinical keratoconus were basically based on the topographic pattern. Asymmetric bowtie with skewed radial axes (AB/SRAX), asymmetry bowtie with inferior steep (AB/IS) and symmetric bowtie with skewed radial axes (SB/SRAX) presented in the topography and no slit-lamp keratoconus findings were collected.[27–29] The best spectacle-corrected vision of subclinical keratoconus patients was not affected. The control group comprised candidates for refractive surgery without any manifestations earlier, with regular astigmatism. All corneal topographic maps were obtained using a computer-assisted videokeratoscope (TMS-4; Tomey Corporation, Nagoya, Japan). The tomographic and tonometry indices were acquired from the Pentacam HR Scheimpflug tomography system (Oculus Optikgeräte GmbH, Wetzlar, Germany) and Corvis ST (Oculus Optikgeräte GmbH, Wetzlar, Germany), respectively.

Overall, 354 images were available, including 170 keratoconus, 28 subclinical keratoconus and 156 normal topographic pictures. The topographies were divided into the following three datasets: (a) training, (b) test, and (c) subclinical test datasets. We used 134 keratoconus and 120 normal images in the training dataset and 36 keratoconus and 36 normal images in the test dataset. Twenty-eight subclinical keratoconus cases were assigned to the subclinical test datasets. We used only one image from each patient and assigned them as subclinical keratoconus test dataset. The test and subclinical test dataset did not involve the training process.

Deep Learning Architecture and Visualization

The proposed methods for the keratoconus detection were based on CNN.[12] Three well-known CNN models were adopted for transfer learning and analysis, namely the VGG16 model,[30] InceptionV3 model,[31] and ResNet152 model.[32] The VGG16 model was mainly stacked with a series of convolutional and pooling layers for image feature extraction and then connected to the fully-connected layer for classification, which was viewed as the extension of the classic AlexNet model.[33] The InceptionV3 model reorganized
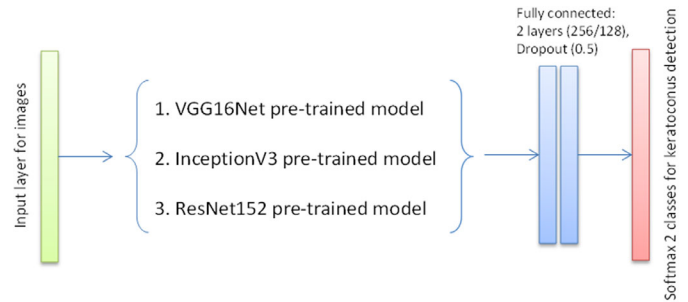

Figure 1. Architecture of the present CNNs for keratoconus binary classification.

the common convolutional and pooling layers into the so-called Inception module, which comprised three convolutions and one pooling to widen the network layer to acquire more detailed image features and improve prediction accuracy. By contrast, the ResNet152 model introduced the shortcut connection before and after convolutional layers to make a plain network into a residual network, which can build an ultra-deep network without problems of gradient vanishing or exploding to gain accuracy from the considerably increased depth. In this study, the three CNN models were adopted as pretrained models for transfer learning separately, wherein the pretrained model weights were downloaded from ImageNet (http://www.image-net.org/). Figure 1 illustrates the architecture of present CNNs for the keratoconus binary classification. The input layer was for images of corneal topography that were $128 \times 128 \times 3$, $296 \times 240 \times 3$, and $296 \times 240 \times 3$ in dimensions, respectively. The last two dense layers were redefined for fine-tuning, keeping all original model layers unchanged. For the training process, training data were augmented using the random shift and rotation of 5% or less and a random horizontal flip. For all cases, training tasks were conducted for more than 300 epochs with the Adam optimizer, with a learning rate of 0.00001 and a dropout rate of 0.5 implemented in the last fully connected layers. After the training was completed, test datasets were applied to analyze the performance of CNN. For accessing the predictability in subclinical keratoconus cases of the CNN models, subclinical test datasets were applied as well.

Once the CNN models were trained and validated, it was essential to evaluate and analyze the performance of the network through visual explanations. In general, visualization can be achieved by reconstructing pixel-wise images from high-level feature maps through Deconvnet.[34] Moreover, to further localize class-discriminative regions, Grad-CAM[35] used the gradient information of the last convolutional layer
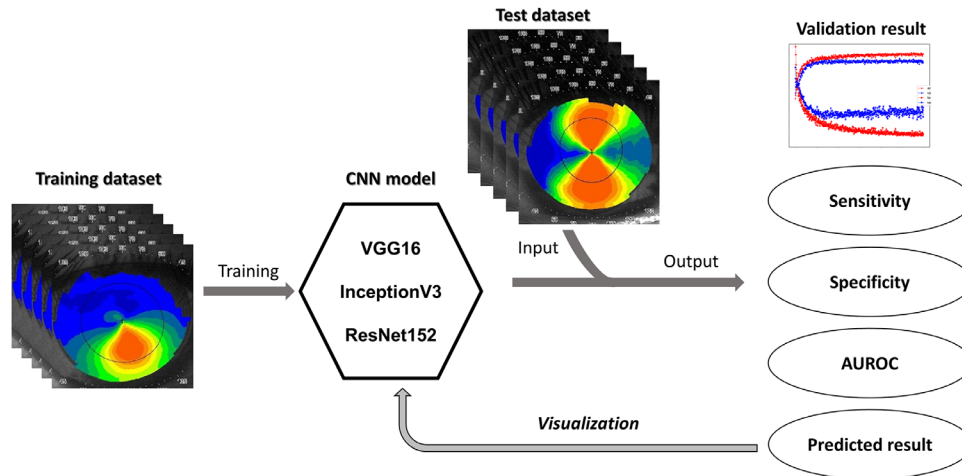
Figure 2. The brief procedure of our proposed model.

to determine the weights of feature maps. Hence, the pixel-wise discriminative features, as well as the class-discriminative heat map of diopter images, can be displayed. The brief procedure of the present study is presented in Figure 2.

Statistical Analysis

We evaluated the performance of our DL algorithm. The results of the training process were defined as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The methods to access the algorithm were as below:

- Accuracy: (TP+TN)/(TP+TN+FP+FN)
- Sensitivity: TP/(TP+FN)
- Specificity: TN/(TN+FP)
- Area under the receiver operating characteristic curve (AUROC): the area under the cumulative distribution function of the sensitivity on the y-axis versus the cumulative distribution function of (1-specificity) on the x-axis.

For demographic data, the *t*-test was used to compare continuous data, and the χ^2 test was used to compare categorical data between the study and control groups. A *P* value < 0.05 was considered statistically significant. Pearson correlation coefficient was used to estimate the correlation between the probability of keratoconus with topographic, tomographic and tonometry indices. All statistical analyses were performed using SPSS software (SPSS 24.0; SPSS Inc., Chicago, IL, USA).

Results

The keratoconus group comprised 94 patients, 64 men and 30 women, with the mean age of 29.78 ± 9.23 years. The control group comprised 84 patients, 46 men and 38 women, who were candidates of refractive surgery, with the mean age of 25.49 ± 9.26 years. The keratoconus group was older than the control group, whereas no significant differences were noted regarding sex distribution (Table 1). Descriptive statistics of topographic parameters between the two groups are presented in Table 2. All indexes in the keratoconus group were significantly different from the control group except for the cylinder index in the left eye.

All the three CNN models had fair accuracy after training, and no signs of overfitting were noted when we applied the test dataset (Fig. 3). The accuracy, sensitivity, specificity, and AUROC are presented in Table 3. Regarding accuracy, VGG16 and InceptionV3 were both 0.931, and ResNet152 was 0.958. All the three models had acceptable accuracy for keratoconus prediction. The sensitivity was 0.917 in VGG16 and InceptionV3 and 0.944 in ResNet152. The specificity of VGG16 and InceptionV3 was 0.944 and that of the ResNet152 model was 0.972. Figure 4 displays the AUROC of three models, 0.956 in VGG16, 0.987 in InceptionV3, and 0.995 in ResNet152. In our dataset, all models reached acceptable performance, with the ResNet152 being the best.

Taking VGG16 as an example, we demonstrated the results of DL classification in Figure 5. The images shown in Figures 5A and 5B were clinically diagnosed keratoconus, and the possibility of keratoconus

Table 1. Baseline Characteristics of the Keratoconus Group and the Control Group

	Keratoconus Group (n = 94)	Control Group (n = 84)	*P* Value
Age (years)	29.78 ± 9.23	25.49 ± 9.26	0.023
Gender			0.09
Male	64	46	
Female	30	38	

Table 2. Topographic Parameters of the Keratoconus Group and the Control Group

	Keratoconus Group (n = 94)	Control Group (n = 84)	*P* Value
AveK (D)			
OD	46.15 ± 3.88	44.43 ± 1.93	< 0.001
OS	46.30 ± 4.48	46.30 ± 4.48	< 0.001
Cyl (D)			
OD	4.92 ± 6.93	2.96 ± 1.77	0.013
OS	3.35 ± 2.72	2.97 ± 1.66	0.279
SRI			
OD	1.03 ± 0.64	0.58 ± 0.37	< 0.001
OS	0.97 ± 0.61	0.65 ± 0.31	< 0.001
SAI			
OD	2.32 ± 1.82	1.15 ± 1.02	< 0.001
OS	2.18 ± 1.61	1.14 ± 0.92	< 0.001

AveK, average keratometry; D, diopter; Cyl, cylinder; SRI, surface regularity index; SAI, surface asymmetric index.
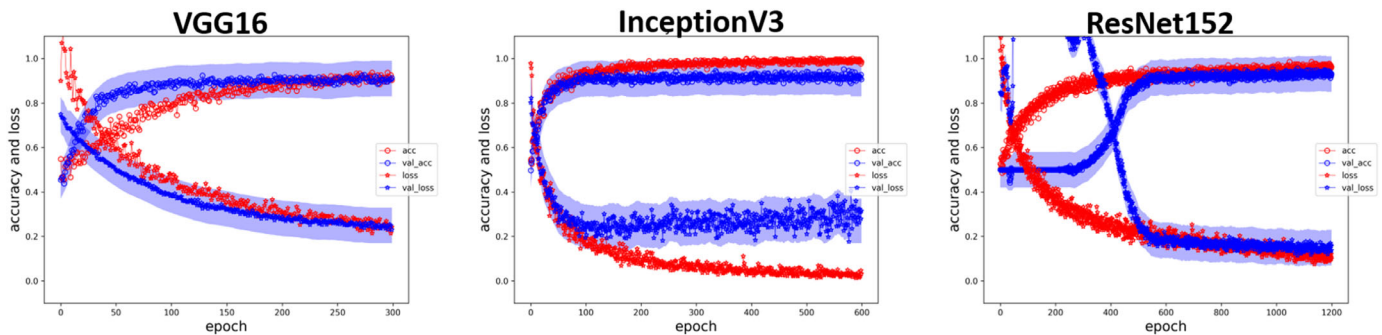

Figure 3. The training results of CNN. The *red dots* indicate the accuracy of the training group and the red asterisks indicate the error rate of the training group. The *blue dots* with shaded band indicate the accuracy of the test group and the *blue asterisks* with shaded band indicate the error rate of the test group. As the training progresses, the accuracy increases and the error rate decreases, thereby indicating no overfitting during the training process.

Table 3. Results of Three CNN Models

Model	Accuracy	Sensitivity	Specificity	AUROC
VGG16	0.931	0.917	0.944	0.956
InceptionV3	0.931	0.917	0.944	0.987
ResNet152	0.958	0.944	0.972	0.995

predicted by the algorithm in Figures 5A and 5B was 90% and 92%, respectively. Figure 5C had a relatively flat topographic feature, and the algorithm thought it was nonkeratoconus with 88% possibility. Figure 5D

was regular astigmatism and predicted as 81% nonkeratoconus feature by the algorithm. Therefore the current algorithm successfully differentiated between nonkeratoconus and keratoconus.

The visualization procedures were applied to the VGG16 model with pixel-wise discriminative features and class-discriminative heat maps for better interpretation (Fig. 6). In Figure 6, the pixel-wise discriminative features outline the gradient difference of topographic maps toward the temporal-lower quadrant in the upper case and superior-inferior quadrant in the
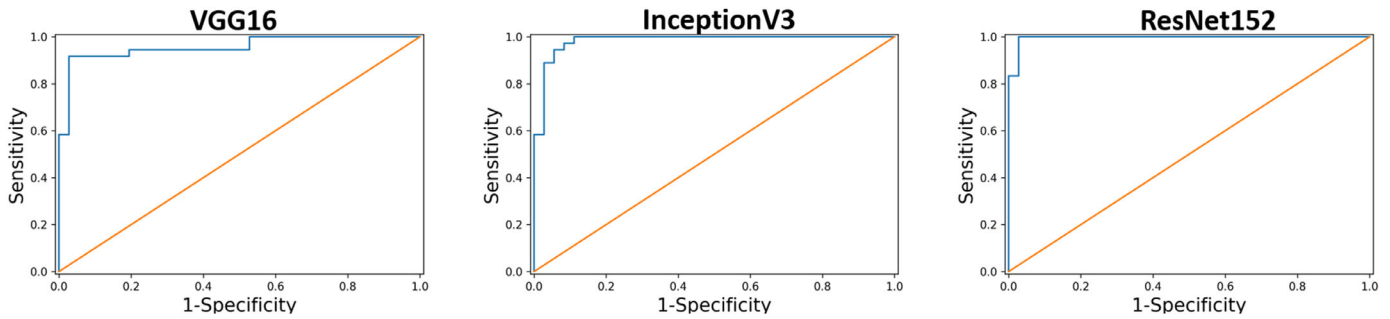
Figure 4. AUROC of the CNN. The AUROC was 0.956 in VGG16 (*left*), 0.987 in InceptionV3 (*middle*), and 0.995 in ResNet152 (*right*).
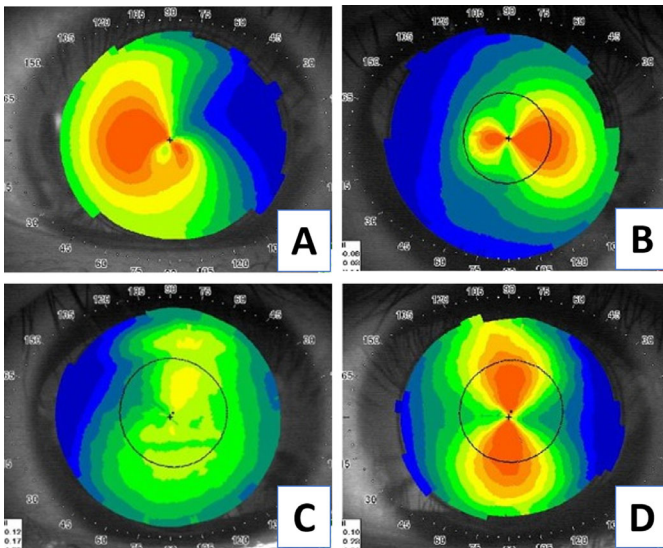

Figure 5. The example of the trained CNN. (**A**) and (**B**) are the keratoconus group, and (**C**) and (**D**) are the control group. The algorithm predicted that (**A**) is keratoconus in 90%, and (**B**) is keratoconus in 92%; the (**C**) is nonkeratoconus in 88%, and (**D**) is nonkeratoconus in 81%.

lower case. Furthermore, the class-discriminative heat map method implied that the CNN model focused on the largest gradient difference of topographic maps. Similar to pixel-wise discriminative features, the class-discriminative heat map revealed that the significance extended from the central to temporal-lower quadrant in the upper case and was in the central region in the lower case, which are diagnostic clues that ophthalmologists use in clinical practice.

The predictability of subclinical keratoconus by CNN model was showed in Table 4. We used VGG16 as a representation and also compared the probability of keratoconus with keratoconus index (KCI),[6,36] keratoconus severity index (KSI),[20] Belin/Ambrósio enhanced ectasia display deviation (BAD-D),[37] and Corvis biomechanical index (CBI).[38] The probability of keratoconus ranged from 0% to 97% and

the accuracy was 28.5% (8/28) if we set the cutoff value as 50%. Overall, the probability of keratoconus was most correlated with KCI (correlation coefficient $r = 0.63$). Taking the accurate cases into consideration, the probability was highly correlated with KSI (correlation coefficient $r = 0.68$). No correlation was observed between the predictive model with BAD-D and CBI. The ground truth of the predictive model was based on the pattern of the topographic images and no topographic parameters were involved. The positive correlation with the topographic indices demonstrated the diagnostic process of the model was reliable.

Discussion

For early diagnosis or predicting the development of keratoconus, several parameters, such as Kmax more than 47.2D,[24] I-S[25] more than 1.4, and KISA more than 60[26] have been used, but the difficulty includes high false-positive rates and the complexity and the number of available parameters that may make the interpretation difficult.[39] Consequently, several machine-learning methods with an anterior segment analyzer or topography have been developed for the same purpose. For example, Smadja et al.[40] developed a method for automatizing the detection of keratoconus based on a tree classification. They imaged 372 eyes with a dual Scheimpflug analyzer and a total of 55 parameters were obtained. After the training and pruning process, they reached a sensitivity of 100% and a specificity of 100% for the discriminative ability between normal and keratoconus. Hidalgo et al.[41] trained a support vector machine algorithm with 22 parameters obtained from Pentacam measurements. The accuracy was 88.8%, with an average sensitivity of 89.0% and specificity of 95.2%. However, the tree discriminant rule and machine algorithm remain
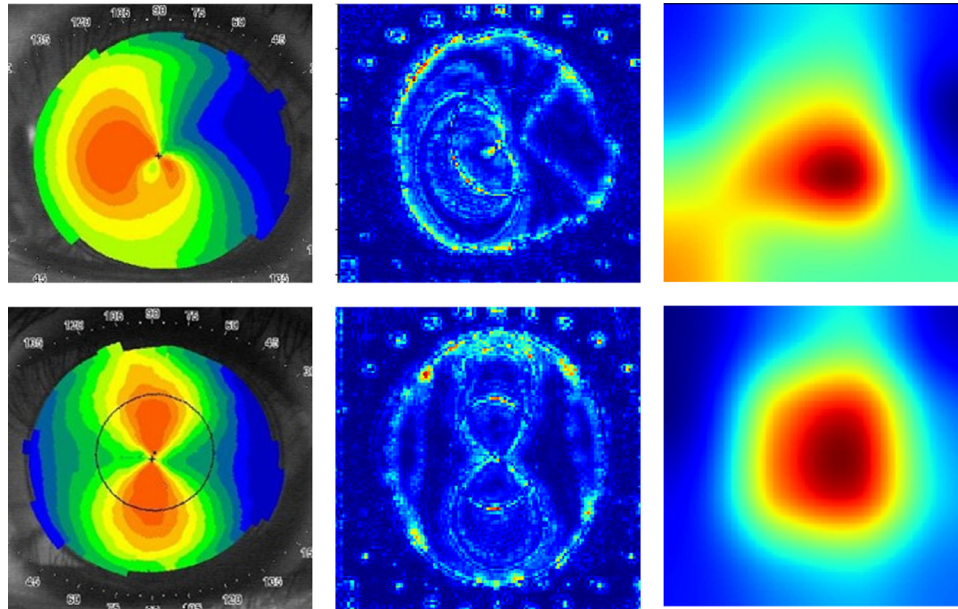
Figure 6. The visualization of the trained CNN. The first column is the original topographic image. The second column is the pixel-wise discriminative features, and it can outline the gradient difference of the topographic maps clearly. The third column is the class-discriminative heat map visualization method, and it reveals that the most significant area in the topographic images lies in the area of the greatest gradient difference.

impractical in terms of quick and simple clinical application and generalization.

In the current study, we trained three CNNs with topographic images without manual segmentation. We did not include topographic parameters in the training process. Our results revealed comparable sensitivity, specificity, and accuracy by training with color-coded maps and CNN models (Table 3). The sensitivity and specificity were over 0.90 in all models, and the specificity of ResNet152 was 0.972. High sensitivity suggested a low false-negative prediction rate, which implied that the present trained CNN models were suitable for keratoconus screening. Moreover, it was applicable in daily clinical situation owing to its high specificity, which means good prediction power for normal controls. Our models could identify not only the typical inferior steep pattern but also other typical keratoconus topographic patterns, such as central cone and asymmetric bowtie with a skewed radial axis (AB/SRAX) (Fig. 5A).

The color scale of each topography map was different in our study. In the keratoconus group, the maximum of color scale ranged from 43.4 to 69.9 while the minimum ranged from 32.6 to 51.5. In the control group, the maximum of color scale was in the range of 42.8 to 51.1 while the minimum was in the range of 37.7 to 45.2. This meant that CNN models were trained with data having varied color scales. Nevertheless, none of color scale information was given in the

training process, as well as in the validation process. As a result, the CNN models could learn more pattern features of topography map for classification rather than the features of fixed color scale. According to the good performance of model validation, it was not the color scale but the pattern of topographic features that mattered the classification process and it has two important application. First of all, it is believed that this method could work across different platforms (e.g., Nidek, Pentacam, Galilei, etc.) as long as it shows the correct topographic pattern even the color scale is different among these platforms. Second, it may be able to train the models with front, back, or topographic maps with reasonable topographic patterns. The front topographic corneal maps were used in this study and it may have comparable results with different topographic maps from different platforms.

Clinical manifested keratoconus was easily diagnosed by our CNN models, while the accuracy of subclinical keratoconus diagnosis was barely satisfactory. There were several explanations. First of all, it was estimated that 50% of subclinical keratoconus progressed to keratoconus within 16 years, and the greatest risk was the first six years of onset.[42] The clinical course of subclinical keratoconus was a long-term progression, and it was reasonable that the prediction was unsatisfactory when based solely on one topography image. Second, the accuracy of a model was dependent on the cutoff value set by

Table 4. Keratoconus Probability Prediction of Subclinical Keratoconus Cases[a] by VGG16 and Comparing With Topographic, Tomographic, and Tonometry Indexes

| Case No. | Probability of keratoconus (%) | TMS Indexes (Topography) | | Pentacam (Tomography) | Corvis (Tonometry) |
		KCI	KSI	BAD-D	CBI
1	97%	15.7%	35%	0.23	0
2	89%	24.2%	37.5%	0.22	0
3	89%	29.3%	29.4%	0.5	0.07
4	81%	0%	0%	2.75	0.67
5	73%	23.8%	31.7%	3.03	0.08
6	67%	34.3%	22.3%	0.19	0
7	64%	0%	0%	0.5	0
8	55%	0%	0%	2.1	0.02
9	44%	41.2%	37.2%	0.36	0
10	37%	0%	19.3%	0.64	0
11	36%	26.5%	31.7%	1.81	0.82
12	20%	0%	0%	0.94	0
13	14%	0%	15.3%	0.49	0.06
14	12%	0%	0%	2.82	1
15	10%	0%	0%	0.92	0
16	7%	0%	0%	1.83	0.03
17	4%	0%	20.8%	1.84	0.05
18	4%	0%	0%	1.27	0.05
19	3%	0%	0%	1.57	0.07
20	3%	0%	0%	0.91	0
21	2%	0%	15.5%	1.61	0.25
22	1%	0%	20%	2.37	0
23	1%	0%	17.9%	2.11	0
24	1%	0%	17.4%	1.62	0.01
25	1%	0%	15.8%	2.38	0
26	0%	0%	0%	2.01	0.9
27	0%	0%	0%	1.79	0.71
28	0%	0%	17.9%	1.87	0.23

[a]The ground truth of subclinical keratoconus was based on topographic pattern interpreted by clinicians.

investigators. It was usually set at the value of 50% in most circumstances but can also be adjusted according to different clinical situations. For the purpose of screening, we may set the value around 30% to 40% and find as many subclinical cases as possible. The accuracy of subclinical keratoconus test group will increase from 28.5% to 39.2% (11/28) if we set the cutoff value at 30%. In our opinion, we may improve the predictability by replacing the training dataset with longitudinal topography maps and tested with several topography images. On the other hand, adjusting the cutoff value may be another effective method to improve the predictability if we don't have longitudinal topographic maps for training.

We adopted a data augmentation and dropout layer in the training process to prevent overfitting. Data were augmented using random shift and rotation within 5% and random horizontal flip during each epoch. Therefore, the more the training times (epoch), the more the diversity of the training dataset. Moreover, not including the test dataset into the training process made the accuracy more reliable. In the last two fully-connected layers, we implanted a dropout rate of 0.5 to prevent overtraining of the algorithm.[43] Figure 2 reveals that the error rate of the test group decreased after training, indicating no overfitting during the algorithm development and making the algorithm reliable.

The nature of the hidden layer in the CNN model prevented us from interpreting the decision base of the algorithm. Therefore we performed the pixel-wise discriminative feature through Deconvnet[34] and developed the class-discriminative heat map of diopter images through Grad-CAM.[35] Deconvnet could highlight the detail in images with high resolution, and Grad-CAM could localize the most significant region in the image. Therefore we needed to determine what and where the CNN model read during the classification process under these two methods. Initially, these two methods were performed to visualize the VGG16 model. In the pixel-wise discriminative feature method, the model focused on the shapes of diopter contours, which is very crucial to evaluate corneal topography. In the heat map of the prediction layer by Grad-CAM, the model localized a hot region corresponding to the high variation or gradient of diopter contours in topographic images. This visualization approach helped us outline the detail and the region of interest with the highest significance in the topographic image reading by the VGG16 model. We noted that it focused on the region similar to the one relied on by ophthalmologists. Furthermore, we tried to visualize the other two CNN models, but results were difficult to interpret because of the nature of the CNN model architecture. Further effort was mandatory for these issues to facilitate a better understanding of the model and to make the model more interpretable. Nonetheless, we believed the visualization of the VGG16 model could provide ophthalmologists insights to re-examine the possible features of keratoconus detection.

Keratoconus is a bilateral, progressive corneal degenerative disease that is garnering significance because of the increase in the number of laser refractive surgeries worldwide, as well as for its incidence in children.[44–46] Preoperative screening of keratoconus may prevent post-LASIK ectasia (PLE). Multiple objective metrics obtained using contemporary corneal imaging devices have been proposed for the diagnosis and severity evaluation of keratoconus. Nonetheless, these highly complex and numerous parameters can pose a clinical challenge for ophthalmologists. Therefore we believe our models could aid in simplifying keratoconus screening with high accuracy in the clinical setting or telemedicine.

Currently, diagnosing keratoconus remains controversial. Scheimpflug-based tomographic indices, the curvature of the posterior cornea, corneal thickness, and biomechanical assessments all help in detecting clinical and subclinical keratoconus.[47–50] Furthermore, the integration of several diagnostic approaches has exhibited enhanced detection ability.[51] In addition to these modalities, the Placido disc-based topography can evaluate the ocular surface and tear film, which are crucial in clinical evaluation. To the best of our knowledge, this study is the first to develop an AI method to diagnose the corneal disease based on corneal topography. We believe this method can be of great benefit in the clinical scenario, alone or combined with other devices.

The limitations of this study are its retrospective nature and the limited number of patients for algorithm training. Moreover, the final diagnosis of topographic images was decided by one institute. Therefore future studies can achieve further progress in this aspect by including more cases and disease types. In addition, future studies can involve different specialists to help arrive at a consensus regarding the diagnosis.

In conclusion, the current study revealed that the CNN model had high sensitivity and specificity in identifying patients with keratoconus through topographic images. In addition, it exhibited a reasonable diagnostic ability after applying visualization methods. Nevertheless, further study is mandatory to explore the applicability of this model in daily clinical practice and large-scale screening.

Acknowledgments

References

1. Krachmer JH, Feder RS, Belin MW. Keratoconus and related noninflammatory corneal thinning disorders. *Surv Ophthalmol*. 1984;28:293–322.
2. Chatzis N, Hafezi F. Progression of keratoconus and efficacy of pediatric [corrected] corneal collagen cross-linking in children and adolescents. *J Refract Surg*. 2012;28:753–758.
3. Jacobs DS, Dohlman CH. Is keratoconus genetic? *Int Ophthalmol Clin*. 1993;33:249–260.
4. Sherwin T, Brookes NH. Morphological changes in keratoconus: pathology or pathogenesis. *Clin Exp Ophthalmol*. 2004;32:211–217.
5. Ferdi AC, Nguyen V, Gore DM, Allan BD, Rozema JJ, Watson SL. Keratoconus natural progression: a systematic review and meta-analysis of 11 529 eyes. *Ophthalmology*. 2019;126:935–945.

6. Maeda N, Klyce SD, Smolek MK, Thompson HW. Automated keratoconus screening with corneal topography analysis. *Invest Ophthalmol Vis Sci*. 1994;35:2749–2757.

7. Hustead JD. Detection of keratoconus before keratorefractive surgery. *Ophthalmology*. 1993;100:975.

8. Mas Tur V, MacGregor C, Jayaswal R, O'Brart D, Maycock N. A review of keratoconus: diagnosis, pathophysiology, and genetics. *Surv Ophthalmol*. 2017;62:770–783.

9. Rabinowitz YS, McDonnell PJ. Computer-assisted corneal topography in keratoconus. *Refract Corneal Surg*. 1989;5:400–408.

10. Dastjerdi MH, Hashemi H. A quantitative corneal topography index for detection of keratoconus. *J Refract Surg*. 1998;14:427–436.

11. Holladay JT. Keratoconus detection using corneal topography. *J Refract Surg*. 2009;25:S958–962.

12. LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature*. 2015;521:436–444.

13. Tompson J, Jain A, LeCun Y. Joint training of a convolutional network and a graphical model for human pose estimation. *Adv Neural Information Processing Syst*. 2014;27:1799–1807.

14. Hinton G, Deng L, Yu D. Deep neural networks for acoustic modeling in speech recognition. *IEEE Signal Processing Magazine*. 2012;29:82–97.

15. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103:167–175.

16. Cheung CY, Tang F, Ting DSW, Tan GSW, Wong TY. Artificial intelligence in diabetic eye disease screening. *Asia Pac J Ophthalmol (Phila)*. doi:10.22608/APO.201976. Published online ahead of print.

17. Zheng C, Johnson TV, Garg A, Boland MV. Artificial intelligence in glaucoma. *Curr Opin Ophthalmol*. 2019;30:97–103.

18. Schmidt-Erfurth U, Waldstein SM, Klimscha S, et al. Prediction of Individual Disease Conversion in Early AMD Using Artificial Intelligence. *Invest Ophthalmol Vis Sci*. 2018;59:3199–3208.

19. Reid JE, Eaton E. Artificial intelligence for pediatric ophthalmology. *Curr Opin Ophthalmol*. 2019;30:337–346.

20. Smolek MK, Klyce SD. Current keratoconus detection methods compared with a neural network approach. *Invest Ophthalmol Vis Sci*. 1997;38:2290–2299.

21. Accardo PA, Pensiero S. Neural network-based system for early keratoconus detection from corneal topography. *J Biomed Inform*. 2002;35:151–159.

22. Hwang ES, Perez-Straziota CE, Kim SW, Santhiago MR, Randleman JB. Distinguishing highly asymmetric keratoconus eyes using combined scheimpflug and spectral-domain OCT analysis. *Ophthalmology*. 2018;125:1862–1871.

23. Lopes BT, Ramos IC, Salomao MQ, et al. Enhanced tomographic assessment to detect corneal ectasia based on artificial intelligence. *Am J Ophthalmol*. 2018;195:223–232.

24. Ramos IC, Correa R, Guerra FP, et al. Variability of subjective classifications of corneal topography maps from LASIK candidates. *J Refract Surg*. 2013;29:770–775.

25. Rabinowitz YS, McDonnell PJ. Computer-assisted corneal topography in keratoconus. *Refract Corneal Surg*. 1989;5:400–408.

26. Rabinowitz YS, Rasheed K. KISA% index: a quantitative videokeratography algorithm embodying minimal topographic criteria for diagnosing keratoconus. *J Cataract Refract Surg*. 1999;25:1327–1335.

27. Rabinowitz YS, Yang H, Brickman Y, et al. Videokeratography database of normal human corneas. *Br J Ophthalmol*. 1996;80:610–616.

28. Li X, Yang H, Rabinowitz YS. Keratoconus: classification scheme based on videokeratography and clinical signs. *J Cataract Refract Surg*. 2009;35:1597–1603.

29. Henriquez MA, Hadid M, Izquierdo L, Jr. A systematic review of subclinical keratoconus and forme fruste keratoconus. *J Refract Surg*. 2020;36:270–279.

30. Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint:1409.1556.

31. Szegedy C, Wei L, Yangqing J, et al. Going deeper with convolutions. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2015:1–9.

32. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 2016;770–778.

33. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*. 2012;1097–1105.

34. Zeiler MD, Fergus R. Visualizing and Understanding Convolutional Networks. In: Fleet DP, Pajdla T, Schiele B, Tuytelaars T, eds. *The European Conference on Computer Vision (ECCV)*. Zurich, Switzerland: Springer; 2014:818–833.

35. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-based

Localization. *IEEE International Conference on Computer Vision (ICCV)*: *The Institute of Electrical and Electronics Engineers (IEEE)*; 2017;618–626.

36. Maeda N, Klyce SD, Smolek MK. Comparison of methods for detecting keratoconus using videokeratography. *Arch Ophthalmol*. 1995;113:870–874.

37. Luz A, Lopes B, Hallahan KM, et al. Enhanced Combined Tomography and Biomechanics Data for Distinguishing Forme Fruste Keratoconus. *J Refract Surg*. 2016;32:479–494.

38. Vinciguerra R, Ambrosio R, Jr, Elsheikh A, et al. Detection of Keratoconus With a New Biomechanical Index. *J Refract Surg*. 2016;32:803–810.

39. Maeda N, Klyce SD, Smolek MK. Neural network classification of corneal topography. Preliminary demonstration. *Invest Ophthalmol Vis Sci*. 1995;36:1327–1335.

40. Smadja D, Touboul D, Cohen A, et al. Detection of subclinical keratoconus using an automated decision tree classification. *Am J Ophthalmol*. 2013;156:237–246 e231.

41. Ruiz Hidalgo I, Rodriguez P, Rozema JJ, et al. Evaluation of a Machine-Learning Classifier for Keratoconus Detection Based on Scheimpflug Tomography. *Cornea*. 2016;35:827–832.

42. Li X, Rabinowitz YS, Rasheed K, Yang H. Longitudinal study of the normal eyes in unilateral keratoconus patients. *Ophthalmology*. 2004;111:440–446.

43. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:12070580*. 2012.

44. Randleman JB, Dupps WJ, Jr, Santhiago MR, et al. Screening for Keratoconus and Related Ectatic Corneal Disorders. *Cornea*. 2015;34:e20–22.

45. Moshirfar M, Heiland MB, Rosen DB, Ronquillo YC, Hoopes PC. Keratoconus Screening in Elementary School Children. *Ophthalmol Ther*. 2019;8:367–371.

46. Saad A, Gatinel D. Combining placido and corneal wavefront data for the detection of forme fruste keratoconus. *J Refract Surg*. 2016;32:510–516.

47. de Sanctis U, Loiacono C, Richiardi L, Turco D, Mutani B, Grignolo FM. Sensitivity and specificity of posterior corneal elevation measured by Pentacam in discriminating keratoconus/subclinical keratoconus. *Ophthalmology*. 2008;115:1534–1539.

48. Ucakhan OO, Cetinkor V, Ozkan M, Kanpolat A. Evaluation of Scheimpflug imaging parameters in subclinical keratoconus, keratoconus, and normal eyes. *J Cataract Refract Surg*. 2011;37:1116–1124.

49. Ambrosio R, Jr, Dawson DG, Salomao M, Guerra FP, Caiado AL, Belin MW. Corneal ectasia after LASIK despite low preoperative risk: tomographic and biomechanical findings in the unoperated, stable, fellow eye. *J Refract Surg*. 2010;26:906–911.

50. Roberts CJ, Dupps WJ, Jr. Biomechanics of corneal ectasia and biomechanical treatments. *J Cataract Refract Surg*. 2014;40:991–998.

51. Ambrosio R, Jr, Lopes BT, Faria-Correia F, et al. Integration of Scheimpflug-Based Corneal Tomography and Biomechanical Assessments for Enhancing Ectasia Detection. *J Refract Surg*. 2017;33:434–443.