A Deep Learning–Based Framework for Accurate Evaluation of Corneal Treatment Zone After Orthokeratology

Yong Tang1,*, Zhao Chen2,*, Weijia Wang3, Longbo Wen2, Linjing Zhou3, Mao Wang4, Fan Tang4, He Tang5, Weizhong Lan2,6, and Zhikuan Yang2,7

1 School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
2 Aier School of Ophthalmology, Central South University, Changsha, China
3 School of Information and Software Engineering, University of Electronic Science and Technology of China, Chengdu, China
4 Information Center, Aier Eye Hospital Group, Changsha, China
5 School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu, China
6 Guangzhou Aier Eye Hospital, Jinan University, Guangzhou, China
7 Hunan Province Optometry Engineering Technology Research Center, Changsha, China

Correspondence: He Tang, School of Electronic Science and Engineering, University of Electronic Science and Technology of China, Chengdu 610054, China. e-mail: tanghe@uestc.edu.cn

Weizhong Lan, Aier School of Ophthalmology, Central South University, Changsha 410000, China. e-mail: lanweizhong@aierchina.com

Received: March 9, 2021
Accepted: June 27, 2021
Published: December 21, 2021

Keywords: myopia; deep learning; orthokeratology; corneal topography

Citation: Tang Y, Chen Z, Wang W, Wen L, Zhou L, Wang M, Tang F, Tang H, Lan W, Yang Z. A deep learning–based framework for accurate evaluation of corneal treatment zone after orthokeratology. Transl Vis Sci Technol. 2021;10(14):21, https://doi.org/10.1167/tvst.10.14.21

Purpose: Given the robust effectiveness of inhibiting myopia progression, orthokeratology has gained increasing popularity worldwide. However, identifying the boundary and the center of reshaped corneal area (i.e., treatment zone) is the main challenging task in evaluating the performance of orthokeratology. Here we present automated deep learning algorithms to solve the challenges.

Methods: A total of 6328 corneal topographical maps, including 2996 axial subtractive maps and 3332 tangential subtractive maps, were collected from 2044 myopic patients who received orthokeratology. The boundary and the center of the treatment zones were annotated by experts as ground truths using axial subtractive maps and tangential subtractive maps, respectively. The algorithms based on neural network structures of fully convolutional networks (FCNs) and convolutional neural networks (CNNs) were developed to automatically identify the boundary and the center of the treatment zone, respectively.

Results: The algorithm of FCNs identified the treatment zone boundaries with an accuracy intersection over union (IoU) of 0.90 ± 0.06 (mean ± SD; range, 0.60–0.97). The algorithm of CNNs also identified the treatment zone centers with an average deviation of 0.22 ± 0.22 mm (range, 0.01–1.66 mm).

Conclusions: These results show that a deep learning–based solution is able to provide an automatic and accurate tool to accomplish the two main challenges of orthokeratology.

Translational Relevance: Deep learning in orthokeratology can shorten the time while maintaining accurate results in clinical practice, which enables clinicians to help more patients daily.

Introduction

The orthokeratology (OK) lens is a rigid reverse-geometry contact lens that is designed to temporarily correct myopia by reshaping the cornea.1 Accumulating evidence has shown the efficacy of an OK lens on inhibiting myopia progression2 and that the lens has gained increasing popularity worldwide.3 A standard OK lens has four curves, from the center to the periphery, namely, a base curve, a reverse curve, an alignment curve, and a peripheral curve. Of these curves, the radius of the base curve is designed to be larger than that of the central cornea to achieve the treatment by inducing a flattened cornea, designated as the treatment zone. The location and size of the treatment...
zone are two critical parameters in evaluating the performance of the OK lens treatment. In practice, this is determined by examining the subtractive map between the topography measured before and after OK lens treatment using corneal topographers. Significant dislocation of the treatment zone (i.e., deviation of the center of the treatment zone from the pupil center) or a too small treatment zone leads to poor visual acuity, visual disturbance such as glare and double vision, and poor antimyopia effect. An imperfect treatment zone usually leads to multiple adjustments of the lens’s parameters or even ordering a new lens.

Despite the importance, there is no standard technique to assess the center and the boundary of the treatment zone. In general clinical practice, clinicians have to make decisions based on their experience. This decision-making procedure is very subjective and inevitably introduces variations among clinicians. In addition, the decision-making procedure in some ambiguous cases is particularly challenging for less experienced clinicians. In recent years, several scholars have attempted to develop semiobjective techniques to facilitate the decision-making procedure, based on a third-party commercial software. However, these techniques were only applicable for cases when the topography formed a complete, continuous, “bull’s-eye” pattern (i.e., an all-round perfect contact for the reverse curve, the cornea, and the usually well-centered location), which are not always the perfect cases. Another significant drawback of these techniques is that they are time-consuming and require clinicians to outline the contour of the treatment zone manually, followed by additional analysis using third-party software.

Thanks to the remarkable developments of deep learning algorithms, recent years have seen significant advances in artificial intelligence (AI). As a new branch of AI, deep learning has sparked tremendous research interest and led to applications in many fields, including health care and medical diagnosis. The adoption of deep learning in medical image analysis, for instance, has demonstrated competitive performance in a range of analysis tasks, such as classification, detection, segmentation, registration, and location. Deep learning has also shown exciting potential in ophthalmology based on optical coherence tomography and fundus photographs. In this study, we proposed and developed a deep learning approach to evaluate the treatment effect after orthokeratology. Specifically, we developed and trained a deep learning neural network with an encoder–decoder structure to identify the boundary of the treatment zone and developed a deep learning neural network with a convolutional neural network structure to determine the center of the treatment zone. Both algorithms achieved promising performance in validation data sets. With this approach, the critical evaluation process can be conducted fully automatically.

### Methods

#### Image Data Set

The overall flowchart of this study is shown in Figure 1. First, 6328 anonymized corneal topography maps were collected from Changsha Aier Eye Hospital, Aier Eye Hospital Group. The image data set comprised axial subtractive maps \( n = 2996 \) and tangential subtractive maps \( n = 3332 \) from patients who received orthokeratology treatment between 2015 and 2018. The corneal topography of these patients was measured using the Pentacam HR tomographer (Oculus GmbH, Wetzlar, Germany), which is a noninvasive anterior segment tomographer based on rotating Scheimpflug technology. For each eye, corneal topographies at two different stages were taken. Specifically, the first was taken before the OK lens was worn \( t = 0 \) days and the second was taken approximately 3 months \( t = 90 \) days after the OK lens treatment, when the corneal parameters were stabilized. Then, the subtractive map was generated by subtracting the second map from the first map. Finally, the subtractive map was extracted as an image file for later annotation. Subtractive maps can show the practitioner about the effect of orthokeratology on the corneal surface. Two types of subtractive maps were employed for analysis. Axial subtractive maps reflected the changes of optical power in axial corneal surfaces and were used to determine the area of the treatment zone in the study. Tangential subtractive maps offered a better representation of the changes in cornea and were used to determine the centration of the treatment zone. All the maps were extracted as images with a size of 300 \( \times \) 300 pixels, corresponding to 10.63 \( \times \) 10.63 mm; the diameter of the corneal topography was 254 pixels, or 9 mm. The study was approved by the institute’s ethical committee (AIER2018IRB27). Given the nature of the data and the study design, participants’ informed consent was not required.

#### Image Annotation

To provide the ground truths for training and validating the deep learning algorithms, we invited clinicians to manually annotate all images. First, the image files of axial subtractive maps were manually segmented by three experienced clinicians independently to indicate the boundary of the treatment zone. Then the results were reviewed by another expert, who
is one of the authors, to ensure data quality. For a given axial subtractive map, one segmentation result was randomly chosen from the qualified results to serve as the ground truth. Similarly, three experienced clinicians were invited to identify the center of the treatment zone in the images of tangential subtractive maps independently. The ground truth of the center was obtained by averaging the positions of the three annotations to eliminate bias. The annotations were conducted using a tool developed in house.

**Deep Learning Methods**

The automatic analysis of the OK treatment maps consisted of two tasks, namely, the identification of the treatment zone in axial subtractive maps and the
Identification of the centers in tangential subtractive maps. Since these two tasks were independent, we developed two different deep learning algorithms to achieve the objectives of these two tasks. Furthermore, the two algorithms were trained using two separate data sets of axial subtractive maps and tangential subtractive maps to identify the boundaries and the centers of the treatment zones, respectively.

Identification of the Treatment Zone Boundary

The first task was to automatically identify the boundary of the treatment zone in an axial subtractive map. A semantic segmentation approach was used for the pixel-level classification, in which all pixels of a given axial subtractive image were labeled as either one of the two parts, namely, the treatment zone and the remaining nonflattened region. Deep learning models have achieved state-of-the-art performance in medical image semantic segmentation. In this study, we adopted a deep neural network based on the fully convolutional network (FCN) structure. It comprised two parts, including one encoder and one decoder. The encoder performed the downsampling to convert the original image to a smaller size.

Identification of the Treatment Zone Center

In the second task, identifying the treatment zone center, the output of an input image is a tuple comprising the two coordinates of the treatment zone center. Therefore, we needed a neural network structure that was capable of extracting features by way of dimensionality reduction. In this study, we designed a four-layer CNN to calculate the coordinates of the
treatment zone center. In each convolutional layer, a convolutional operation was performed using a kernel in a sliding window approach to scan the input image. Afterward, a rectified linear unit activation layer was conducted to introduce nonlinearity before a max-pooling operation was performed to shrink the dimensionality. The four convolutional layers were performed sequentially before the final two coordinates were output to indicate the treatment zone center. Figure 2B presents the detailed structure of the CNN model.

We developed the deep learning models in Python (version 3.7.3) using the open-source deep learning library PyTorch (version 1.1.0). Additional Python libraries used in this study were NumPy (version 1.16.2), Pandas (version 0.24.2), and Matplotlib (version 3.0.3). The models were trained and evaluated using the Nvidia-SMI 418.67 Tesla P40 (Nvidia, Santa Clara, CA, USA) graphical processing unit with the compute unified device architecture toolkit (version 10.1.105).

**Algorithm Development**

The boundary and the center of the treatment zone were pixelwise annotated by experts before they were used to develop the algorithm. Both the data sets of the axial subtractive maps and the tangential subtractive maps were randomly split into a training set (90%) and a validation set (10%) for deep learning model training and evaluation, respectively. First, deep learning neural networks of the FCN structure were trained to identify the boundary of the treatment zone based on the axial subtractive maps. At the same time, the CNN structures were trained to identify the treatment zone centers based on tangential subtractive maps.

**Evaluation of Treatment Zone Boundary Identification**

To evaluate the performance of the deep learning models in identifying the treatment zone boundary, we adopted the evaluation metric of intersection over union (IoU = TP/(TP + FP + FN)), where TP is true positive, FP is false positive, FN is false negative. The IoU was calculated using the predicted area obtained by the proposed model and the ground truth annotated by the human expert for each validation sample. A larger value of IoU closer to 1 indicated better performance compared to poor performance with a smaller IoU. Therefore, the main task in developing the segmentation algorithm was to maximize the IoU value. We reported the average IoU as the performance of a given deep learning model over the samples in the whole validation set. Besides IoU, in line with segmentation literature, we also reported the value of the Dice similarity coefficient (DSC = 2TP/(TP + FP + TP + FN)) as the secondary evaluation metric to describe the overlapping of the ground truth and the prediction. To visualize the result of treatment zone identification, we drew the boundaries of the ground truth area and the predicted area in the same axial subtractive map.

**Evaluation of Treatment Zone Center Identification**

The prediction of the treatment zone centers was to find the two values of coordinates in the tangential subtractive map. Given the two positions of the ground truth annotated by the experts and the prediction obtained by the proposed CNN model, adopting the Euclidean distance between the two centers to indicate the deviation was straightforward. The deviation was described as pixels in the image and later converted into the physical distance in millimeters. The optimization target of the model was to minimize the gap between the predicted center and the ground truth center. As described above, the ground truth center was obtained by averaging the three centers annotated by the three experts. To visualize the result of treatment zone center identification, the two centers were plotted in the same tangential subtractive map.

**Results**

**Data Description**

A total of 6328 maps of 2044 patients treated with the OK lens were included in this study. The average age at the initiation of OK lens wear was 13.56 ± 2.78 years, ranging from 8 to 23 years. Among them, 909 were male (44.47%) and 1135 were female (55.53%). As a result, 2996 axial subtractive maps and 3332 tangential subtractive maps were obtained and anonymized to develop the deep learning algorithms.

**Performance of Treatment Zone Boundary Identification**

For the identification of the treatment zone boundary in axial subtractive maps, the interclinician IoU was 0.54 ± 0.11, 0.50 ± 0.11, and 0.83 ± 0.09, respectively (all P < 0.01). Despite the significantly good agreement between these three experienced clinicians, only one segmentation result was randomly chosen to serve as the ground truth for each axial subtractive map,
Figure 3. Histogram of the performance in validation data sets. (A) Histogram of IoU for treatment zone identification in the validation set of axial subtractive maps. Larger IoU values indicate good segmentation of treatment zones. (B) Histogram of deviation (millimeters) for treatment zone center identification in the validation set of tangential subtractive maps. Small values of deviation indicate good identification of treatment zone centers.

Figure 4. Examples of treatment zone identification in the axial subtractive maps with the best (A) and the worst performance (B). The white contour was annotated by AI, while the black contour was annotated by the human expert. (A) Good segmentation of the treatment zone with the greatest IoU of 0.96. (B) Poor segmentation of the treatment zone with the smallest IoU of 0.66, which was due to the relatively great decentration of the OK lens and the resultant unclear treatment boundary. Cases with this type of corneal change usually lead to borderline unacceptable visual quality and require adjustment of the lens parameter to achieve better lens centration.

to reduce possible bias of the readers. Based on the data set of the ground truth, we first trained the FCN model on the training set, then evaluated the model on another independent validation set. The performance in the training stage was a mean ± SD IoU of 0.92 ± 0.04 and DSC of 0.98 ± 0.02. The final results of the segmentation model in the validation set were very promising, with an IoU of 0.90 ± 0.06 and a DSC of 0.94 ± 0.04. The distribution of IoU values is illustrated in Figure 3A. The results indicated that the FCN deep learning structure can accurately identify the treatment zone boundary in the axial subtractive map with encouraging performance. Figure 4 shows examples of the best and worst performance in the segmentation of the treatment zone.

Performance of Treatment Zone Center Identification

Similarly, we trained and evaluated the CNN deep learning model to identify the treatment zone center
using the training and validation sets of the tangential subtractive maps, respectively. In the training stage, the CNN model achieved an average deviation of 5.33 ± 4.42 pixels between the ground truth center and the predicted center, namely, 0.19 ± 0.16 mm. The performance on the validation set was 6.32 ± 6.23 pixels or 0.22 ± 0.22 mm. The distribution of deviation in the validation set is plotted in Figure 3B. The results demonstrate that the deep learning model can identify the treatment zone center with high accuracy. Figure 5 shows examples of the best and worst performance in the identification of the treatment zone center.

**Discussion**

Deep learning is a new branch of artificial neural network characterized by multiple layers, a large number of neurons, and complicated network structures.\(^\text{12}\) By training the connectivity matrices, a deep learning network can effectively learn the abstract representations of the training samples in a variety of inference tasks with unprecedented performance. To our knowledge, this is the first time deep learning has been applied in facilitating OK lens treatment. It was found that a deep learning–based approach can identify the boundary and the center of the treatment zone automatically, with precision comparable to that of human experts and instant efficiency.

Hiraoka et al.\(^\text{9}\) were among the first scholars to attempt to define the treatment zone after OK lens treatment with computer assistance. They manually outlined the contour of the treatment zone with 16 discrete dots based on the annotator’s personal experience and then approximated an ellipse and its center using a customized data analysis program written in a third-party programming language (MATLAB; The MathWorks, Natick, MA, USA). The authors reported the technique had good reproducibility (the repeatability between examiners had not been reported in the literature). A later study by our team in which this technique was adopted showed a much poorer level of intraindividual repeatability,\(^\text{30}\) indicating the performance of the approach relies heavily on the annotator’s experience. More recently, Mei et al.\(^\text{11}\) introduced another technique based on image analysis software (Image-Pro Plus; Media Cybernetics Corporation, Rockville, MD, USA). Similarly, the boundary of the treatment zone was manually depicted by the annotator. The technique was reported to demonstrate “excellent” reproducibility by one examiner and repeatability between two examiners. However, the authors admitted several inherited drawbacks of the technique to restrict its application in practice. For instance, the treatment zone could not be drawn completely when its boundary was discontinuous due to lens decentration or

---

**Figure 5.** Examples of treatment zone center identification in the tangential subtractive maps with the best (A) and the worst performance (B). The white dot was annotated by AI. The black dots were annotated by the three human experts, and then the averaged location was represented by a red dot. (A) Good identification of the treatment center with the smallest deviation of 0.01 mm. (B) Poor identification of the treatment zone center with the greatest deviation of 1.66 mm, which was due to a rare case called “central island” (i.e., an abnormally raised ridge on the cornea after a failed OK lens treatment); therefore, the judgment by AI was misleading.
incomplete acquisition of image data by the topographer. In addition, they pointed out that the manual depiction required highly detailed operation skills with the software; therefore, they did not suggest the method for extensive analysis.

Unlike the aforementioned studies, our approaches are based on advanced deep learning algorithms that can learn the abstract experience of professional experts. Therefore, the trained algorithms can mimic the decision making of experts in evaluating the treatment zone. Our results show that the algorithms developed performed very well in the identification of both the boundary and the center of the treatment zone. The level of accuracy fully met the need for clinical practice. Compared with the previous techniques, the current approach has several advantages. First, the annotation was conducted by three independent experts, but not one or two experts, to avoid individual bias. Second, the large data set covered all possible patterns of the treatment zones, including irregular, discontinuous boundaries (e.g., Fig. 4B) or even some rare cases such as “central island” (e.g., Fig. 5B). This diversity of the data set ensures the generalization as well as the robustness of our algorithms in real practice. Additionally, the developed algorithm can assist practitioners in completing the time-consuming evaluation task almost instantly, which can significantly improve clinical efficiency. Furthermore, our deep learning–based algorithms can also be continually improved by learning from growing data accumulated in practices, leading to better accuracy. The most exceptional advantage is that the fully automated procedure provides a standardized approach, which could minimize the judgment variation between follow-up visits and between clinicians.

However, the study also has several limitations. First, the deep learning algorithms require all input images to be in the same format, and the users need to export the images with an identical standard before applying the algorithms. However, this can be solved in future algorithm development by training different images to be compatible with all formats. Second, the images were all captured using the Pentacam HR tomographer. It is unknown to what extent the performance of the algorithms developed would be in other topographers. However, according to the nature of the topographical images, the performance should be at a comparable level. Although different formats of the topographical images are used by other instruments, the procedures and algorithms are transferrable to shorten the development process. Meanwhile, with the availability of images taken by other tomographers, it is possible to simply train current models to be able to analyze those new images. Third, although average satisfactory performance has been achieved by the approach, some errors might happen in some rare cases (e.g., Figs. 4B and 5B). However, this could be easily noticed by clinicians in practice. With the assistance of AI annotations on the images, clinicians can easily evaluate the magnitude of the treatment zone dislocation of the lens or the overlapping proportion.

In conclusion, this study shows the capabilities of deep learning models to assess the treatment zones and centers after orthokeratology with performance competitive with that of human experts, which can be employed as an automated solution to facilitate the assessment and reduce interindividual subjectivity during the follow-up visits of orthokeratology. However, similar to other algorithm-based approaches, the performance of the system needs further evaluation using other data sets for external validation. Therefore, this fully automated system has been published online for public test and application (see Supplemental Appendix and Video). Meanwhile, the system can be also employed as a teaching platform to train practitioners’ evaluation on the lens fitting in the future.

Acknowledgments

The annotations of the topographical maps were supported in part by Chao Zhou, Jianhua Li, and Jiwen Yang.

Supported by grants from the Hunan Provincial Science and Technology Plan Project (2019SK2051), the Science Fund for Distinguished Young Scientists (2019JJ20034) from Hunan Provincial Science and Technology Department, the Innovation Methodology Project (2017IM010700) of the Ministry of Science and Technology, the Fundamental Research Funds for the Central Universities of Central South University (2019ZZTS367), the Science Research Foundation of Aier Eye Hospital Group (AR1903D7), and Hunan Province International Science and Technology Cooperation Base (2020CB1002).

Disclosure: Y. Tang, None; Z. Chen, None; W. Wang, None; L. Wen, None; L. Zhou, None; M. Wang, None; F. Tang, None; H. Tang, None; W. Lan, None; Z. Yang, None

* YT and ZC contributed equally to the work and therefore should be treated as co-first authors.

References

1. VanderVeen DK, Kraker RT, Pineles SL, et al. Use of orthokeratology for the prevention of myopic
progression in children: a report by the American Academy of Ophthalmology. *Ophthalmology*. 2019;126:623–636.

2. Lee YC, Wang JH, Chiu CJ. Effect of orthokeratology on myopia progression: twelve-year results of a retrospective cohort study. *BMC Ophthalmol*. 2017;17:243.

3. Holden BA, Fricke TR, Wilson DA, et al. Global prevalence of myopia and high myopia and temporal trends from 2000 through 2050. *Ophthalmology*. 2016;123:1036–1042.

4. Hiraoka T, Okamoto C, Ishii Y, Kakita T, Oshika T. Contrast sensitivity function and ocular higher-order aberrations following overnight orthokeratology. *Invest Ophthalmol Vis Sci*. 2007;48:550–556.

5. Santodomingo-Rubido J, Villa-Collar C, Gilmartin B, Gutierrez-Ortega R. Factors preventing myopia progression with orthokeratology correction. *Optom Vis Sci*. 2013;90:1225–1236.

6. Chen Z, Niu L, Xue F, et al. Impact of pupil diameter on axial growth in orthokeratology. *Optom Vis Sci*. 2012;89:1636–1640.

7. Hiraoka T, Okamoto C, Ishii Y, Okamoto F, Oshika T. Recovery of corneal irregular astigmatism, ocular higher-order aberrations, and contrast sensitivity after discontinuation of overnight orthokeratology. *Br J Ophthalmol*. 2009;93:203–208.

8. Carracedo G, Espinosa-Vidal TM, Martinez-Alberquilla I, Batres L. The topographical effect of optical zone diameter in orthokeratology contact lenses in high myopes. *J Ophthalmol*. 2019;2019:1082472.

9. Hiraoka T, Mihashi T, Okamoto C, Okamoto F, Hirohara Y, Oshika T. Influence of induced decentralized orthokeratology lens on ocular higher-order wavefront aberrations and contrast sensitivity function. *J Cataract Refract Surg*. 2009;35:1918–1926.

10. Chen Z, Xue F, Zhou J, et al. Prediction of orthokeratology lens decentration with corneal elevation. *Optom Vis Sci*. 2017;94:903–907.

11. Mei Y, Tang Z, Li Z, Yang X. Repeatability and reproducibility of quantitative corneal shape analysis after orthokeratology treatment using Image-Pro Plus software. *J Ophthalmol*. 2016;2016:1732476.

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

13. Stead WW. Clinical implications and challenges of artificial intelligence and deep learning. *JAMA*. 2018;320:1107–1108.
27. Jin X, Li X, Xiao H, et al. Video Scene Parsing with Predictive Feature Learning. In: *IEEE International Conference on Computer Vision, ICCV 2017, Venice, Italy, October 22-29, 2017*. IEEE Computer Society; 2017:5581–5589.

28. Klein A, Warszawski J, Hillengaß J, Maier-Hein KH. Automatic bone segmentation in whole-body CT images. *Int J Comput Assist Radiol Surg*. 2019;14:21–29.

29. Bokhovkin A, Burnaev E. Boundary loss for remote sensing imagery semantic segmentation. In: Lu H, Tang H, Wang Z, eds. *Advances in Neural Networks—ISNN 2019*. Cham, Switzerland: Springer International; 2019:388–401.

30. Li X, Wang L, Chen Z, Yang Z. Influence of treatment zone decentration on corneal higher-order wavefront aberrations and axial length elongation after orthokeratology. *Chin J Optom Ophthalmol Vis Sci*. 2017;2019(9):540–547.