Multiple-Image Deep Learning Analysis for Neuropathy Detection in Corneal Nerve Images

Fabio Scarpa, PhD, Alessia Colonna, MSc, and Alfredo Ruggeri, MSc

Purpose: Automated classification of corneal confocal images from healthy subjects and diabetic subjects with neuropathy.

Methods: Over the years, in vivo confocal microscopy has established itself as a rapid and noninvasive method for clinical assessment of the cornea. In particular, images of the subbasal nerve plexus are useful to detect pathological conditions. Currently, clinical information is derived through a manual or semiautomated process that traces corneal nerves and achieves their descriptors (eg, density and tortuosity). This is tedious and subjective. To overcome this limitation, a method based on a convolutional neural network (CNN) for the classification of images from healthy subjects and diabetic subjects with neuropathy is proposed. The CNN simultaneously analyzes 3 nonoverlapping images, from the central region of the cornea. The algorithm automatically extracts features, without the need for neither nerve tracing nor parameter extraction nor montage/mosaicking, and provides an overall classification for each image trio.

Results: On a dataset composed by images from 50 healthy subjects and 50 subjects with neuropathy, the algorithm achieves a classification accuracy of 96%. The proposed method improves the results obtained using a traditional method that traces nerves and evaluates their density and tortuosity.

Conclusions: The proposed method provides a completely automated analysis of corneal confocal images. Results demonstrate the potentiality of the CNN in identifying clinically useful features for corneal nerves by analysis of multiple images.

Key Words: corneal nerves, confocal microscopy, convolutional neural network, multiple-image

(*Cornea* 2019;00:1–6)

Accurate diagnoses in ophthalmology usually depend on the recognition of typical features and their changes in the main ocular structures, for example, nerve fibers in the corneal nerve plexus. Thanks to in vivo confocal microscopy (IVCM), nerves are imaged in a fast and noninvasive way. Because of these attributes, IVCM became an indispensable tool for studying corneal physiology and disease.

The cornea is one of the most densely innervated tissues in the human body,[1,2] and the correlation between morphometric parameters of corneal nerves (density, tortuosity, etc.) and a wide group of ocular (eg, dry eye syndrome) and systemic diseases (eg, diabetic neuropathy) has been widely demonstrated.[3–17]

A variety of methods for quantifying corneal nerve parameters have been recently proposed in the literature.[18] However, most of these methods are based on a manual or semiautomatic tracing of nerve fibers, which is a tedious, subjective (not reproducible), and time-consuming process. A few completely automatic software applications were developed, each with its own specific advantages and disadvantages (especially in low-quality images), and several morphological parameters were proposed.[19–27] However, none of them is currently used in clinical practice, and standard criteria for this analysis are still missing. In addition, because each image covers only a small part of the cornea, one single image may not fully describe the general corneal nerve appearance and thus may not be sufficient to achieve a correct diagnosis. To cope with this, wide-field mosaics can be built,[28,29] but their construction requires acquisition of several images, mostly overlapped, and a registration process that can create an artifact (eg, because of the misalignment of the images) in the final wide-field image. These are the main reasons why corneal nerve analysis has not become a widespread clinical practice yet, despite its great potentiality reported in the scientific literature.

This work bypasses the tracing problem and proposes a method based on a convolutional neural network (CNN), a deep learning technique, which directly associates IVCM images to healthy subjects or to subjects with diabetic neuropathy. The proposed method considers simultaneously different nonoverlapping images, so as to enlarge the analyzed corneal area, without building their mosaics.

CNNs are increasingly being used in the field of image processing, with interesting results, and also in the medical field,[30] particularly in ophthalmology.[31–34] This study investigates whether the CNN can also be successfully used for corneal nerve multiple-image analysis.

MATERIAL

In this study, confocal images of the subbasal corneal nerve plexus from 50 diabetic (type 1 or 2) with neuropathy and 50 age-matched healthy subjects (53 ± 13 years, 58 men/42 women) were used.

Received for publication July 8, 2019; revision received September 4, 2019; accepted September 6, 2019.
From the Department of Information Engineering, University of Padova, Padova, Italy.
The authors have no funding or conflicts of interest to disclose.
F. Scarpa and A. Colonna have contributed equally.
Correspondence: Fabio Scarpa, PhD, Department of Information Engineering, University of Padova, Via Gradenigo 6/b, 35131 Padova, Italy (e-mail: fabio.scarpa@unipd.it).

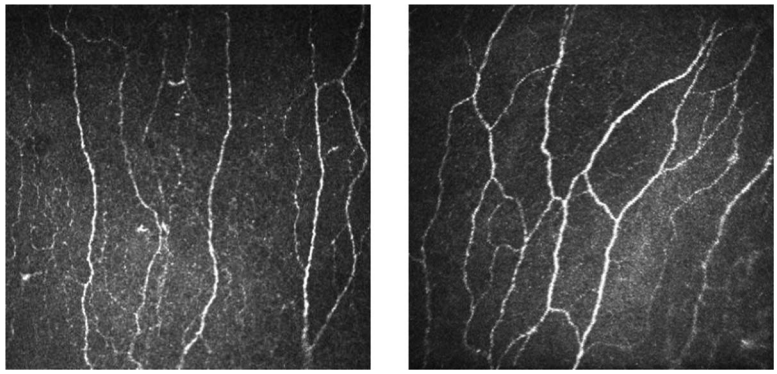

FIGURE 1. Representative images from a healthy (left) and a pathological (right) subject.

For each subject, 3 nonoverlapping images were considered from the left and right eye separately, acquired from the central part of the cornea, with no information about the mutual position of the images. Thus, the dataset of this study consists of 600 confocal images. They were acquired using the Heidelberg Retina Tomograph-II with the Rostock Cornea Module (Heidelberg Engineering GmbH, Heidelberg, Germany). Each image, saved in a gray-scale digital format, covers an area of 400 × 400 μm (384 × 384 pixels), showing a small region of the corneal nerve structure. Figure 1 shows 2 representative images.

Image acquisition was performed in different clinical centers, and each image was anonymized to delete any patient information. Because the acquisition of these images was approved by the respective local ethical review committees, occurred with informed consent, and followed the tenets of the Declaration of Helsinki, no specific further ethical approval was sought for the retrospective analysis of the resulting compilation of images.

METHODS

Preprocessing

In the peripheral area of images, problems can arise because of the characteristic curvature of corneal layers and the possible misalignment of the instrument to the corneal apex during image acquisition. Other common problems include spatial distortion, partial volume effect (sometimes corneal structures belonging to layers adjacent to the subbasal nerve plexus also appear in the image), illumination drift, and blurring.

Thus, the first step of the proposed automatic procedure consists of a crop of the most external area (10 pixels, empirically determined) of the analyzed image. Moreover, each image is resized by a factor of 0.7. The final size of the image is 256 × 256 pixels. The bicubic function used to resize the image also allows a partial noise reduction.

The dataset consists of 3 nonoverlapping images of the eye of each subject. Thus, to simultaneously analyze all images of each eye, 6 (ie, 3!) blocks of 3 images each were built, and within them, the order of the images was

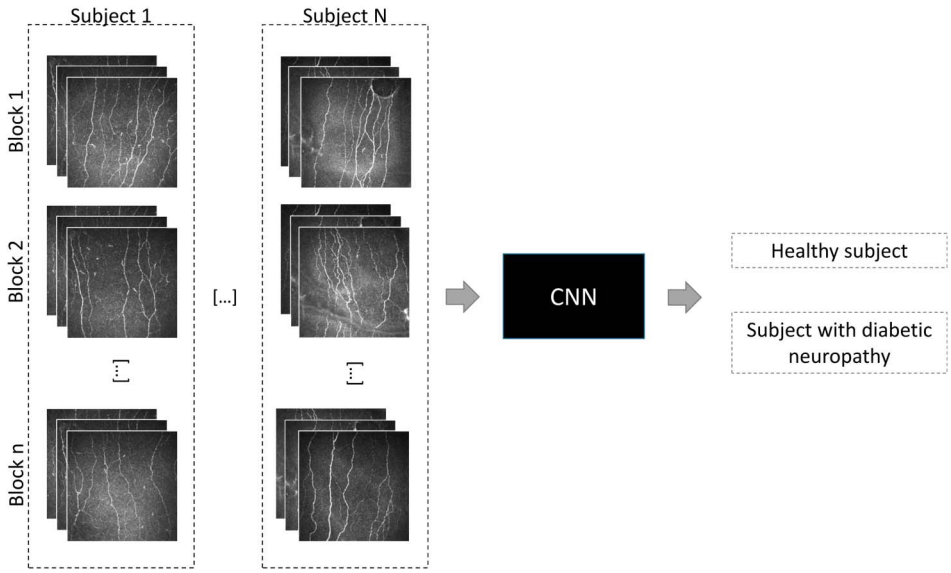

FIGURE 2. Diagram of the proposed algorithm. Input data correspond to N subjects, each composed by n blocks of 3 images. The algorithm uses the CNN to provide a binary classification (healthy/pathological).

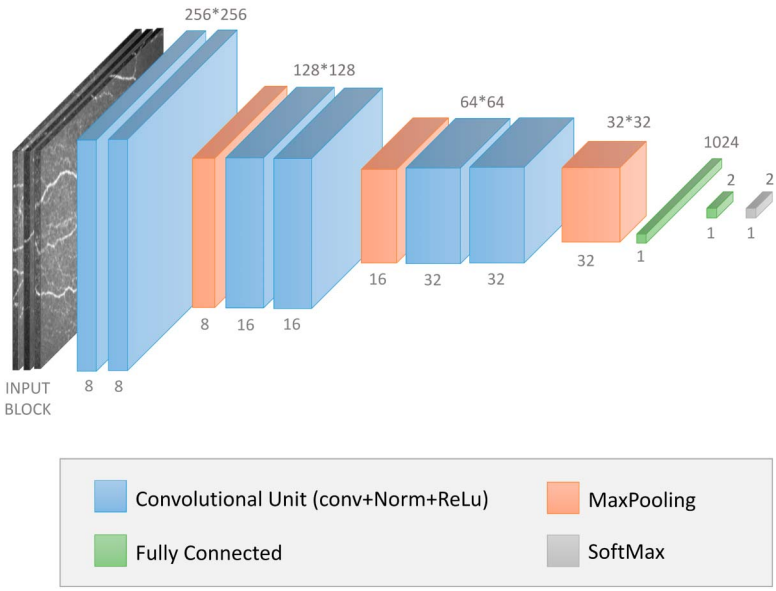

FIGURE 3. Convolutional neural network architecture for healthy/pathological classification. Input is a block of 3 images. Blue boxes correspond to a convolutional unit (which consist of a convolutional layer, followed by a batch normalization and a ReLU activation layer). Orange boxes correspond to activation maps obtained by max-pooling. The 2 green boxes represent the fully connected layers, and the gray one represents the softmax layer. The number of channels (number of filters/kernels) is denoted at the bottom of each box. The *xy*-size is provided at the top of each box.

rearranged. This technique should ensure that the final classification is not influenced by the order in which the images of each subject are considered because, in this dataset, there is no information about their mutual position in the cornea.

CNNs are known for their need of large datasets: for this reason, a data augmentation technique was used, consisting of flipping each image horizontally. Thus, for each block of 3 images, a new block obtained by horizontally flipping each of its images was derived.

In this way, 2400 blocks with 3 images each were used to train the CNN (12 blocks for each of the left and right eyes of 100 subjects), as shown in Figure 2.

Neuropathy Classification

The proposed CNN classifies each block of 3 images as healthy or pathological. The CNN is composed by a feature extraction part and a classification part (the architecture of the CNN is shown in Figure 3).

The first part of the architecture consists of subsequent application of convolutional units [convolution, batch normalization, and rectified linear unit (ReLU)] and down-sampling layers (max-pool layer). CNNs usually present an architecture in which the input size gets smaller and smaller from the start to the end of the network, whereas the number of channels increases. The proposed CNN progresses from a small number of filters (8 filters with kernel size 7×7, in the first 2 convolutional units), detecting low-level features, to a larger number of filters (32 filters with kernel size 3×3, in the last 2 convolutional units), each looking at different high-level features. The intermediate layers have 16 filters with kernel size 3×3. Zero padding is used during every convolutional layer. Moreover, after each convolutional layer, a ReLU layer is used to introduce nonlinearities, which allow models to learn patterns that are more complex. The output of the last layer is a $32 \times 32 \times 32$ representation of the input.

The achieved representation is supplied to 2 fully connected layers, which represent the classification part. After the final fully connected layer, a softmax nonlinearity is used to normalize the result into a binominal distribution over the 2 classes, healthy and pathological subjects. This layer also provides a score that represents the confidence of the final classification.

The parameters of the convolutional and fully connected layers were randomly initialized from a zero-mean Gaussian distribution. Different values of hyperparameters were tried: SGDM optimization with a momentum of 0.9 was adopted, the training was performed on 10 epochs, the batch size was set to 128, and the initial learning rate was set to 0.0001. After training, the classification of a multiple-image block took less than 1 ms.

A cross-validation strategy was used: training and evaluation were repeated 5 times, and each time, the CNN was trained on data from 80 subjects and evaluated on data from the other 20 subjects (10 healthy and 10 pathological each time). In this way, the CNN classified all available subjects, but each time it was trained on data in which the subjects (both left and right eyes) under test were not included.

TABLE 1. Performance Achieved by the Proposed CNN Either on a Single Block of 3 Images or on a Single Subject, Using Cross-Validation on 100 Subjects: Training on 80 Subjects and Evaluation on the Other 20, Repeated for 5 Times

	Mean, %	SD, %
Single block		
Accuracy	97	2
Sensitivity	98	3
Specificity	96	6
Whole subject		
Accuracy	96	4
Sensitivity	98	4
Specificity	94	9

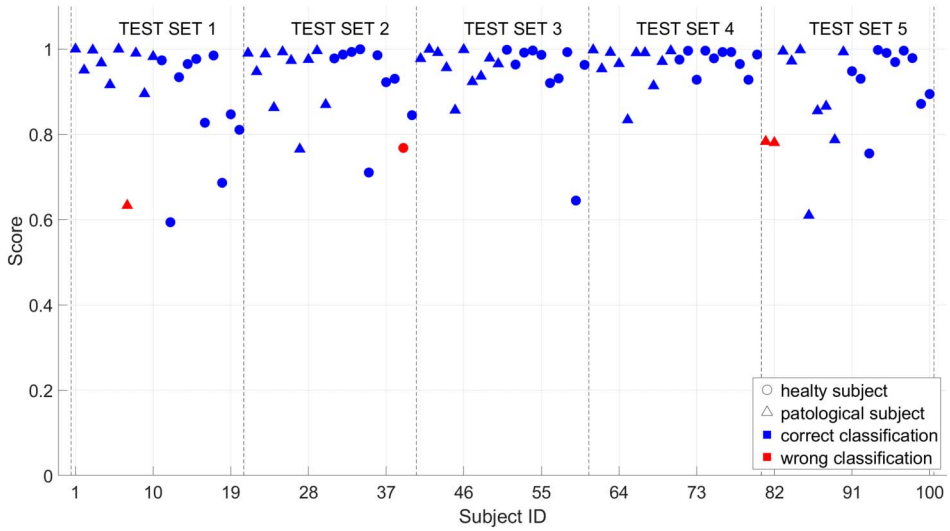

FIGURE 4. Classification scores for each subject. Mean scores of blocks are reported for each subject. Dots represent healthy subjects, while triangles represent pathological ones. Blue-colored shapes (dot or triangle) correspond to correctly classified subjects, and red-colored shapes correspond to incorrectly classified subjects.

RESULTS

Cross-validation was used to evaluate the proposed classification on 100 subjects: training on 80 subjects and evaluation on the other 20, repeated 5 times. Results were obtained from 5 different test sets, each composed of 20 subjects (10 healthy and 10 pathological).

For each subject, 12 blocks of 3 images were considered for the left and for the right eye, and each block was classified by the proposed CNN. Results from each block revealed a good capability of the CNN to achieve the correct classification. The mean accuracy is 97%, as shown in Table 1.

A final classification was derived for each subject, considering the prevalent classification of her/his blocks. A subject was considered properly classified if both right and left eyes were correctly classified. The validation of the proposed CNN resulted in a mean accuracy of 96% (96/100 subject were correctly classified).

Furthermore, the classification score for each subject was investigated: this score represents the probability to belong to the predicted class (high score = high probability to belong to the class). Results for each subject were derived by the mean probability of these blocks and are shown in Figure 4. The average score was approximately 93%, as shown in Table 2. Incorrect and correct classifications were also considered separately with a mean score of 94% for the correct ones and 74% for the incorrect being obtained. This classification score could therefore be used as an index of the reliability of the classification proposed by the CNN.

The proposed method was also compared with a method previously developed that analyzes the nerve density and morphology of each single image: nerve fibers are automatically traced,[25] and density and 2 tortuosity indexes (short-range and long-range tortuosity) are derived for each image.[13] Based on these morphological measures, a final classification is first derived for each image and finally for the whole subject (obtained by averaging the classifications of all its images). To compare the results, both methods were evaluated on the same dataset of 100 subjects. Density and tortuosity were different in healthy versus pathological subjects, as expected from the literature.[4,9,13,18,20,23,24,26] However, the accuracy obtained using nerve density and tortuosity in the classification of each subject was much lower than that obtained by the proposed method, as shown in Table 3. The CNN appears to be able to identify features that better describe neuropathy conditions.

TABLE 2. Classification Scores

Mean	SD	Minimum Value	Maximum Value
93%	9%	59%	100%

DISCUSSION

In recent years, IVCM has definitely increased the clinical interest in the corneal structure. IVCM allows acquiring images of various layers in the human cornea, in a fast and noninvasive way. Many studies revealed correlations between morphometric parameters of the corneal nerves and a wide group of ocular and/or systemic diseases. However, parameter estimation is based on a manual or semiautomatic tracing of nerves structures, which is a tedious, subjective (not reproducible), and time-consuming process. Difficulties in obtaining an accurate nerve tracing and long run times are likely the main reasons why corneal nerve analysis has not become part of the mainstream clinical diagnostic process yet.

The proposed method is based on a CNN. It aims to find the correlation between corneal nerve images and the presence of diabetic neuropathy, bypassing the tracing problem and obtaining a direct classification between healthy and pathological subjects. This work considered 3 corneal images of the left eye and 3 of the right eye, of 100 subjects (50 healthy and 50 pathological), using cross-validation to evaluate results.

The proposed method obtained a mean accuracy of 96% (96/100 subjects were correctly classified), demonstrating the potential of the CNN in identifying clinically useful features. These results outperform the ones obtained by traditional methods using nerve density and tortuosity automated estimation. These results denote that the CNN

TABLE 3. Classification Accuracy Achieved by the Proposed Multiple-Image Analysis and by a Previously Developed Automated Analysis of Nerve Density and Morphology on the Same Dataset (100 Subjects)

Proposed CNN	3 images simultaneously classified
	96%
Automated nerve tracing and parameter estimation	Mean classification of 3 images
Density	53%
Short-range tortuosity	62%
Long-range tortuosity	58%

provides a more complex analysis of the 3 images, which is more similar to the human clinical process than the one obtained by classifying each image separately and taking the average, as performed in previous works.[13,26] Indeed, clinical practice involves the observation of several images at the same time, looking for some features that reveal the presence of pathologies (in this case, neuropathy due to diabetes). These features could be present only in some of the acquired images (even only in one) because of the limited field of view of IVCM. Those features are, however, crucial for the final diagnosis, even if they are present in only 1 image out of 3. The proposed CNN analysis appears to reproduce what clinicians do during the diagnostic process, simultaneously analyzing multiple images and providing an overall classification for the whole subject.

An additional advantage of the proposed method is that analysis of multiple images is performed without the need for their montage/mosaicing, which requires acquisition of several images, mostly overlapped, is time consuming, and can introduce artifacts as a result of registration errors.

The dataset of this study consisted of 3 images for the eye of each subject; thus, the CNN was implemented to analyze 3 images simultaneously. However, better results may probably be achieved by increasing the number of simultaneously analyzed images.

The score associated with the classification of each block was also investigated. The proposed CNN presents high scores in correct classifications (94%), whereas scores are lower (74%) in incorrect classifications, as shown in Figure 4. The classification score could be used, as an example, to assess whether the proposed classification is reliable or whether more images are required.

In conclusion, the automated classification, obtained by the proposed CNN, and its confidence score provide an interesting quantitative analysis that could be useful for the diagnostic clinical process.

REFERENCES

1. Müller LJ, Marfurt CF, Kruse F, et al. Corneal nerves: structure, contents and function. *Exp Eye Res*. 2003;76:521–542.
2. Marfurt CF, Cox J, Deek S, et al. Anatomy of the human corneal innervation. *Exp Eye Res*. 2010;90:478–492.
3. Patel SV, McLaren JW, Hodge DO, et al. Confocal microscopy in vivo in corneas of long-term contact lens wearers. *Invest Ophthalmol Vis Sci*. 2002;43:995–1003.
4. Kallinikos P, Berhanu M, O'Donnell C, et al. Corneal nerve tortuosity in diabetic patients with neuropathy. *Invest Ophthalmol Vis Sci*. 2004;45:418–422.
5. Benítez del Castillo JM, Wasfy MA, Fernandez C, et al. An in vivo confocal masked study on corneal epithelium and subbasal nerves in patients with dry eye. *Invest Ophthalmol Vis Sci*. 2004;45:3030–3035.
6. Patel DV, McGhee CN. In vivo confocal microscopy of human corneal nerves in health, in ocular and systemic disease, and following corneal surgery: a review. *Br J Ophthalmol*. 2009;93:853–860.
7. De Cillà S, Ranno S, Carini E, et al. Corneal subbasal nerves changes in patients with diabetic retinopathy: an in vivo confocal study. *Invest Ophthalmol Vis Sci*. 2009;50:5155–5158.
8. Cruzat A, Pavan-Langston D, Hamrah P. In vivo confocal microscopy of corneal nerves: analysis and clinical correlation. *Semin Ophthalmol*. 2010;25:171–177.
9. Efron N. The Glenn A. Fry award lecture 2010: ophthalmic markers of diabetic neuropathy. *Optom Vis Sci*. 2011;88:661–683.
10. Ferrari G, Grisan E, Scarpa F, et al. Corneal confocal microscopy reveals trigeminal small sensory fiber neuropathy in amyotrophic lateral sclerosis. *Front Aging Neurosci*. 2014;6:278.
11. Winter K, Scheibe P, Köhler B, et al. Local variability of parameters for characterization of the corneal subbasal nerve plexus. *Curr Eye Res*. 2016;41:186–198.
12. Parissi M, Randjelovic S, Poletti E, et al. Corneal nerve regeneration after collagen cross-linking treatment of keratoconus: a 5-year longitudinal study. *JAMA Ophthalmol*. 2016;134:70–78.
13. Scarpa F, Ruggeri A. Development of clinically based corneal nerves tortuosity indexes. In: Jorge Cardoso M, Arbel T, Melbourne A, et al, eds. *MICCAI 2017. Lecture Notes in Computer Science*. Cham, Switzerland: Springer; 2017;10554:219–226.
14. Benkhatar H, Levy O, Goemaere I, et al. Corneal neurotization with a great auricular nerve graft: effective reinnervation demonstrated by in vivo confocal microscopy. *Cornea*. 2018;37:647–650.
15. Arrigo A, Rania L, Calamuneri A, et al. Early corneal innervation and trigeminal alterations in Parkinson Disease: a Pilot Study. *Cornea*. 2018;37:448–454.
16. Zemaitiene R, Rakauskiene M, Danileviciene V, et al. Corneal esthesiometry and sub-basal nerves morphological changes in herpes simplex virus keratitis/uveitis patients. *Int J Ophthalmol*. 2019;12:407–411.
17. Zhang J, Zhao Z, Shao C, et al. Degeneration of corneal sensation and innervation in patients with Facial Paralysis: a cross-sectional study using in vivo confocal microscopy. *Curr Eye Res*. 2019:1–7.
18. Dehghani C, Pritchard N, Edwards K, et al. Fully automated, semi-automated, and manual morphometric analysis of corneal subbasal nerve plexus in individuals with and without diabetes. *Cornea*. 2014;33:696–702.
19. Scarpa F, Grisan E, Ruggeri A. Automatic recognition of corneal nerve structures in images from confocal microscopy. *Invest Ophthalmol Vis Sci*. 2008;49:4801–4807.
20. Zhivov A, Winter K, Hovakimyan M, et al. Imaging and quantification of subbasal nerve plexus in healthy volunteers and diabetic patients with or without retinopathy. *PLoS One*. 2013;8:e52157.
21. Ziegler D, Papanas N, Zhivov A, et al. Early detection of nerve fiber loss by corneal confocal microscopy and skin biopsy in recently diagnosed type 2 diabetes. *Diabetes*. 2014;63:2454–2463.
22. Dabbah MA, Graham J, Petropoulos I, et al. Dual-model automatic detection of nerve-fibres in corneal confocal microscopy images. *Med Image Comput Comput Assist Interv*. 2010;13:300–307.
23. Dabbah MA, Graham J, Petropoulos IN, et al. Automatic analysis of diabetic peripheral neuropathy using multi-scale quantitative morphology of nerve fibres in corneal confocal microscopy imaging. *Med Image Anal*. 2011;15:738–747.
24. Petropoulos IN, Alam U, Fadavi H, et al. Rapid automated diagnosis of diabetic peripheral neuropathy with in vivo corneal confocal microscopy automated detection of diabetic neuropathy. *Invest Ophthalmol Vis Sci*. 2014;55:2071–2078.
25. Guimarães P, Wigdahl J, Ruggeri A. A fast and efficient technique for the automatic tracing of corneal nerves in confocal microscopy. *Transl Vis Sci Technol*. 2016;5:7.
26. Colonna A, Scarpa F, Ruggeri A. Segmentation of corneal nerves using a U-Net-Based convolutional neural network. In: Stoyanov D, et al, eds. *Computational Pathology and Ophthalmic Medical Image Analysis.*

OMIA 2018, COMPAY 2018. Lecture Notes in Computer Science. Cham, Switzerland: Springer; 2018;11039:185–192.
27. Sturm D, Vollert J, Greiner T, et al. Implementation of a quality index for improvement of quantification of corneal nerves in corneal confocal microcopy images: a Multicenter Study. *Cornea*. 2019;38:921–926.
28. Poletti E, Wigdahl J, Guimarães P, et al. Automatic montaging of corneal sub-basal nerve images for the composition of a wide-range mosaic. *Conf Proc IEEE Eng Med Biol Soc*. 2014;2014:5426–5429.
29. Lagali NS, Allgeier S, Guimarães P, et al. Wide-field corneal subbasal nerve plexus mosaics in age-controlled healthy and type 2 diabetes populations. *Sci Data*. 2018;5:180075.
30. Biswas M, Kuppili V, Saba L, et al. State-of-the-art review on deep learning in medical imaging. *Front Biosci (Landmark Ed)*. 2019;24:392–426.
31. Gulshan V, Peng L, Coram M, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316:2402–2410.
32. Grewal PS, Oloumi F, Rubin U, et al. Deep learning in ophthalmology: a review. *Can J Ophthalmol*. 2018;53:309–313.
33. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, et al. Artificial intelligence in retina. *Prog Retin Eye Res*. 2018;67:1–29.
34. Ting DSW, Pasquale LR, Peng L, et al. Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol*. 2019;103:167–175.