A Coarse to Fine Corneal Ulcer Segmentation Approach Using U-net and DexiNed in Chain

Helano Miguel B. F. Portela[1]([✉]), Rodrigo de M. S. Veras[1],
Luis Henrique S. Vogado[1], Daniel Leite[2], Jefferson A. de Sousa[3],
Anselmo C. de Paiva[3], and João Manuel R. S. Tavares[4]

[1] Departamento de Computação, Universidade Federal do Piauí,
Teresina, Brazil
`rveras@ufpi.edu.br`
[2] Departamento de Medicina Especializada, Universidade Federal
do Piauí, Teresina, Brazil
[3] Núcleo de Computação Aplicada,
Universidade Federal do Maranhão, São Luiz, Brazil
`paiva@nca.ufma.br`
[4] Instituto de Ciência e Inovação em Engenharia Mecânica
e Engenharia Industrial, Departamento de Engenharia Mecânica,
Faculdade de Engenharia, Universidade do Porto, Porto, Portugal
`tavares@fe.up.pt`

Abstract. A corneal ulcer is one of the most frequently appearing diseases that may affect eye health. The proper measurement of corneal ulcer lesions enables the physician to evaluate the treatment effectiveness and assist in decision-making. This article presents the solution for ulcer segmentation as a pixel-wise classification task, and proposes a novel coarse-to-fine method to extract corneal ulcers from ocular staining images. This study combines two classical convolutional neural networks (CNNs), known as U-net and DexiNed, following Morphological Geodesic Active Contour as a post-processing operation. We trained the CNNs using 358 point-flaky corneal ulcer images and evaluated its performance in 91 flaky corneal ulcer images. Our approach achieved 70.50% of Dice Coefficient on average, 87.4% of Recall, and 99.0% of Specificity, and True Dice Coefficient of 63.7%. These results corroborate our approach's efficacy and efficiency.

Keywords: Computer-aided diagnosis · Image segmentation · Deep learning · Eye health

1 Introduction

Many corneal diseases may affect eye health, such as Pytherigium, Infection, Conjunctival nevus and Corneal Ulcer. The corneal ulcer is one of the most frequently appearing of these, and it is defined as an inflammatory or even

more severe condition. It may lead in some cases to epithelial layer disruption or corneal stroma disruption. There are some potential causes of corneal ulcers, such as topical steroid usage, contact lens usage, trauma and ocular disorders, leading to perforation, scarring and vision loss [3]. Corneal ulcers can be classified into three general types, considering their shape and distribution: point-like, point-flaky mixed and flaky. Figure 1 shows samples of these three types.
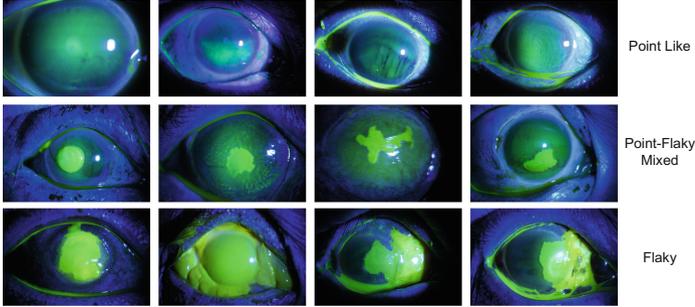

Fig. 1. Image samples from the SYSUTech-SYSU dataset, with the top row depicting point like corneal ulcers, the middle row depicting point-flaky mixed, and the bottom row depicting flaky corneal ulcers.

Usually, the point-like type appears at an early stage, when there are most chances of success in its treatment. This type of corneal ulcer has numerous ulcer dot distribution patterns that can appear anywhere within the corneal tissue. Therefore, it is not reasonably possible to segment it manually. A flaky corneal ulcer usually indicates a much more severe corneal disease. It has a uniform shape with clear boundaries, and may lead to scars and even vision loss. A point-flaky mixed corneal ulcer is a combination of point-like and flaky corneal ulcers. It indicates corneal disease with a severity degree, which lies between the aforementioned types. Measuring corneal ulcer lesion extension plays a crucial role in the treatment, as such a measurement may assist the specialist in the treatment follow-up.

The present study consisted of developing a computational method for corneal ulcer segmentation in ocular staining images. Hence, we evaluated and compared different CNN architectures and post-processing techniques found in the state-of-the-art. During the research process, we defined the following specific objectives: evaluate the U-net, DexiNed and LinkNet CNN architectures applied to this problem of image segmentation; estimate different pre- and post-processing image operations; and train the method using only point-flaky mixed corneal ulcer images, and validate it using flaky corneal ulcer images.

The remainder of this article has the following structure. Section 2 presents related works as to corneal lesion image segmentation; Sect. 3 details the evaluated CNN architectures, used image dataset, applied data preparing operations, and the adopted evaluation metrics. Section 4 presents the proposed approach.

Section 5 presents the results and discussion. Finally, Sect. 6 concludes the article and indicates future directions for research.

2   Related Works

We carried out a literature review looking for state-of-the-art articles related to computer-aided diagnosis solutions to segment corneal lesions. The survey aimed to identify and classify the works available in the literature based on the techniques employed, image dataset, year of publication and application domain. In this context, we can highlight the articles of Sun et al. [15], Deng et al. [5], Patel et al. [12], Deng et al. [4], Lima et al. [10], and Liu et al. [11].

We noticed that only the work of Sun et al. [15] uses a CNN-based approach to segment corneal ulcer images. In contrast, the other methods are based either on classical clustering or classification algorithms. In this context, our work contributes to exploring the limits of applying CNNs to segment corneal lesions, specifically corneal ulcer lesions.

Table 1 summarises the works found in the reviewed literature in terms of the year of publication, used technique(s), number of used images, and the application domain. In all of these works, the images used for training and test purposes were from the same dataset, and none of them used publicly available datasets, except for the CLID dataset used in Lima et al. [10].

Table 1. Summary of the works found in the reviewed literature in terms of the year of publication, used technique(s), number of used images and application domain.

Work	Year	Technique(s)	N. of images	Domain
Sun et al. [15]	2017	Path-based CNN	48	Corneal Ulcer
Deng et al. [5]	2018	SVM with Superpixel	150	Corneal Ulcer
Patel et al. [12]	2018	Random Forest and Active Contour	50	Corneal Ulcer
Deng et al. [4]	2018	Iterative k-means, Morphological Operations Region growing	48	Corneal Ulcer
Liu et al. [11]	2019	Gaussian mixture modeling and Otsu method	150	Corneal Ulcer
Lima et al. [10]	2020	Random Forest Classifier	30	Infection, Pterygium and Conjunctival nevus

3   Materials and Methods

This study aimed to propose an automatic method for corneal ulcer segmentation. We performed experiments using different combinations of the U-net, LinkNet and DexiNed CNNs architectures applied to the SUSTech-SYSU [6]

dataset. Combined with that, we used other post-processing techniques such as Binary Threshold, Otsu Threshold, Geodesic Active Contour, Fill holes and Morphological Operations. We evaluated the models' performance using four different metrics to identify the proposed method's best settings.

3.1   Evaluated CNN Architectures

U-net [14] is a convolutional neural network proposed for biomedical image segmentation. The general U-net architecture consists of two paths: a contracting path capable of capturing the image's context, and an expansive path capable of building the segmented image. The primary strategy that differentiates the U-net architecture from the other fully connected ones is combining the feature maps from the contraction layers with their symmetric correlated feature maps from the expansion layers. This characteristic allows the propagation of context information to high-resolution feature maps.

Chaurasia et al. [2] proposed **LinkNet** aiming to provide a semantic segmentation approach using less computational complexity comparing to other CNN architectures. LinkNet is based on encoders and decoders concepts, and is designed to perform a convolutional operation followed by a max-pooling operation on its output data; after that, there are four encoders blocks, followed by four decoders blocks. Finally, the architecture applies a sequence of full convolution, followed by a simple convolution and another full convolution as output.

DexiNed (Dense Extreme Inception Network for Edge Detection) [13] is a convolutional neural network for edge detection. It is built using a stack of filters that predict an edge map based on an input image. DexiNed comprises two sub-network architectures: Dense Extreme Inception Network (Dexi) and an up-sampling block (UB). Whereas the Dexi architecture has an image as input, the up-sampling block gets a feature map from the Dexi architecture block. The resulting architecture generates thin edge maps avoiding edge losses in the deep layers. DexiNed provides two outputs: Pred-a and Pred-f. The upsampling block returns six edge map outputs; by calculating the average from these six edge maps, one gets the Pred-a output, and by fusing these six edge maps, one gets the Pred-f output. In this work, we use DN-a and DexiNed-a to refer to the DexiNed model using the Pred-a output, and DN-f to refer to the DexiNed model using the Pred-f output.

3.2   Image Dataset

SUSTech-SYSU [6] is a dataset for automatically segmenting and classifying corneal ulcers from ocular fluorescein staining images. It was prepared to supply the lack of high-quality datasets to develop segmentation and classification algorithms for corneal diseases. The dataset contains 712 ocular staining images, and the segmentation ground truth of flaky corneal ulcers: 263, 358 and 91 images for point-like, point-flaky and flaky general types, respectively. The dataset also

provides the three-fold class labels for each image: 1) labels in terms of general ulcer pattern, 2) labels in terms of its specific ulcer pattern, and 3) the corresponding ulcer severity degree.

The work of Gross et al. [8] is the only official article published using the SUSTech-SYSU dataset. It describes a CNN based image classification approach to identify different types of Corneal Ulcers from fluorescein staining images.

3.3   Evaluation Metrics

In this work, we use the term 'positive' to designate ulcer areas and 'negative' to define non-ulcer areas. We calculated the confusion matrix to obtain the segmentation Recall (R), Specificity (S) and Dice Coefficient (DC).

In order to evaluate the segmentation quality, we calculated two more metrics: Average Dice Coefficient (ADC) and TDC (True Dice Coefficient). We can define the Average Dice Coefficient (ADC) as the mean value of all DC_i divided by the number of the images in a given dataset (n), i.e.:

$$ADC = \frac{\sum_{i=1}^{n} DC_i}{n}. \tag{1}$$

We consider that a "good" image segmentation result should have a DC value over a threshold (t). The True Dice Coefficient (TDC) metric for a dataset d is the number of automatic segmentation executions that achieved $DC_i > t$ divided by the total number of images. To compute the TDC metric, a score for each image (i) is calculated based on the Dice Coefficient as:

$$\{ score_i = 0, \text{if } DC < 0.7, score_i = 1, \text{otherwise.} \tag{2}$$

Given a dataset with n images, the final TDC value is defined as the mean of all per-image scores:

$$TDC = \frac{\sum_{i=1}^{n} score_i}{n}. \tag{3}$$

According to Genctav et al. [7], Dice scores greater than 0.7 indicate a remarkable similarity between the segmented and ground truth regions. So, we considered the threshold $t = 0.7$, to calculate the TDC.

4   Proposed Method

This work proposes an automatic segmentation method that combines two CNNs Architectures: U-net for image segmentation and DexiNed, which was initially proposed for edge detection, but it is here combined with the U-net output for image segmentation. We then apply 300 operations of the Morphological Geodesic Active Contour (MorphGAC) algorithm [1]. Figure 2 illustrates the proposed method's process.
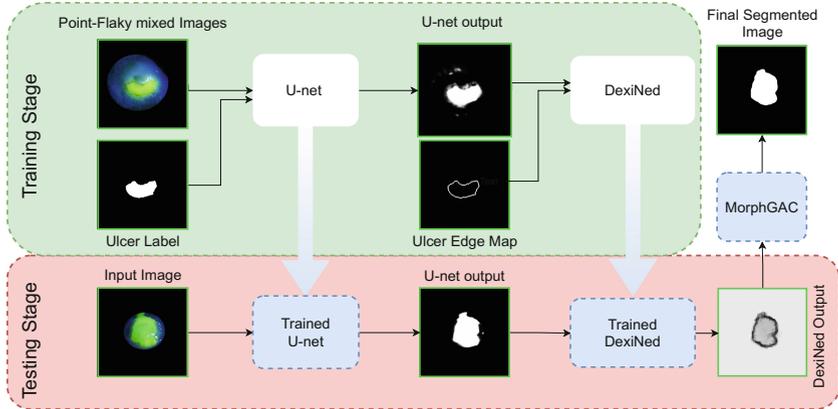
Fig. 2. Flowchart of the proposed method.

One should note that we do not propose any corneal segmentation method in this study. Instead, we focus exclusively on the corneal ulcer lesion segmentation, assuming that an automatic [4,9] or even manual [15] process had previously segmented the cornea. Therefore, the first performed preprocessing step is the corneal area segmentation using the corneal ground truth provided by the SUSTech-SYSU.

Once the training sets were prepared, we submitted the 358 point-flaky images and their corneal ulcer labels to the U-net. The U-net model was trained during 70 epochs using its classical architecture [14]. The number of 70 epochs was empirically defined, varying the number of epochs in increments of 10 until the model converges. The U-net output is a grayscale image that may contain undesirable segmented areas. Besides, a grayscale image is not the ideal final result for an image segmentation method. Thus, we then use the DexiNed to refine it.

To train the DexiNed model, we generated an edge map for each of the ulcer cornea labels. We did this by applying two iterations of erosion in the original ulcer label using a 3×3 structuring element. Then, we subtracted the resulting image from the erosion operation to the original ulcer label resulting in the ulcer edge map.

We submitted the 91 corneal flaky images (Fig. 3(a) to the trained U-net model. The U-net model (Fig. 3(b)) was then connected to the DexiNed input. Finally, on the DexiNed-a output (Fig. 3(c)), 300 iterations of the MorphGAC were executed to get the final result (Fig. 3(d)) using the Otsu threshold method to find the threshold value for the MorphGAC parameter. After that, we calculated the quality metrics using the final segmentation achieved by the proposed method (Fig. 3(d)) against the corresponding ground truth (Fig. 3(e)).
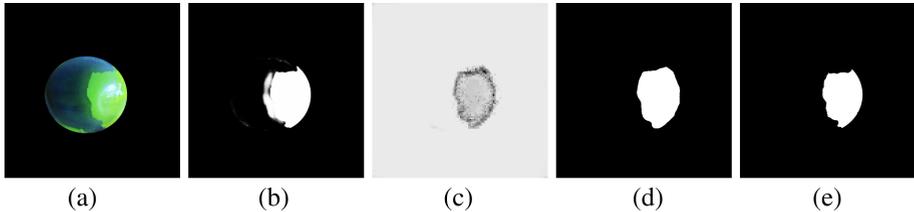
Fig. 3. Input image (a), U-net output (b), DexiNed-a output (c), Final result after submitted to MorphGAC (d), and corresponding ground truth provided by the SUSTech-SYSU dataset (e).

5   Results and Discussion

To find the best possible approach, we tested the combination of the U-net, DexiNed and LinkNet models, and each of these CNN models separately with different post-processing methods.

In the experiments, we used the following settings. **Dataset:** 358 Point-flaky images as training data and 91 Flaky images as testing data. **Architectures:** U-net, LinkNet, and DexiNed arranged according to the following settings: LinkNet connected to U-net, LinkNet itself, U-net connected to LinkNet, DexiNed connected to U-net, DexiNed itself, U-net connected to DexiNed and U-net itself. LinkNet and U-net were set with 1e-5 of learning rate and executed for 70 epochs, and DexiNed with a learning rate of 1e-4 and 1000 iterations at most.

Post-processing: For each of the architectures previously mentioned, we tested several combinations of Otsu threshold, MorphGAC and morphological operations.

Table 2 indicates the best results found for each combination. It is essential to point out that we did not use the point-like images for testing, because the used dataset does not provide the point-like corneal ulcer ground truth.

Table 2. Results obtained using combinations of the DexiNed (DN), LinkNet (LN) and U-net (UN) architectures, and MorphGAC, Otsu thresholding, Binary thresholding and Morphological operations. (Best values found in bold.)

Experimental settings	ADC (%)	R (%)	S (%)	TDC (%)
LN → UN	06.50 ± 0.069	00.00 ± 00.00	00.00 ± 00.00	00.00
LN → Otsu → fill holes	15.70 ± 17.00	40.50 ± 37.00	84.30 ± 30.10	00.00
UN → LN → Otsu	23.00 ± 20.10	38.10 ± 21.50	91.70 ± 19.90	00.00
DN-a → UN	26.90 ± 25.50	23.60 ± 29.10	95.50 ± 26.80	08.70
DN-a → Otsu → fill holes → DN-f	66.70 ± 25.50	91.20 ± 18.30	98.80 ± 01.10	49.40
DN-a → MorphGAC → Otsu value	68.30 ± 25.20	89.20 ± 19.30	$\mathbf{99.00 \pm 01.00}$	53.80
UN → binary threshold	70.30 ± 28.40	$\mathbf{96.40 \pm 11.80}$	98.30 ± 02.10	60.40
UN → Otsu → morph. op.	$\mathbf{74.10 \pm 27.10}$	92.10 ± 31.90	98.70 ± 01.90	62.60
UN → DN-a + MorphGAC	70.50 ± 25.10	87.40 ± 21.50	$\mathbf{99.00 \pm 01.10}$	**63.70**

As to the ADC metric, we obtained the best result (74.10%) with the U-net model trained for 70 epochs, with its output submitted to the Otsu threshold, fill holes operation and erosion, in this specific sequence and using a 3×3 structuring element for the erosion operation.

The results of each experiment for the Recall (R) metric are indicated in Table 2. The Recall is a critical metric as it represents the method's capability to identify pixels that correspond to the lesion area correctly. We achieved the best result (96.40%) using the U-net model combined with a classical binary threshold operation to get an utterly binary image from the U-net output directly.

The S column of Table 2 indicates the specificity metric results. This metric represents the method's capability to identify pixels that correspond to non-lesion areas correctly. We achieved a value of 99.00% for this metric as the best result by using the U-net model with its output connected to the DexiNed model. Using the DexiNed-a output, we applied MorphGAC using the Otsu threshold function to set its threshold value.

The same settings previously mentioned also achieved the TDC metric's best results with a value of 63.70%. This metric indicates the percentage of the testing data that the method could correctly segment, considering the criteria defined in Sect. 3.1.

As one can notice from the third row of Table 2, the results using the DexiNed-a output combined with the MorphGAC using the Otsu threshold parameter are promising. These results suggest that the DexiNed CNN architecture may be a reasonable option for image segmentation problems, although it was primarily designed for edge detection.

Figure 4 depicts the results of the proposed segmentation method (in red) overlapped with the correspondent ground truths (in white). The figure shows the worst (Figs. 4(a)), median (Figs. 4(b)) and best (Fig. 4(c)) cases obtained as to the Dice Coefficient.
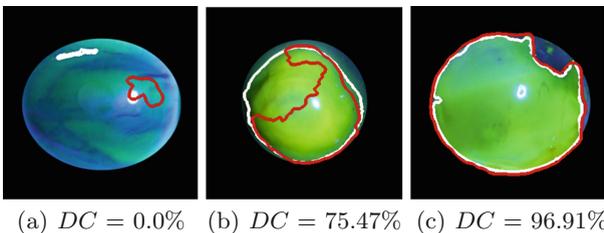

(a) $DC = 0.0\%$ (b) $DC = 75.47\%$ (c) $DC = 96.91\%$

Fig. 4. Examples of corneal ulcer images from the test dataset of 91 flaky corneal ulcers: The results of the proposed segmentation method (in red) are overlapped with the correspondent ground truths (in white). (Color figure online)

Although there are several approaches for corneal lesion segmentation, there was no benchmark image dataset in the literature. For a long time, this fact prevented a direct comparison between the existing methods over the years. Only

in 2020, two corneal lesion image datasets were publicly released: SUSTech-SYSU [6] (used in this work) and the CLID dataset [10]. Therefore, to the best of our knowledge, this is the first method applied to corneal ulcer segmentation using the SUSTech-SYSU dataset.

From Table 3, one can verify that the number of images used in this work is higher than the number of images included in the datasets used by Sun et al. [15], Deng et al. [4] and Lima et al. [10], and only lower than the number used in the works of Deng et al. [5] and Liu et al. [11]. It is essential to point out that our method uses different corneal ulcer lesions for training and testing stages using 449 images. Additionally, we evaluated it with all 91 flaky images available on the dataset, which are the only ones that have clear ground truth for validation.

Table 3. Comparison of the results obtained by the proposed method with the ones of the state-of-the-art.

Work	Dataset size	ADC (%)	R (%)	S (%)	TDC (%)
Sun et al. [15]	48	86.00 ± 07.30	82.00 ± 11.20	–	–
Deng et al. [5]	150	–	–	–	–
Deng et al. [4]	48	87.90	–	–	–
Lima et al. [10]	30	87.82	98.05	98.20	82.00
Liu et al. [11]	150	88.05 ± 06.11	–	–	–
Proposed Method	449	70.50 ± 25.10	87.40 ± 21.50	$99.00 \pm \pm 01.10$	63.70

Although the methods presented in the state-of-the-art had achieved better results, we believe that our method is relevant because it could generalize the features from the point-flaky images to segment the flaky corneal images. Based on that, we can train our method to segment point-like corneal ulcers, bringing up the possibility of using it to assist physicians in measuring point-flaky corneal ulcers. Not to mention that some of those methods are not entirely automatic as the one proposed by Lima et al. [10].

6   Conclusion

We proposed an automatic segmentation method for corneal lesion images that was applied to the SUSTech-SYSU image dataset. The new method uses two different CNNs: U-net, originally proposed for image segmentation, and DexiNed, initially designed for edge detection. We also tested various combined post-processing techniques (Binary Threshold, Otsu threshold, Fill holes and Geodesic Active Contour) to improve the CNN model outputs.

We found that the combination of the U-net output connected to the DexiNed model achieved better overall results when using the DexiNed-a prediction output with MorphGAC using the Otsu threshold value as parameter.

It is essential to point out that we only used point-flaky corneal images with not accurate ground truth to train our model. Considering that the used dataset does not provide the point-like ulcer ground truth, we used only the flaky corneal ulcer images to test the proposed models. However, we achieved encouraging results. Thus, we think that our model could generalize the training data (358 point-flaky corneal images) to segment the test data (91 flaky corneal ulcer images).

In the future, we intend to apply the proposed method on point-like corneal ulcer images and perform a manual validation by ophthalmologists. We believe that our method would be able to successfully segment point-like corneal ulcer images.

References

1. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comput. Vision **22**(1), 61–79 (1997)
2. Chaurasia, A., Culurciello, E.: LinkNet: exploiting encoder representations for efficient semantic segmentation. In: 2017 IEEE Visual Communications and Image Processing (VCIP), pp. 1–4. IEEE (2017)
3. Cohen, E.J., Laibson, P.R., Arentsen, J.J., Clemons, C.S.: Corneal ulcers associated with cosmetic extended wear soft contact lenses. Ophthalmology **94**(2), 109–114 (1987)
4. Deng, L., Huang, H., Yuan, J., Tang, X.: Automatic segmentation of corneal ulcer area based on ocular staining images. In: Medical Imaging 2018: Biomedical Applications in Molecular, Structural, and Functional Imaging, vol. 10578, p. 105781D. International Society for Optics and Photonics (2018). https://doi.org/10.1117/12.2293270
5. Deng, L., Huang, H., Yuan, J., Tang, X.: Superpixel based automatic segmentation of corneal ulcers from ocular staining images. In: 2018 IEEE 23rd International Conference on Digital Signal Processing (DSP), pp. 1–5. IEEE (2018)
6. Deng, L., Lyu, J., Huang, H., Deng, Y., Yuan, J., Tang, X.: The SUSTech-SYSU dataset for automatically segmenting and classifying corneal ulcers. Sci. Data **7**(1), 1–7 (2020)
7. Gençtav, A., Aksoy, S., Onder, S.: Unsupervised segmentation and classification of cervical cell images. Pattern Recogn. **45**(12), 4151–4168 (2012). https://doi.org/10.1016/j.patcog.2012.05.006
8. Gross, J., Breitenbach, J., Baumgartl, H., Buettner, R.: High-performance detection of corneal ulceration using image classification with convolutional neural networks. In: Proceedings of the 54th Hawaii International Conference on System Sciences, p. 3416 (2021)
9. Hezekiah, J.D., Chacko, S.: A review on cornea imaging and processing techniques. Current Med. Imaging **16**(3), 181–192 (2020)
10. Lima, P.V., et al.: A semiautomatic segmentation approach to corneal lesions. Comput. Electr. Eng. **84**, 106625 (2020)
11. Liu, Z., Shi, Y., Zhan, P., Zhang, Y., Gong, Y., Tang, X.: Automatic corneal ulcer segmentation combining Gaussian mixture modeling and Otsu method. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 6298–6301. IEEE (2019)

12. Patel, T.P., et al.: Novel image-based analysis for reduction of clinician-dependent variability in measurement of the corneal ulcer size. Cornea **37**(3), 331–339 (2018). https://doi.org/10.1097/ICO.0000000000001488
13. Poma, X.S., Riba, E., Sappa, A.: Dense extreme inception network: towards a robust CNN model for edge detection. In: The IEEE Winter Conference on Applications of Computer Vision, pp. 1923–1932 (2020)
14. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
15. Sun, Q., Deng, L., Liu, J., Huang, H., Yuan, J., Tang, X.: Patch-based deep convolutional neural network for corneal ulcer area segmentation. In: Cardoso, M.J., et al. (eds.) FIFI/OMIA -2017. LNCS, vol. 10554, pp. 101–108. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-67561-9_11