

AES-CSFS: an automatic evaluation system for corneal sodium fluorescein staining based on deep learning

Shaopan Wang^{*} , Jiezhou He^{*}, Xin He^{*}, Yuwen Liu^{*}, Xiang Lin, Changsheng Xu, Linfangzi Zhu, Jie Kang, Yuqian Wang, Yong Li, Shujia Guo, Yunuo Zhang, Zhiming Luo and Zuguo Liu

Abstract

Background: Corneal fluorescein sodium staining is a valuable diagnostic method for various ocular surface diseases. However, the examination results are highly dependent on the subjective experience of ophthalmologists.

Objectives: To develop an artificial intelligence system based on deep learning to provide an accurate quantitative assessment of sodium fluorescein staining score and the size of cornea epithelial patchy defect.

Design: A prospective study.

Methods: We proposed an artificial intelligence system for automatically evaluating corneal staining scores and accurately measuring patchy corneal epithelial defects based on corneal fluorescein sodium staining images. The design incorporates two segmentation models and a classification model to forecast and assess the stained images. Meanwhile, we compare the evaluation findings from the system with ophthalmologists with varying expertise.

Results: For the segmentation task of cornea boundary and cornea epithelial patchy defect area, our proposed method can achieve the performance of dice similarity coefficient (DSC) is 0.98/0.97 and Hausdorff distance (HD) is 3.60/8.39, respectively, when compared with the manually labeled gold standard. This method significantly outperforms the four leading algorithms (Unet, Unet++, Swin-Unet, and TransUnet). For the classification task, our algorithm achieves the best performance in accuracy, recall, and F1-score, which are 91.2%, 78.6%, and 79.2%, respectively. The performance of our developed system exceeds seven different approaches (Inception, ShuffleNet, Xception, EfficientNet_B7, DenseNet, ResNet, and ViT) in classification tasks. In addition, three ophthalmologists were selected to rate corneal staining images. The results showed that the performance of our artificial intelligence system significantly outperformed the junior doctors.

Conclusion: The system offers a promising automated assessment method for corneal fluorescein staining, decreasing incorrect evaluations caused by ophthalmologists' subjective variance and limited knowledge.

Keywords: artificial intelligence, corneal epithelium defect, deep learning, image analysis, sodium fluorescein staining

Received: 3 June 2022; revised manuscript accepted: 13 December 2022.

Introduction

The cornea is located in the anterior part of the eyeball. It is directly exposed to the external environment, which can easily be influenced by

traumatic factors,¹ environmental factors,² and infectious and inflammatory diseases.³ The corneal epithelium, as the outermost physiological structure of the cornea, is particularly

Ther Adv Chronic Dis

2023, Vol. 14: 1–12

DOI: 10.1177/
20406223221148266

© The Author(s), 2023.
Article reuse guidelines:
[sagepub.com/journals-](https://sagepub.com/journals-permissions)
[permissions](https://sagepub.com/journals-permissions)

Correspondence to:

Zuguo Liu
Institute of Artificial
Intelligence, Xiamen
University, 4th Floor, 4221-
122, South Xiang'an Road,
Xiamen 361102, Fujian,
China

Eye Institute of Xiamen
University, School
of Medicine, Xiamen
University, Xiamen, China.

Fujian Provincial
Key Laboratory of
Ophthalmology and
Visual Science, Xiamen
University, Xiamen, China

Department of
Ophthalmology, Xiang'an
Hospital of Xiamen
University, Xiamen,
Xiamen University, Xiamen, China

Xiamen Eye Center of
Xiamen University, Xiamen
University, Xiamen, China.
zuguo.liu@xmu.edu.cn

Zhiming Luo
Institute of Artificial
Intelligence, Xiamen
University, Xiamen, China

School of Informatics,
Xiamen University, 422
Siming South Road,
Xiamen 361005, Fujian,
China.

zhiming.luo@xmu.edu.cn

Shaopan Wang
Changsheng Xu
Institute of Artificial
Intelligence, Xiamen
University, Xiamen, China

Eye Institute of Xiamen
University, School
of Medicine, Xiamen
University, Xiamen, China

Fujian Provincial
Key Laboratory of
Ophthalmology and
Visual Science, Xiamen
University, Xiamen, China

Jiezhou He
Institute of Artificial
Intelligence, Xiamen
University, Xiamen, China

Xin He
Eye Institute of Xiamen University, School of Medicine, Xiamen University, Xiamen, China
Fujian Provincial Key Laboratory of Ophthalmology and Visual Science, Xiamen University, Xiamen, China
Department of Ophthalmology, The First Affiliated Hospital of Xiamen University, Xiamen, China

Yuwen Liu
Eye Institute of Xiamen University, School of Medicine, Xiamen University, Xiamen, China
Fujian Provincial Key Laboratory of Ophthalmology and Visual Science, Xiamen University, Xiamen, China
National Institute for Data Science in Health and Medicine, Xiamen University, Xiamen, China

Xiang Lin
Jie Kang
Department of Ophthalmology, Xiang'an Hospital of Xiamen University, Xiamen University, Xiamen, China

Linfangzi Zhu
Yuqian Wang
Yong Li
Shujia Guo
Yunuo Zhang
Eye Institute of Xiamen University, School of Medicine, Xiamen University, Xiamen, China
Fujian Provincial Key Laboratory of Ophthalmology and Visual Science, Xiamen University, Xiamen, China

*These authors contributed equally to this work and thus should be considered co-first authors.

problematic.⁴⁻⁶ Ophthalmologists used corneal fluorescein sodium staining to test the integrity of the corneal epithelium.⁷ Different staining shapes, the number of stained dots, and the size of the staining area are often used to distinguish various illnesses or disease severity. In addition, accurate measurement of the defect area is essential to assess the corneal repair function and the effectiveness of the treatment.^{8,9} However, the evaluation of corneal sodium fluorescein staining relies heavily on the subjective perception of the clinician. Therefore, there exists a great need for automated quantitative analysis of corneal sodium fluorescein staining.

Researchers have accomplished a wide variety of related work. Pritchard *et al.*¹⁰ developed a method for digital imaging to evaluate punctate corneal staining. Gaussian edge detection and RGB color system were combined to detect corneal erosions.¹¹ The corneal ulcer area was identified by *k*-means clustering.¹² A joint Otsu and Gaussian mixture modeling method was performed to segment corneal ulcers.¹³ The limitations of these studies are the relatively small amount of data studied, traditional insufficient digital technology, or the insufficient generalization capacity of the model.

With the advent of massive data, the deep neural network has wildly succeeded in various visual tasks¹⁴⁻¹⁹ and achieved satisfactory results in the last decades. In the field of ophthalmology, deep neural network methods have been used for optic disk segmentation,²⁰ cataract image classification,²¹ eye tracking,²² detection of diabetic retinopathy,²³ and iris segmentation,²⁴ which brings great convenience to multiple clinical tasks, such as disease screening, diagnosis, and prevention.

In this study, we showcase a novel and effective way to quantify corneal sodium fluorescein staining, which can accurately evaluate the sodium fluorescein staining of the entire corneal area. Furthermore, the evaluation system can automatically detect the corneal stained area. Our experimental results demonstrated that deep learning-based artificial intelligence technology could reduce the variability of clinicians in interpreting corneal staining results and improve the reliability of measurement results, which provides a powerful tool for disease management.

Materials and methods

Image data set collection

We stained the patient's cornea after moistening the fluorescein impregnated strips with a drop of sterile normal saline without applying a local anesthetic. Subsequently, the patient was asked to blink three times to ensure that the fluorescein dye was evenly applied throughout the tear film. Images were taken with a digital slit-lamp camera 3 min after the instillation of fluorescein sodium dye solution. The image was saved in JPG format, 1624 × 1232 in size. In our experiment, we focused on the staining of the corneal region, so that, each photo was designed to put the corneal part in the center of the image. Considering the model's generalization ability, images at different magnifications are retained. In this prospective study, 1896 slit-lamp images were consecutively collected between October 2021 and May 2022.

Inclusion and exclusion criteria

Inclusion criteria: Patients with corneal epithelial injury of any cause should be included in this study.

Exclusion criteria: Patients with the following conditions will be excluded from the research plan because these factors may lead to false positive sodium fluorescein staining: severe corneal trauma, corneal mass, corneal edema, post-keratoplasty, pterygium, corneal leakage, and secretion.

Data set division

In the field of machine learning or deep learning, it is very important to divide the training set, validation set, and test set reasonably. To reduce information leakage and improve the efficiency of the reaction model, we split the data set in this experiment into a training set, validation set, and test set at a ratio of 7:1:2. One image per patient was selected to form the data set. Each image is augmented with a random enhancement algorithm to make the training model more generalizable. The details of each set of various computer vision tasks are presented in Table 1.

Image preprocessing

We labeled the corneal region with Labelme software (version 4.6.0). The corneal region was

Table 1. Data set division is used in this study.

Tasks	All image	Train image >	Validation image	Test image
Cornea segmentation	363	254	36	73
Lesion segmentation	254	249	35	70
Classification score	1279	895	128	256

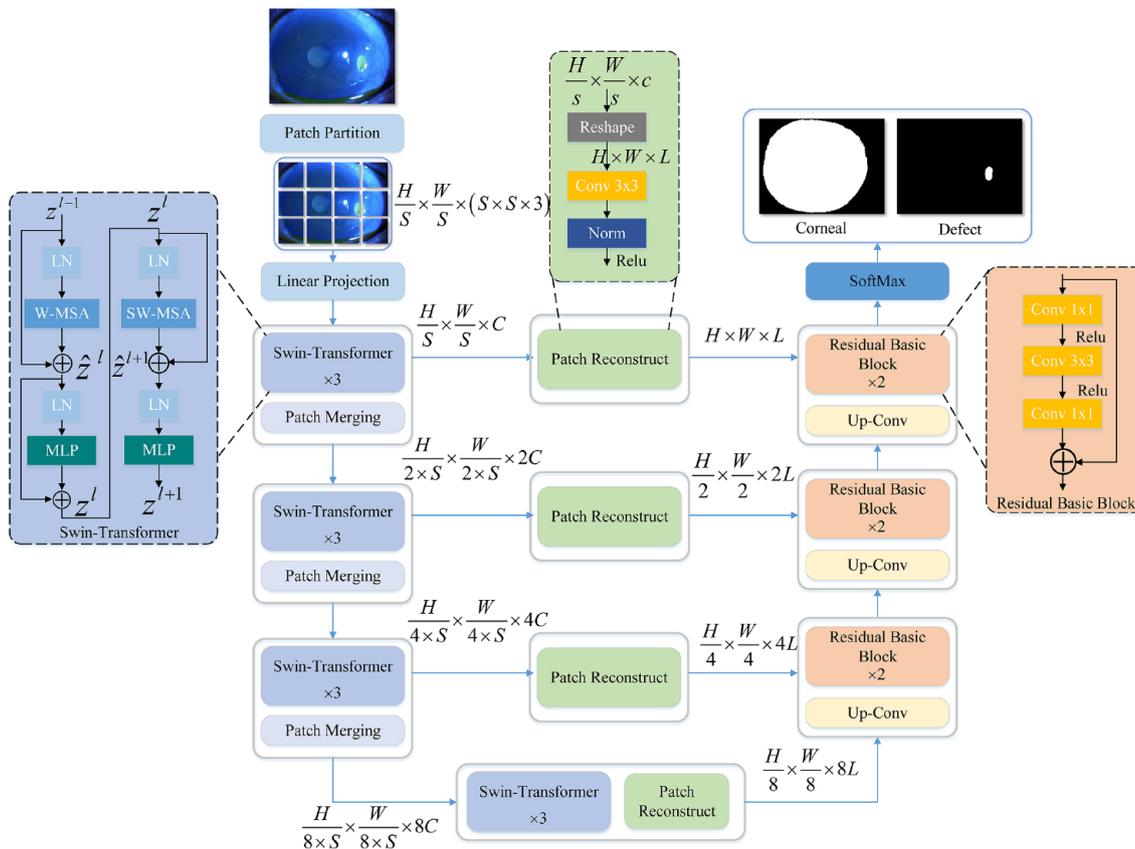


Figure 1. Images segmentation neural network structure for cornea boundary and corneal lesion region. C, channels of images; Conv., convolution; H, the height of the image; Norm, normalization; Up-Conv, up convolution; W, the weight of images.

labeled as the foreground, while the other areas were labeled 0 as the background. Because of the particular corneal structure, padding resize is employed to transmit it into the neural network for segmentation. We used the steps in Figure 1 to extract the corneal region. Patchy defects in corneal epithelium significantly affect the patient's vision and, in severe cases, lead to blindness in patients.^{25,26} We selected bigger corneal epithelial defects for manual labeling; 254/249 images were used to train cornea and patchy defect area

segmentation, respectively, and 73/70 images were used to verify the performance of model training.

Image segmentation

We proposed a novel U-shaped structure segmentation network for successfully segmenting the cornea and huge defect region. The network employed the Swin-Transformer²⁷ to replace the standard convolution process, allowing for a more

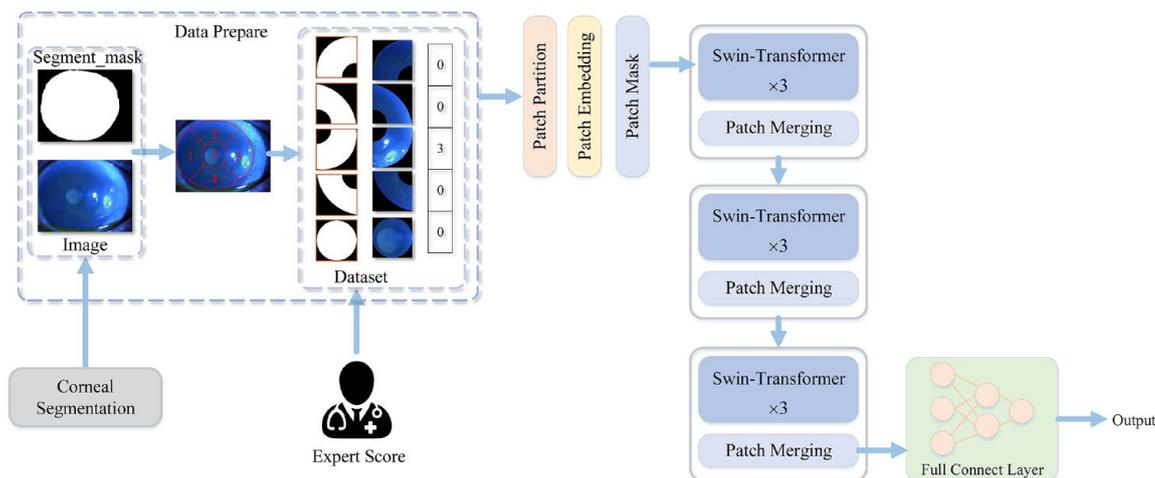


Figure 2. Flow chart of corneal fluorescein sodium staining score system.

significant consideration of global information in feature extraction. Simultaneously, Swin-Transformer has a quicker reasoning speed than other vision transformers' (ViT) structures.²⁸ To extract image features at various scales, we employed a patch merge layer to accomplish 2×2 feature map scaling. A convolutional neural network was still used for feature recovery in the upper sampling part to ensure the image edge's detailed features were better. The structure is shown in Figure 1.

Automatic corneal staining scoring system

We divided the extracted corneal region into five subdomains. We scored the stained images according to the NEI Workshop on Clinical Trials in Dry Eyes,²⁹ and two experienced ophthalmologists scored each subdomain. When a picture was consistently scored, it was included in the study. We used the framework in Figure 2 to train a four-classification neural network model, which directs the allocation of five subareas of the cornea to one of the four grades of 0–3. Of the 1279 images, 895 scored by the experts in total were used for the training of the model. The remaining photos were used for the performance test of the neural network.

Implementation details

The software used in this experiment is Python (version 3.7.11). The network architecture is built on PyTorch (version 1.10.1), an open-source

Python machine learning library primarily used in deep learning research. PyTorch supports GPU computing, which significantly speeds up application execution. This is especially important in the era of big data. In this experiment, we used four RTX2080Ti graphics cards.

Model evaluation and data analysis

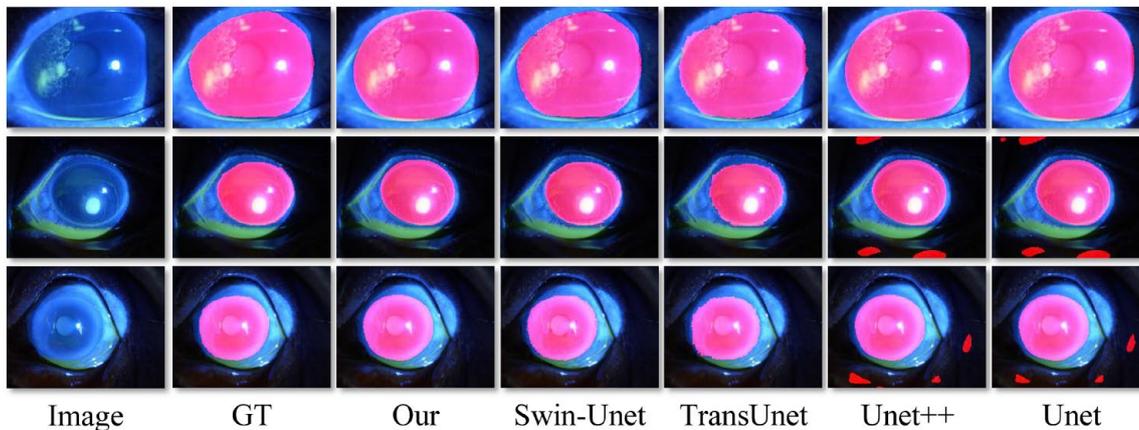
Segmentation: Two commonly used objective criteria, dice similarity coefficient (DSC)³⁰ and the Hausdorff distance (HD),³¹ are used to evaluate the medical image segmentation performance quantitatively. The DSC is often used to calculate the similarity of closed regions. The value range is [0,1]. The higher the value, the nearer the two sets are. HD calculates the distance between two sets. The smaller the value, the greater the similarity between the two sets. Images labeled by experts are considered the ground truth (GT).

Classification: In the classification task, accuracy means the percentage of correctly classified instances, precision represents the proportion of the true positives against all the positives, recall reveals the ratio of the true positive against the sum of true positive and false negative, while F1-score shows the harmonic mean of precision and sensitivity.^{32,33} Here, we used the above four indicators to evaluate the performance of the corneal fluorescein sodium staining score system, also called four-classification neural network.

Table 2. Quantitative evaluation of corneal boundary segmentation performance.

Indicators	Our	Swin-Unet	TransUnet	Unet++	Unet
DSC	0.98236	0.97284	0.84684	0.88959	0.78959
HD	3.60555	5.12714	10.22517	10.16368	23.16368

DSC, dice similarity coefficient; HD, Hausdorff distance; Our, our proposed method. Bold values indicate that the result is the best of this metric.

**Figure 3.** Performance of corneal area segmentation on various methods.

Visualization results of corneal region segmentation by the above five different segmentation methods. From left to right: original image, ground truth, Swin-Unet, TransUnet, Unet++ , Unet. Red represents the region segmented from the neural network model.

DSC, dice similarity coefficient; HD, Hausdorff distance; GT, ground truth; Our, our proposed method.

Results

Segmentation of corneal boundary

In this research, we focused on the staining of the corneal region. Therefore, we first identified the corneal boundary from the original image using a segmentation model based on Swin-Transformers (Figure 1). In our approach, the DSC for identification of cornea boundary was 0.9824, while HD was 3.605 (Table 2). In addition, classic and excellent performance methods of medical image segmentation, such as Unet,³⁴ Unet++,³⁵ Swin-Unet,³⁶ and TransUnet,³⁷ were performed to compare with our proposed method in the task of corneal region segmentation. Compared with Unet, Unet++, TransUnet, and Swin-Unet can improve DSC by 12.6%, 7.25%, and 23.2%, respectively, while decreasing HD by 56.1%, 55.9%, and 77.9%, respectively. Among the methods mentioned above, our proposed method achieves the best performance; it can increase DSC by 24.4% and decrease HD by 84.4%.

In addition, we presented visualization results of corneal segmentation using five different segmentation methods (Figure 3). Visually, we can intuitively find that the segmentation performance of our process is the closest to GT. In addition, it can be seen that our model has good performance in corneal staining images of different enlarged sizes, which indicates that our model has good generalization ability.

Segmentation of corneal lesion area

Similar to the corneal boundary segmentation task, we used several different methods the above-mentioned for patchy corneal staining region segmentation. Using our way, the DSC for detecting corneal staining lesions was 0.9776, while HD was 8.3909 (Table 3). Compared with Unet, the three methods, including Unet++, TransUnet, and Swin-Unet, can improve DSC by 20.3%, 18.5%, and 22.9%, respectively, while decreasing HD by 37.1%, 35.3%, and 66.9% respectively.

Table 3. Quantitative evaluation of patchy corneal lesion segmentation performance.

Indicators	Our	Swin-Unet	TransUnet	Unet++	Unet
DSC	0.97763	0.91376	0.88089	0.89431	0.74340
HD	8.39098	12.26036	23.97022	23.32689	37.06745

DSC, dice similarity coefficient; HD, Hausdorff distance; Our, our proposed method. Bold values indicate that the result is the best of this metric.

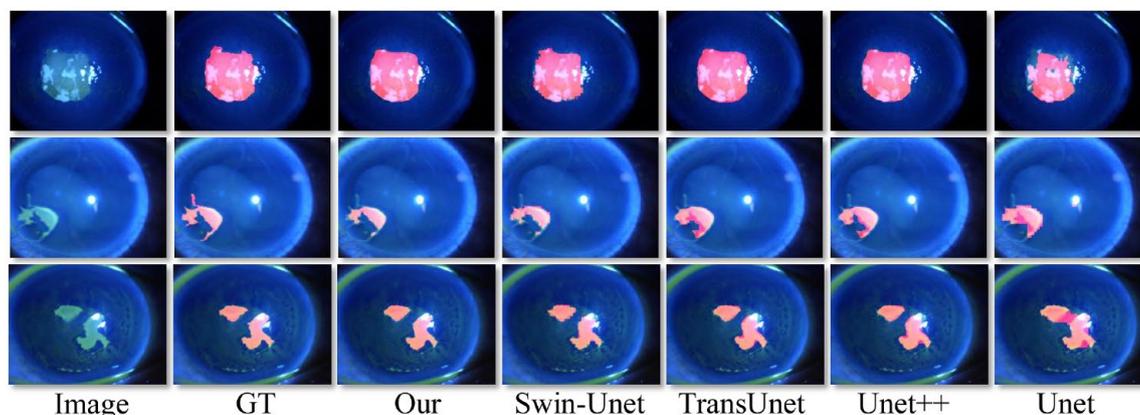


Figure 4. Performance of detection of patchy corneal staining region on several methods. Visualization results of patchy lesion area segmentation using the above five different segmentation ways. From left to right: original image, ground truth, our proposed method, Swin-Unet, TransUnet, Unet++, Unet. Red represents the region segmented from the model. DSC, dice similarity coefficient; HD, Hausdorff distance; GT, ground truth; Our, our proposed method.

Our proposed method increased DSC by 31.5% and reduced HD by 77.4%. We finally achieved the best performance in identifying corneal staining lesions by far surpassing the second place.

We visually generated the corneal patchy defect area's mask (the red region) and drew it automatically on the original image. Three representative segmentation outcomes with different methods are shown in Figure 4. It can be seen that whether the defect shape is complex or simple, our model can reasonably identify the location and boundary of the cornea defect. Our proposed method achieved the best segmentation results.

Corneal fluorescein sodium staining score system

A score for each cornea subdomain was obtained using the method in Figure 2. Accuracy, as the most commonly used index in classification models, cannot reasonably reflect the prediction ability of models when there are unbalanced

samples. Therefore, we also introduced precision, recall, and F1-score to more comprehensively evaluate the performance of the classification model. Inception,³⁸ ShuffleNet,³⁹ Xception,⁴⁰ EfficientNet_B7,⁴¹ DenseNet,^{42,43} ResNet, and ViT⁴⁴ have been used for medical image classification tasks and achieved promising performance. Here, compared with the other seven methods, our developed model achieves the best performance in accuracy, precision, and F1-score, which are 91.2%, 78.6%, and 79.2%, respectively, and the implementation of precision achieved 84.244% (Table 4).

Comparison of artificial intelligence technology and ophthalmologists in corneal staining evaluation

We randomly selected three doctors who were asked to rate corneal staining images. Then, we compared the evaluation results of these three doctors with the results of machine learning. Our developed artificial intelligence-based approach

Table 4. Performance of corneal fluorescein staining evaluation system using different frameworks.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Inception	85.7	73.5	63.9	68.5
ShuffleNet	69.1	35.6	40.3	37.8
Xception	87.4	76.2	45.3	56.8
EfficientNet_B7	80.5	84.2	49.6	62.4
DenseNet	79.1	57.8	69.2	63.0
ResNet	85.9	91.4	63.2	74.7
VIT+ mask	87.6	82.1	73.5	77.6
Our	92.5	86.3	79.2	82.6

Our, our proposed method. Bold values indicate that the result is the best of this metric.

Table 5. Comparison of artificial intelligence system with various ophthalmologists in corneal staining evaluation results.

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
Our	91.2	84.2	78.6	79.2
Junior ophthalmologist 1	53.9	26.4	29.7	27.3
Junior ophthalmologist 2	52.0	26.0	25.0	24.9
Junior ophthalmologist 3	58.0	45.8	56.7	42.4
Senior ophthalmologist 1	90.5	84.9	76.4	80.4
Senior ophthalmologist 2	91.3	84.2	75.9	79.8

Our, our proposed method. Bold values indicates that the result is the best of this metric.

far outperformed junior ophthalmologists on all four measures. In addition, we can see a big difference between the three ophthalmologists in the evaluation of corneal staining, especially in precision, recall, and F1-score. Table 5 shows their respective performance.

Discussion

Corneal fluorescein sodium staining is the most commonly used way in ophthalmology to assess corneal integrity. The evaluation of corneal staining and accurate automatic detection of the stained lesion area is very important for various clinical purposes, such as the evaluation of patient symptoms, quantitative study of lesions, monitoring of drug treatment, and management of the patient's condition.⁴⁵

At present, the emerging field of deep learning has exerted a significant impact on medical imaging. As effective methods used in biomedical image segmentation, Unet, Unet++, TransUnet, and Swin-Unet deep learning frameworks have achieved promising performance on many medical vision tasks, such as iris segmentation,⁴⁶ pupil segmentation,⁴⁷ retinal vessel segmentation,⁴⁸ skin lesion segmentation, and classification.⁴⁹ However, these methods cannot model global features well and are prone to lose a wealth of detailed local information. Deep learning techniques suitable for massive data have yet to be fully utilized to solve the aforementioned clinical problem.

This article introduces an artificial intelligence method based on deep learning to automatically

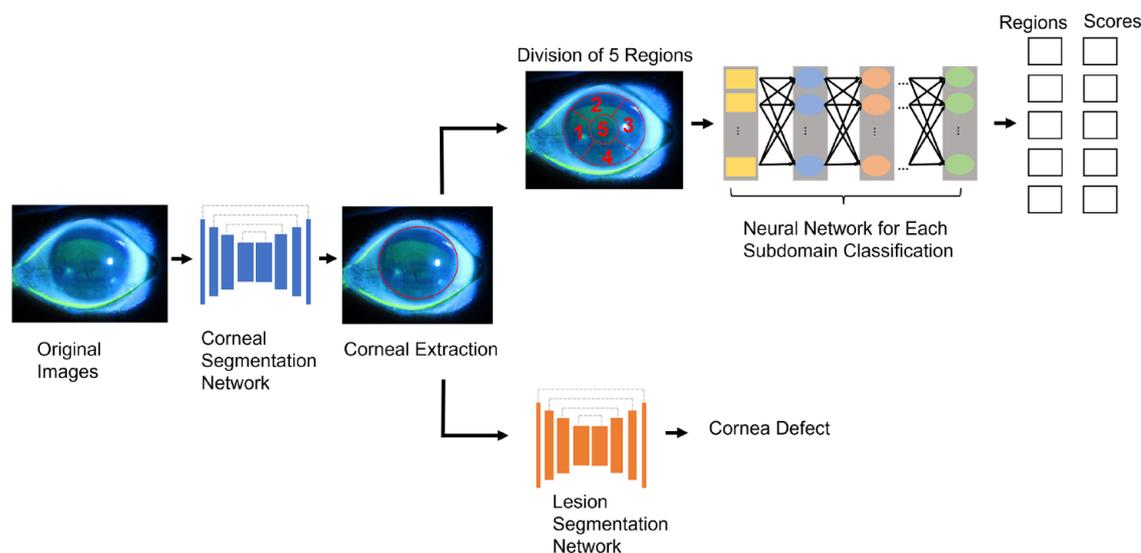


Figure 5. Schematic overview of the entire corneal staining evaluation system.

The corneal region was manually labeled with a small amount of original image to train a neural network model to identify the corneal boundary accurately. For the extracted cornea, first of all, we divided the cornea into five regions, with 0–3 grades in each subdomain, sequentially manual scoring, and then trained a four-part neural network model to evaluate the score of corneal fluorescein sodium staining automatically. Furthermore, for the staining pictures of patchy defects, we leveraged the image segmentation method to train a neural network model to segment the patchy defects automatically.

and accurately assess corneal fluorescein sodium staining. As shown in Figures 3 and 4, our proposed method can perfectly identify the corneal region and the corneal epithelial defect area. The best performance of the segmentation task may result from Swin-Transformer. It is used as the backbone in the segmentation model’s down-sampling part. It can model global features better than the backbone based on convolution. Patch merging is also used for scaling to extract global features at different sizes and decrease computer resources. The patch reconstructs module transforms the patch-based feature map into the standard feature matrix. In the up-sampling part, residual basic is used as the foundation to achieve pixel-level restoration. Finally, the corneal and lesion region segmentation is accomplished.

In contrast to the typical convolution-based classification model, the classification model’s backbone uses a Swin-Transformer, which pays more attention to the context information and can have global modeling features. Based on the segmentation results, this article uses a patch mask to mask the padding region. It does not involve feature modeling to decrease interference to the target region and lower the degree of computation. Therefore, we achieved the best accuracy, recall, and F1-score (Table 4). As for failing to reach the

best precision, we argue that the transformer model has global features, and the input image is significantly unbalanced, which results in poor performance of our proposed model in this task. Moreover, the mask mechanism is also introduced into the traditional VIT model (VIT+mask) to optimize it to ensure the fairness of the comparative experiment. Ultimately, our artificial intelligence system outperformed the young ophthalmologists (Table 5). We hypothesize that the following reasons may explain the above phenomenon. First of all, the excess sodium fluorescein dye on the corneal surface may also be one of the reasons for the poor performance of young doctors, who may mistakenly believe that this is a region of corneal epithelial defect. Second, young ophthalmologists lack rich clinical experience and may have poor judgment in Grades 1 and 2 of corneal staining, but machine learning has learned the knowledge of experts. Finally, the interpretation of corneal staining results depends heavily on the doctor’s individual subjective feelings. At the same time, the machine’s standard is fixed and will not be affected by multiple factors, such as working environment and mental state.

This study also has limitations. Punctate staining areas may be challenging to detect and count. Kourukmas and Rodriguez analyzed punctate

corneal lesion staining regions using digital image analysis techniques,^{50,51} but such methods are suitable for a small amount of data, and their ways perform poorly in the era of massive data. Therefore, the future research direction is implementing the accurate quantification of fine punctate staining, which is still very challenging work. In addition, it is costly and time-consuming for doctors to label images, so researchers need to consider unsupervised machine learning methods to process the big data generated in outpatient clinics in future work. Finally, in deep learning for image segmentation and classification, one image is regarded as a sample, and the gap between the image and the target region is viewed as a loss to leverage backpropagation to update the neural network parameters and carry out network optimization. Power analysis is not used in this application scenario. It is well-known that the cost of image labeling by physicians is prohibitive. Furthermore, according to expert experience, in the sample collection labeling stage, as far as possible to be consistent with the distribution of actual samples, and to ensure the diversity and complexity of samples, each sample has a certain independence and representativeness. In a word, incremental learning was used in our study; that is to say, when the sample size we collected is enough to allow the neural network model to learn the characteristics and perform well, we do not collect data.

Conclusion

Our proposed method performs state-of-the-art vision tasks, including identifying corneal regions, detecting patchy corneal epithelial defects, and automatically assessing corneal staining scores (Figure 5). Moreover, this method is a fast, simple and trustworthy way. Besides, our proposed deep neural network framework has better generalization ability. The development of such a system offers the possibility of future applications in practical clinical scenarios.

Declarations

Ethics approval and consent to participate

Considering that all of the data are anonymous in our clinical practice and patient information is not subject to leakage in data acquisition, data analysis, and paper writing, we adopted the simple form of verbal informed consent in this study

plan. The detailed method is that before collecting the corneal fluorescein sodium photos of the patients, we will show the patients the documents approved by the Ethics Committee of Xiamen University School of Medicine, inform the patients of the detailed research plan, make clear that the patients understand the purpose of this study, research content, image resource usage, and then obtain the images after the patients' consent. The plan was conducted according to the World Medical Association Declaration of Helsinki, and the protocol has been approved by the Ethics Committee of Xiamen University School of Medicine with the approval number XDYX2022004.

Consent for publication

Not applicable.

Author contributions

Shaopan Wang: Conceptualization; Methodology; Writing – original draft.

Jiezhou He: Conceptualization; Methodology; Writing – original draft.

Xin He: Conceptualization; Methodology; Writing – original draft.

Yuwen Liu: Conceptualization; Methodology; Writing – original draft.

Xiang Lin: Methodology; Writing – original draft.

Changsheng Xu: Methodology; Writing – original draft.

Linfangzi Zhu: Methodology; Writing – original draft.

Jie Kang: Methodology; Writing – original draft.

Yuqian Wang: Methodology; Writing – original draft.

Yong Li: Methodology; Writing – original draft.

Shujia Guo: Methodology; Writing – original draft.

Yunuo Zhang: Methodology; Writing – original draft.

Zhiming Luo: Conceptualization; Writing – original draft; Writing – review & editing.

Zuguo Liu: Conceptualization; Writing – original draft; Writing – review & editing.

Acknowledgements

None.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the National Key R&D Program of China (No. 2018YFA0107304) and the National Natural Science Foundation of China (NSFC, No. 81870627). The funders have no role in the study design, data collection and analysis, decision on publishing, or manuscript preparation.

Competing interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Availability of data and materials

Data and related materials used to support this study are available from corresponding authors upon request.

ORCID iD

Shaopan Wang  <https://orcid.org/0000-0001-8044-8398>

References

- McMahon TT and Robin JB. Corneal trauma: I – classification and management. *J Am Optom Assoc* 1991; 62: 170–178.
- Niu L, Li L, Xing C, *et al.* Airborne particulate matter (PM(2.5)) triggers cornea inflammation and pyroptosis via NLRP3 activation. *Ecotoxicol Environ Saf* 2021; 207: 111306.
- Lobo AM, Agelidis AM and Shukla D. Pathogenesis of herpes simplex keratitis: the host cell response and ocular surface sequelae to infection and inflammation. *Ocul Surf* 2019; 17: 40–49.
- Jin M, Wang Y, An X, *et al.* Phenotypic and transcriptomic changes in the corneal epithelium following exposure to cigarette smoke. *Environ Pollut* 2021; 287: 117540.
- Raizman MB, Hamrah P, Holland EJ, *et al.* Drug-induced corneal epithelial changes. *Surv Ophthalmol* 2017; 62: 286–301.
- Ardan T, Němcová L, Bohuslavová B, *et al.* Reduced levels of tissue inhibitors of metalloproteinases in UVB-irradiated corneal epithelium. *Photochem Photobiol* 2016; 92: 720–727.
- Morgan PB and Maldonado-Codina C. Corneal staining: do we really understand what we are seeing? *Cont Lens Anterior Eye* 2009; 32: 48–54.
- American Academy of Ophthalmology (AAO). External disease and cornea. AAO, San Francisco, CA, 2010.
- Krachmer J, Mannis M and Holland E. *Cornea fundamentals, diagnosis and management*. 3rd ed. St Louis, MO: Mosby, 2010.
- Pritchard N, Young G, Coleman S, *et al.* Subjective and objective measures of corneal staining related to multipurpose care systems. *Cont Lens Anterior Eye* 2003; 26: 3–9.
- Chun YS, Yoon WB, Kim KG, *et al.* Objective assessment of corneal staining using digital image analysis. *Invest Ophthalmol Vis Sci* 2014; 55: 7896–7903.
- Deng L, Huang H, Yuan J, *et al.* Automatic segmentation of corneal ulcer area based on ocular staining images. In: *Medical imaging 2018: biomedical applications in molecular, structural, and functional imaging*, 2018, p.105781D. International Society for Optics and Photonics, <https://ui.adsabs.harvard.edu/abs/2018SPIE10578E..1DD/abstract>
- Liu Z, Shi Y, Zhan P, *et al.* Automatic corneal ulcer segmentation combining Gaussian mixture modeling and Otsu method. *Annu Int Conf IEEE Eng Med Biol Soc* 2019; 2019: 6298–6301.
- Lin T-Y, Maire M, Belongie S, *et al.* Microsoft COCO: common objects in context. In: *European conference on computer vision*, 2014, pp. 740–755. Springer, https://link.springer.com/chapter/10.1007/978-3-319-10602-1_48
- He K, Gkioxari G, Dollár P, *et al.* Mask R-CNN. In: *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 2961–2969, <https://ieeexplore.ieee.org/document/8237584>
- Girshick R. Fast R-CNN. In: *Proceedings of the IEEE international conference on computer vision*, 2015, pp. 1440–1448, <https://ieeexplore.ieee.org/document/7410526>
- Everingham M, Van Gool L, Williams CK, *et al.* The Pascal visual object classes (VOC) challenge. *Int J Comput Vis* 2010; 88: 303–338.
- De Brebisson A and Montana G. Deep neural networks for anatomical brain segmentation. In: *Proceedings of the IEEE conference on computer*

- vision and pattern recognition workshops, 2015, pp. 20–28, <https://ieeexplore.ieee.org/document/7301312/>
19. Wachinger C, Reuter M and Klein T. DeepNAT: deep convolutional neural network for segmenting neuroanatomy. *NeuroImage* 2018; 170: 434–445.
 20. Srivastava R, Cheng J, Wong DW, *et al.* Using deep learning for robustness to parapapillary atrophy in optic disc segmentation. In: 2015 IEEE 12th international symposium on biomedical imaging (ISBI), 2015, pp. 768–771. IEEE, <https://ieeexplore.ieee.org/document/7163985>
 21. Zhang H, Niu K, Xiong Y, *et al.* Automatic cataract grading methods based on deep learning. *Comput Methods Programs Biomed* 2019; 182: 104978.
 22. Stember JN, Celik H, Krupinski E, *et al.* Eye tracking for deep learning segmentation using convolutional neural networks. *J Digit Imaging* 2019; 32: 597–604.
 23. 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.
 24. Li Y-H, Huang P-J and Juan Y. An efficient and robust iris segmentation algorithm using deep learning. *Mob Inf Syst* 2019; 2019: 4568929.
 25. Peters M. *BMA A-Z family medical encyclopedia*. London: Dorling Kindersley, 2004.
 26. Foster A and Sommer A. Corneal ulceration, measles, and childhood blindness in Tanzania. *Br J Ophthalmol* 1987; 71: 331–343.
 27. Liu Z, Lin Y, Cao Y, *et al.* Swin transformer: hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF international conference on computer vision, 2021, pp. 10012–10022, <https://arxiv.org/abs/2103.14030>
 28. Dosovitskiy A, Beyer L, Kolesnikov A, *et al.* An image is worth 16x16 words: transformers for image recognition at scale. arXiv preprint arXiv:2010.11929, 2020, <https://arxiv.org/abs/2010.11929>
 29. Lemp MA. Report of the National Eye Institute/ Industry workshop on clinical trials in dry eyes. *CLAO J* 1995; 21: 221–232.
 30. Andrews S and Hamarneh G. *Multi-region probabilistic dice similarity coefficient using the Aitchison distance and bipartite graph matching*. arXiv preprint arXiv:1509.07244, 2015, <https://arxiv.org/abs/1509.07244>
 31. Huttenlocher DP, Klanderman GA and Rucklidge WJ. Comparing images using the Hausdorff distance. *IEEE Trans Pattern Anal Mach Intell* 1993; 15: 850–863.
 32. Yacouby R and Axman D. Probabilistic extension of precision, recall, and F1 score for more thorough evaluation of classification models. In: Proceedings of the first workshop on evaluation and comparison of NLP systems, 2020, pp. 79–91, <https://aclanthology.org/2020.eval4nlp-1.9/>
 33. Tharwat A. Classification assessment methods. *Appl Comput Inform* 2021; 17: 168–192.
 34. Ronneberger O, Fischer P and Brox T. U-net: convolutional networks for biomedical image segmentation. In: International conference on medical image computing and computer-assisted intervention, 2015, pp. 234–241. Springer, https://link.springer.com/chapter/10.1007/978-3-319-24574-4_28
 35. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, *et al.* UNet++: a nested u-net architecture for medical image segmentation. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3–11, https://link.springer.com/chapter/10.1007/978-3-030-00889-5_1
 36. Cao H, Wang Y, Chen J, *et al.* Swin-Unet: Unet-like pure transformer for medical image segmentation. arXiv preprint arXiv:2105.05537, 2021, <https://arxiv.org/abs/2105.05537>
 37. Chen J, Lu Y, Yu Q, *et al.* TransUnet: transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306, 2021, <https://arxiv.org/abs/2102.04306>
 38. Alom MZ, Yakopcic C, Nasrin MS, *et al.* Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network. *J Digit Imaging* 2019; 32: 605–617.
 39. Zhang X, Zhou X, Lin M, *et al.* ShuffleNet: an extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 6848–6856, https://openaccess.thecvf.com/content_cvpr_2018/papers/Zhang_ShuffleNet_An_Extremely_CVPR_2018_paper.pdf
 40. Kassani SH, Kassani PH, Khazaeinezhad R, *et al.* Diabetic retinopathy classification using a modified xception architecture. In: 2019 IEEE international symposium on signal processing

- and information technology (ISSPIT), 2019, pp. 1–6. IEEE, <https://ieeexplore.ieee.org/document/9001846>
41. Zhaputri A, Hayaty M and Laksito AD. Classification of brain tumour MRI images using efficient network. In: 2021 4th international conference on information and communications technology (ICOIACT), 2021, pp. 108–113. IEEE, <https://ieeexplore.ieee.org/document/9563922>
 42. Zhou Y, Li Z, Zhu H, *et al.* Holistic brain tumor screening and classification based on densenet and recurrent neural network. In: International MICCAI brainlesion workshop, 2018, pp. 208–217. Springer, https://link.springer.com/chapter/10.1007/978-3-030-11723-8_21
 43. Lu Z, Bai Y, Chen Y, *et al.* The classification of gliomas based on a Pyramid dilated convolution resnet model. *Pattern Recognit Lett* 2020; 133: 173–179.
 44. Gheflati B and Rivaz H. *Vision transformer for classification of breast ultrasound images*. arXiv preprint arXiv:2110.14731, 2021, <https://arxiv.org/abs/2110.14731>
 45. Joyce PD. Corneal vital staining. *Ir J Med Sci* 1967; 42: 359–367.
 46. Khaki A, Aghagolzadeh A and Cami BR. ISUR: Iris segmentation based on UNet and ResNet. In: 2021 11th international conference on computer engineering and knowledge (ICCKE), 2021, pp. 1–7. IEEE, <https://ieeexplore.ieee.org/document/9721475/>
 47. Gowroju S and Kumar S. Robust pupil segmentation using UNET and morphological image processing. In: 2021 international mobile, intelligent, and ubiquitous computing conference (MIUCC), 2021, pp. 105–109. IEEE, <https://ieeexplore.ieee.org/document/9447658>
 48. Wang Z, Xie Z and Xu Y. Automatic segmentation for retinal vessel using concatenate UNet++. In: International conference on machine learning and big data analytics for IoT security and privacy, 2021, pp. 10–18. Springer, https://link.springer.com/chapter/10.1007/978-3-030-89508-2_2
 49. Chen J, Chen J, Zhou Z, *et al.* *MT-TransUnet: mediating multi-task tokens in transformers for skin lesion segmentation and classification*. arXiv preprint arXiv:2112.01767, 2021, <https://arxiv.org/abs/2112.01767>
 50. Kourukmas R, Roth M, Geerling G. and Automated vs. Human evaluation of corneal staining. *Graefes Arch Clin Exp Ophthalmol* 2022; 260: 2605–2612.
 51. Rodriguez JD, Lane KJ, Ousler GW, *et al.* Automated grading system for evaluation of superficial punctate keratitis associated with dry eye. *Invest Ophthalmol Vis Sci* 2015; 56: 2340–2347.