RESEARCH

Open Access



Automatic diagnosis of keratitis using object localization combined with cost-sensitive deep attention convolutional neural network

Jiewei Jiang¹⁺, Wei Liu²⁺, Mengjie Pei³, Liufei Guo¹, Jingshi Yang¹, Chengchao Wu², Jiaojiao Lu², Ruijie Gao¹, Wei Chen^{4,5}, Jiamin Gong^{1*}, Mingmin Zhu^{6*} and Zhongwen Li^{4,5*}

[†]Jiewei Jiang and Wei Liu contributed equally as first authors.

*Correspondence: gjm@xupt.edu.cn; zmmzhu2010@126.com; li.zhw@qq.com

¹ School of Electronics Engineering, Xi'an University of Posts and Telecommunications, Xi'an 710121. China ² School of Communications and Information Engineering. Xi'an University of Posts and Telecommunications, Xi'an 710121, China ³ School of Computer Science and Technology, Xi'an University of Posts and Telecommunications. Xi'an 710121, China ⁴ Ningbo Eye Hospital, Wenzhou Medical University, Ningbo 315000, China ⁵ School of Ophthalmology and Optometry and Eye Hospital, Wenzhou Medical University, Wenzhou 325027, China ⁶ School of Mathematics and Statistics, Xidian University, Xi'an 710071, China

Abstract

Keratitis is a major cause of corneal blindness worldwide. Early identification and timely treatment of keratitis can deter the disease progression, reaching a better prognosis. The diagnosis of keratitis often requires professional ophthalmologists. However, ophthalmologists are relatively scarce and unevenly distributed, especially in underserved and remote regions, making the early diagnosis of keratitis challenging. In this study, an object localization method combined with cost-sensitive deep attention convolutional neural network (OL-CDACNN) was proposed for the automated diagnosis of keratitis. First, the single shot multibox detector (SSD) algorithm was employed to automatically locate the region of conjunctiva and cornea (Conj Cor) on the original slit-lamp image. Then, the region of Conj Cor was classified using a cost-sensitive deep attention convolutional network (CDACNN) to identify keratitis, other cornea abnormalities, and normal cornea. A total of 12,407 slit-lamp images collected from four clinical institutions were used to develop and evaluate the OL-CDACNN. For detecting keratitis, other cornea abnormalities, and normal cornea, the OL-CDACNN model achieved area under the receiver operating characteristic curves (AUCs) of 0.998, 0.997, and 1.000, respectively, in an internal test dataset. The comparable performance (AUCs ranged from 0.981 to 0.998) was observed in three external test datasets, further verifying its effectiveness and generalizability. Due to reliable performance, our model has a high potential to provide an accurate diagnosis and prompt referral for a patient with keratitis in an automated fashion.

Keywords: Automatic diagnosis of keratitis, Object localization, Cost-sensitive, Deep attention convolutional neural network, Slit-lamp images

Introduction

Corneal blindness that mainly results from keratitis is the fifth leading cause of blindness worldwide, affecting more than 4.2 million people [1, 2]. Compared to other blindness reasons such as age-related macular degeneration and glaucoma, corneal blindness has higher incidence in a relatively young population and thereby putting a heavier burden on both the patients and society [3, 4]. Notably, visual impairment due to keratitis can be avoidable through early diagnosis and appropriate treatment [5, 6]. Otherwise, keratitis



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http:// creativecommons.org/licenses/by/4.0/.

can get worse rapidly, potentially leading to corneal perforation and even permanent vision loss [7, 8].

The diagnosis of keratitis is usually performed by professional ophthalmologists based on the examination of the morphology and color information of the cornea using a slit-lamp microscope [9–11]. However, experienced ophthalmologists are scarce and unevenly distributed, especially in underserved and remote regions, where ophthalmologists cannot meet the increasing demand for ophthalmic patients [12, 13]. In addition, ophthalmological diagnosis is a time-consuming and labor-intensive process [11, 14]. To cross the chasm of manual diagnosis defects in ophthalmology, it is imperative to develop an automatic diagnosis algorithm for the accurate identification of keratitis.

Recently, artificial intelligence (AI) hold tremendous promise for automatically diagnosing ophthalmic diseases such as diabetic retinopathy [15, 16], glaucoma [17, 18], cataract [19, 20], age-related macular degeneration [21, 22], and eyelid tumors [23]. In the field of diagnosing keratitis, several studies have demonstrated the effectiveness of deep learning techniques [24–29]. Kuo et al. [24] utilized DenseNet [30] to distinguish fungal keratitis (FK), reporting an area under the curve (AUC) of 0.65 based on 288 corneal photographs. Similarly, Gu et al. [25] achieved exceptional performance in diagnosing keratitis using Inception-v3 [31] with an AUC of 0.93 on a dataset of 5,325 ocular surface slit-lamp images. Redd et al. [26] further investigated different CNNs to differentiate between bacterial keratitis (BK) and FK, resulting in an improved AUC of 0.86. Ghosh et al. [27] employed CNN and ensemble learning techniques for diagnosing both BK and FK, achieving an impressive AUC of 0.90 with the analysis of 223 slit-lamp images. In addition, Hung et al. [28] also developed a deep learning model based on DenseNet161 [30], which effectively discriminated BK and FK, with an AUC of 0.85. Tiwari et al. [29] employed VGG16 [32] for the automated diagnosis of corneal infections and healed scars, attaining remarkable AUCs of 0.947 and 0.973, respectively.

Although the aforementioned studies demonstrate the potential of advanced deep learning techniques in the application of keratitis diagnosis, it is important to acknowledge that the complicated characteristics of the keratitis lesions and large amounts of noise (e.g., evelids) in the slit-lamp images often lead to relatively low performance in the existing systems. For example, the clinical phenotype of keratitis occurs not only in the region of cornea but also in the region of conjunctiva, such as conjunctival injection (Fig. 1a) [33]. Therefore, keratitis is sometimes misdiagnosed as conjunctivitis, leading to missing the optimal time for treatment initiation. Besides, in slit-lamp images, noises such as eyelashes and eyelids are distributed around the cornea. If the original slit-lamp image is directly input into deep learning, it will inevitably extract noise features which affecting the performance of the classifier. It is necessary to study the object localization algorithm to eliminate the noise and prevent it from being transmitted to the classifier. Moreover, early-stage keratitis often presents atypical lesions on slit-lamp images. Indistinguishable lesion characteristics are present between keratitis and other corneal abnormalities (Fig. 1b), posing a high challenge for the accurate diagnosis of keratitis in an automated fashion.

To address the above-mentioned issues, in this study, we proposed an object localization method combined with cost-sensitive deep attention convolutional neural network (OL-CDACNN) for automatic diagnosis of keratitis. First, the single shot multibox



a.Keratitis

b.Cornea with other abnormities al slit-lamp images, **a**. images of keratitis, **b**. images of other corne

Fig. 1 Several representative corneal slit-lamp images. a. images of keratitis. b. images of other cornea abnormalities. c. images of normal cornea

detector (SSD) [34] was employed to filter out the noises such as eyelashes and eyelids around the cornea, and the single region of cornea and the region of conjunctiva and cornea (Conj_Cor) were automatically located and cropped, respectively. Second, the cost-sensitive deep attention convolutional network (CDACNN) was proposed to enable the classification of keratitis, other cornea abnormalities, and normal cornea. Third, we explored and compared the impact of different localization strategies (the regions of cornea and Conj_Cor) on the performance of the OL-CDACNN for identifying keratitis, other cornea abnormalities, and normal cornea. Furthermore, the developed automated diagnosis model can be integrated into slit-lamp cameras to facilitate early detection of keratitis in resource-limited settings where ophthalmologists are scarce, enabling timely referral for positive cases and preventing the occurrence of corneal blindness caused by keratitis.

Methods

Datasets

In this study, 6567 slit-lamp images (2584×2000 pixels in JPG format) collected from Ningbo Eye Hospital (NEH) between January 2017 and March 2020 were used to develop OL-CDACNN. Additional datasets including 5,840 slit-lamp images derived from three clinical institutions were utilized to externally evaluate the performance of OL-CDACNN. The first one was derived from the outpatient clinics and health screening center at Jiangdong Eye Hospital (JEH), consisting of 1987 images (5784×3456 pixels in JPG format); the second one was collected from the outpatient clinics and inpatient department at Ningbo Ophthalmic Center (NOC), consisting of 2924 images (1740×1534 pixels in PNG format); the last one was obtained from the outpatient

clinics, inpatient department, and dry eye center at Zhejiang Eye Hospital (ZEH), consisting of 929 images (2592×1728 pixels in JPG format).

The diagnosis of each image was determined by three experienced ophthalmologists based on clinical manifestations, ocular examination (such as fluorescein staining of the cornea and corneal confocal microscopy), laboratory tests (such as corneal scraping smear examination and the culture of corneal sample), and follow-up visits. All images with clear diagnosis were classified into three categories: keratitis, other cornea abnormalities, and normal cornea. Two regions of cornea and Conj_Cor were automatically localized and copped on each slit-lamp image to fairly compare the effectiveness of different localization strategies. The slit-lamp images derived from the NEH dataset were randomly divided (70%:15%:15%) into training (4526), validation (1,055), and internal test datasets (986). To prevent leakage and biased assessment of the OL-CDACNN, images from the same individual were assigned to only one same dataset. The training and validation datasets were employed to develop the OL-CDACNN and the internal test dataset was used to evaluate its performance. Detailed information on the datasets NEH, JEH, NOC, and ZEH is summarized in Table 1.

Ethical approval

The study was approved by the Institution Review Board of NEH (identifier, 2020-qtky-017) and adhered to the principles of the Declaration of Helsinki. All anonymous slitlamp images were transferred to research investigators for inclusion. Informed consent was exempted, due to the retrospective nature of the data acquisition and the use of deidentified images.

Overall framework of OL-CDACNN

As shown in Fig. 2, the framework of OL-CDACNN for keratitis diagnosis consists of two stages: automatic localization for the regions of cornea and Conj_Cor (Fig. 2a) and automatic classification of keratitis, other cornea abnormalities, and normal cornea (Fig. 2b). Except for the lesion region of cornea, other keratitis-associated signs can also present in the region of conjunctiva, such as conjunctival injection. Therefore, both the cornea and conjunctiva are regions of interest for the diagnosis of keratitis. In the first stage, the SSD was employed to locate and crop the regions of cornea and Conj_Cor [34]. In the second stage, the two cropped regions were separately input into CDACNN to achieve the classification of keratitis, other cornea abnormalities, and normal cornea. We compared the impact of the two cropped regions on the performance of OL-CDACNN in detail to determine the optimal localization strategy. Patients detected as

Table 1 Distributions of slit-lamp images in four clinical institutions

	NEH			JEH	NOC	ZEH
	Training	Validation	Test	Test	Test	Test
Keratitis	2185	511	483	1186	839	378
Other cornea abnormalities	585	136	130	236	969	237
Normal cornea	1756	408	373	565	1116	314

NEH Ningbo Eye Hospital, JEH Jiangdong Eye Hospital, NOC Ningbo Ophthalmic Center, ZEH Zhejiang Eye Hospital



Fig. 2 The overall framework of automatic diagnosis of keratitis. **a**. Automatic localization for the regions of cornea and Conj_Cor using SSD. **b**. Automatic diagnosis of keratitis, other cornea abnormalities, and normal cornea using CDACNN. *SSD* single shot multibox detector, *CDCNN* cost-sensitive deep attention convolutional network, *Conj_Cor* conjunctiva and cornea, *DA* deep attention

keratitis or other cornea abnormalities are referred to experienced ophthalmologists for further confirmation and treatment.

Automatic localization of the regions of cornea and Conj_Cor

To obtain the optimal localization method, eight object localization algorithms are examined and compared in this study: five two-stage object localization algorithms and three one-stage object localization algorithms. The two-stage object localization algorithms include Faster R-CNN1 [35], Faster R-CNN2, Cascade R-CNN1 [36], Cascade R-CNN2, and TridentNet [37], where the feature extraction networks used in R-CNN1 and R-CNN2 are ResNet50 [38] and ResNet101, respectively, and the feature extraction network used for TridentNet is ResNet50. The one-stage object localization algorithms include RetinaNet1 [39], RetinaNet2, and SSD [34], where the feature extraction networks in RetinaNet1 and RetinaNet2 are ResNet50 and ResNet101, respectively, and the feature extraction networks in RetinaNet1 and RetinaNet2 are ResNet50 and ResNet101, respectively, and the feature extraction networks in RetinaNet1 and RetinaNet2 are ResNet50 and ResNet101, respectively, and the feature extraction networks of SSD is VGG16 [32].

The two-stage detectors first generate a series of candidate frames using the candidate region generation module, then perform classification and regression based on the candidate frames to determine the exact region and category. The one-stage detectors perform classification and regression directly based on the anchor frame and complete the adjustment of the bounding box and category identification in one step.

Automatic diagnosis of keratitis

Compared with the Residual Convolutional Network (ResNet), the Dense Convolutional Network (DenseNet) [30] achieves extraordinary performance with fewer computations and more effectiveness. In this study, cost-sensitive and deep attention mechanisms are applied in the DenseNet to further improve the performance and generalization ability of automatic diagnosis of keratitis. Specifically, a deep attention (DA) module in Fig. 2b is unfolded to analyze its internal structure and implementation principle, as shown in Fig. 3. After 3×3 convolution operation, a channel attention sub-module and a spatial



Fig. 3 An unfolded form of one deep attention module

attention sub-module [40] were adopted to adjust the CDACNN to focus on the lesion of keratitis, enhancing the expression of lesion-related features and suppressing the expression of noise features around the lesion. Then, the feature maps of the attention module and the input module are concatenated and fed into the next attention module. After cascading multiple attentions, the deep attention convolutional network is constructed. Furthermore, the cost-sensitive is incorporated and optimized in the loss function to determine appropriate parameters, facilitating the CDACNN to focus more on the minority category of other cornea abnormalities.

Channel attention mechanism

The feature map of channel attention is generated by exploiting the inter-channel relationship of different features. As each channel for a feature map is considered as a feature detector, channel attention focuses on 'what' is meaningful given an input image. To compute the channel attention efficiently, we squeeze the spatial dimension of the input feature map. Specifically, the input feature map is performed using both average-pooling and max-pooling operations in the channel dimension to aggregate two different spatial context descriptors: **Fc avg** and **Fc max**. The aggregated descriptors are forwarded to a multi-layer perceptron (MLP), then the output of the MLP is summed and activated by the sigmoid to generate the feature map of channel attention: M_c . The mathematical expression for channel attention is shown in Eq. (1).

$$\mathbf{M}_{\mathbf{c}}(\mathbf{F}) = \sigma \left(MLP(AvgPool(\mathbf{F})) + MLP(MaxPool(\mathbf{F}))\right)$$

= $\sigma \left(\mathbf{W}_{\mathbf{1}} \left(\mathbf{W}_{\mathbf{0}} \left(\mathbf{F}_{avg}^{c} \right) \right) + \mathbf{W}_{\mathbf{1}} \left(\mathbf{W}_{\mathbf{0}} \left(\mathbf{F}_{max}^{c} \right) \right) \right)$ (1)

where σ , *AvgPool*, *MaxPool*, **F**, **Fc avg**, and **Fc max** denote the sigmoid function, averagepooling operation, max-pooling operation, input feature map, average-pooled features, and max-pooled features, respectively. **W**₀ and **W**₁ are the weights of MLP.

Spatial attention mechanism

The feature map of spatial attention is generated by utilizing the inter-spatial relationship of different features. Compared with channel attention, spatial attention focuses on 'where' is an informative part, which is complementary to channel attention. To compute the spatial attention, average-pooling and max-pooling operations along the channel axis are applied to aggregate channel information to generate two 2D maps (**Fs avg** and **Fs max**) which denote average-pooled features and max-pooled features. Then, the two 2D maps are concatenated and convolved by a standard 7×7 convolution layer to produce the feature map of spatial attention. The mathematical expression for spatial attention is shown in Eq. (2).

$$\mathbf{M}_{\mathbf{s}}(\mathbf{F}) = \sigma \left(f^{7 \times 7} ([AvgPool(\mathbf{F}); MaxPool(\mathbf{F})]) \right) = \sigma \left(f^{7 \times 7} \left(\left[\mathbf{F}_{avg}^{s}; \mathbf{F}_{max}^{s} \right] \right) \right)$$
(2)

where σ , $f^{7\times7}$, *AvgPool*, *MaxPool*, **F**, **Fs avg**, and **Fs max** denote the sigmoid function, a convolution operation with the filter size of 7×7 , average pooling, maximum pooling, feature map of the input, average-pooled features, and max-pooled features, respectively.

Cost-sensitive method and optimization process

In the NEH dataset, the number of keratitis patients (or normal cornea) is far more than that of other corneal abnormalities patients. The class imbalance problem is common in medical practice. The imbalanced dataset causes the decision boundary of conventional classifiers to be biased towards the majority class [19, 20, 41]. Therefore, this study employs the cost-sensitive method to adjust the weights of different classes in the softmax loss function. Specifically, this study discriminatively determines the cost of misclassification of different classes and assigns a larger weight to the other cornea abnormalities class. For one iterative training stage, *m* samples are selected at random to form a training dataset {[$x^{(1)}$, $y^{(1)}$], [$x^{(2)}$, $y^{(2)}$],..., [$x^{(m)}$, $y^{(m)}$]}, where $x^{(i) \in \mathbb{R}^{(l)}}$ and $y^{(i)} \in \{1,...,k\}$. Here, $x^{(i)}$ denotes the features of the *i*-th sample, and $y^{(i)}$ is the class label. The cost-sensitive loss function can be computed as shown in Eq. (3).

$$L(\theta) = -\frac{1}{m} \left[\sum_{i=1}^{m} \sum_{j=1}^{k} I\left\{ y^{(i)} = j \right\} \times CS\left\{ y^{(i)} = other \ abnormalities \right\} \times \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{s=1}^{k} e^{\theta_s^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^{k} \sum_{j=1}^{n} \theta_{ij}^2$$

$$(3)$$

where *m*, *n*, *k*, and θ denote the number of training samples, the number of input neurons of the softmax function, the number of classes, and trainable parameters, respectively. $I\{y^{(i)}=j\}$ denotes the indicator function $(I\{y^{(i)} is equal to j\}=1 \text{ and } I\{y^{(i)} is not equal to j\}=0)$ while $CS\{y^{(i)}=other abnormalities\}$ is the cost-sensitive weight function $(CS\{y^{(i)} is the other abnormalities class label}\}=C$ and $CS\{y^{(i)} is not the other abnormalities class label}=1)$. In this study, the grid search method is employed to determine the effective cost-sensitive weight parameter *C* within the interval [2–4]. Moreover, a weight decay term $\frac{\lambda}{2} \sum_{i=1}^{k} \sum_{j=1}^{n} \theta_{ij}^2$ is used to penalize the larger trainable weights. Finally, to obtain the optimal parameters θ^* (see Eq. (4)), the mini-batch gradient descent (Mini-batch-GD) is employed to minimize $L(\theta)$ as shown in Eq. (5).

$$\theta^* = \arg\min_{\theta} L(\theta) \tag{4}$$

$$\nabla_{\theta_j} L(\theta) = -\frac{1}{m} \sum_{i=1}^m \left[CS \left\{ y^{(i)} = other \ abnormalities \right\} \times x^{(i)} \\ \times \left(I \left\{ y^{(i)} = j \right\} - p(y^{(i)} = j | x^{(i)}; \theta) \right) \right] + \lambda \theta_j$$
(5)

Results

Experimental environment

This study was conducted using the PyTorch deep learning framework (version 1.7.1) [42], and all models were trained in parallel with four NVIDIA TITAN RTX GPUs. For object localization algorithms, the number of iterations (epochs) was set to 24. For classifier algorithms, a mini-batch size of 32 was utilized on each GPU, resulting in the processing of 128 images in one iteration. The adaptive moment estimation (ADAM) optimizer was employed to optimize the model's parameters. The specific parameters were configured with an initial learning rate of 0.001, β 1 of 0.9, β 2 of 0.999, and a weight decay of 1e-4. The maximum number of iterations for training classifiers was set to 80. During the training process, the loss value on the validation dataset was used as a metric for model selection. After each iteration, the loss value was calculated based on the validation dataset. Finally, the model with the lowest loss value was selected as the optimal model for testing on the test dataset. To increase the diversity of the NEH dataset and prevent overfitting and bias problems during training, data augmentation techniques, including random cropping, random rotations around the image center, and horizontal and vertical flips, were adopted to enlarge the original training dataset by 6 times (from 4526 to 27,156).

Evaluation Metrics and statistical analysis

Average precision (AP) for each class and mean average precision (mAP) for all classes were used as two primary metrics to evaluate the performance of object localization in the regions of cornea and Conj_Cor. AP and mAP are calculated using P (Precision) and R (Recall) indicators, as shown in Eqs. (6-9).

$$P = \frac{TP}{TP + FP} \tag{6}$$

$$R = \frac{TP}{TP + FN} \tag{7}$$

$$AP = \int_0^1 P(R)dR \tag{8}$$

$$mAP = \frac{1}{N_{class}} \int_0^1 P(R) dR \tag{9}$$

where TP (True Positive) is the number of samples in which the region of cornea (or Conj_Cor) is predicted correctly; FP (False Positives) indicates the number of samples in which the predicted label is the region of cornea (or Conj_Cor), but the true label

is not; FN (False Negatives) indicates the number of samples in which the true label is the region of cornea (or Conj_Cor), but the predicted label is not. FN indicator implies region of cornea (or Conj_Cor) is not detected by the model. The AP is a popular metric for evaluating the performance of object detectors by estimating the area under the curve of the precision and recall relationship for a specific class, and the mAP is the mean AP of all classes.

The evaluation metrics of classification models are accuracy, specificity, sensitivity, receiver operating curve (ROC), and area under the ROC (AUC), as shown in Eqs. (10–12). A larger area under the ROC denotes better performance.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$
(10)

$$Specificity = \frac{TN}{TN + FP}$$
(11)

$$Sensitivity = \frac{TP}{TP + FN}$$
(12)

where, TP, FP, TN, and FN denote the numbers of true positives, false positives, true negatives, and false negatives in the classification results, respectively.

The performance of the object localization methods and classification models was evaluated by utilizing the one-versus-rest strategy. All statistical analyses were conducted using Python 3.7.8 and the package of Scikit-learn. The 95% confidence intervals (CI) for accuracy, specificity, and sensitivity were calculated with the Wilson Score Approach, and for AUC, using Empirical Bootstrap with 2000.

Performance comparison of different localization methods

To obtain the optimal localization method, this study compared the performance of eight object localization methods in the NEH internal test dataset and three external test datasets (JEH, NOC, and ZEH). The statistical results of AP and mAP for the localization of the regions of cornea and Conj_Cor are shown in Table 2. It is easy to obtain that the performance of one-stage localization methods for the region of cornea in the NEH test dataset was slightly better than those of two-stage localization methods, in which the SSD method achieved the best AP of 0.9898. It is worth mentioning that all eight methods obtained an AP of 1 for the region of Conj_Cor in the NEH test dataset. Overall, the SSD method achieved the best mAP of 0.9949.

Experimental results in three external test datasets (JEH, NOC, and ZEH) verified the generalization ability of the automatic localization methods. In the JEH dataset, the RetinaNet2 method achieved the best AP of 0.9987 for the region of cornea while the Faster R-CNN1 achieved the best AP of 0.9986 for the region of Conj_Cor. The RetinaNet2 achieved the best mAP of 0.9972. In the NOC dataset, the RetinaNet1 method obtained the best AP of 0.9983 for the region of cornea. All Faster R-CNN1, Faster R-CN2, TridentNet, RetinaNet1, and SSD methods accurately localized the region of Conj_Cor with an AP of 1. Therefore, the RetinaNet1 obtained the best mAP of 0.9992. In the ZEH dataset, both RetinaNet2 and SSD achieved the optimal AP of 0.9999 for the region

Methods		NEH			JEH			NOC			ZEH		
		Cornea	Conj_Cor	mAP									
Two-stage localization methods	Faster R-CNN1	0.9894	1.0000	0.9947	0.9889	0.9986	0.9938	0.9897	1.0000	0.9949	0.9897	1.0000	0.9948
	Faster R-CNN2	0.9894	1.0000	0.9947	0.9890	0.9900	0.9895	0.9791	1.0000	0.9895	0.9897	0.9994	0.9946
	Cascade R-CNN1	0.9894	1.0000	0.9947	0066.0	0.9898	0.9899	0.9899	0.9999	0.9949	0.9901	0.9901	0.9901
	Cascade R-CNN2	0.9894	1.0000	0.9947	0.9901	0.9898	0.9899	0.9898	0.9999	0.9948	0.9901	0.9901	0.9901
	TridentNet	0.9894	1.0000	0.9947	0.9599	0.9898	0.9748	0.9801	1.0000	0.9901	0.9801	1.0000	0.9901
One-stage localization methods	RetinaNet1	0.9894	1.0000	0.9947	0.9972	0.9970	0.9971	0.9983	1.0000	0.9992	0.9997	0.9993	0.9995
	RetinaNet2	0.9896	1.0000	0.9948	0.9987	0.9957	0.9972	0066:0	0.9998	0.9949	0.9999	0.9997	0.9998
	SSD	0.9898	1.0000	0.9949	0.9901	0.9901	0.9901	0.9901	1.0000	0.9951	0.9999	1.0000	1.0000

datasets
n four
nods or
n meth
izatior
t local
objec
eight
ice of
ormar
P perf
d mA
AP an
son of
nparis
Č C
ble

AP average precision, mAP mean average precision, Conj_Cor conjunctiva and cornea

of cornea. All Faster R-CNN1, TridentNet, and SSD accurately localized the region of Conj_Cor with an AP of 1. The SSD method obtained the best mAP of 1. Through the above analysis, it is not difficult to see that one-stage methods did not degrade the localization performance for the regions of cornea and Conj_Cor in the absence of the region proposal network (RPN) module. Conversely, the performance of one-stage methods was slightly better than or equivalent to that of two-stage methods. To visually observe the localization effect of the optimal method SSD, we presented several representative localization results for the regions of cornea and Conj_Cor, as shown in Fig. 4.

Efficiency analysis of different localization methods

To determine the best localization method, we further compared their efficiency and resource utilization, including the model size, trainable parameters, and running time of training and testing. As shown in Table 3, the number of trainable parameters, the model size, and the running time of testing of the SDD method were less than those of the other methods. Specifically, the SSD method only took 0.049 s to complete the localization of one slit-lamp image. In addition, although the training time of the SSD method was not minimal, the training could be performed on the local server in advance without affecting the efficiency of the model after deployment. Based on the above performance and efficiency analysis, the SSD method had higher diagnostic performance



a.Keratitis

abnormities

c.Normal cornea

Fig. 4 Representative localization examples for the regions of cornea and Conj_Cor. The green rectangle represents the predicted boundary by the SSD method and the blue rectangle represents the true boundary labeled by the experienced ophthalmologist. *Conj_Cor* conjunctiva and cornea, *SSD* single shot multibox detector

Methods	Size	Parameters(M)	Training time	Testing time
Faster R-CNN1	315.0 MB	4.1e+07	1.41 h	0.078 s
Faster R-CNN2	460.3 MB	6.0e+07	1.96 h	0.090 s
Cascade R-CNN1	527.1 MB	6.9e+07	1.86 h	0.083 s
Cascade R-CNN2	672.4 MB	8.7e+07	2.35 h	0.101 s
TridentNet	251.4 MB	3.3e+07	3.89 h	0.068 s
RetinaNet1	276.9 MB	3.6e+07	0.85 h	0.063 s
RetinaNet2	422.1 MB	5.5e+07	1.27 h	0.071 s
SSD	187.2 MB	2.5e+07	2.27 h	0.049 s

Гаb	e 3	Efficiency	comparison o	f eigl	ht ol	oject	loca	lization	meth	nod	S
-----	-----	------------	--------------	--------	-------	-------	------	----------	------	-----	---

Bold values represent the best performance among different methods in the same column. Training time indicates the running time of the method in the whole training process. Testing time indicates the average time that the method needs in testing every slit-lamp image. *MB* Mbyte



Fig. 5 Visualization of the separability for the embedding features learned by the CDACNN and DenseNet in the internal test dataset via t-SNE. Different colored point clouds represent the different categories. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. *Conj_Cor* conjunctiva and cornea, *t-SNE* t-distributed stochastic neighbor embedding, *CDACNN* cost-sensitive deep attention convolutional neural network

and faster detection speed, and was selected as the final detection model for the regions of cornea and Conj_Cor.

Effectiveness analysis of deep features of CDACNN and DenseNet

The t-distributed Stochastic Neighbor Embedding (t-SNE) [43] was utilized to visualize the embedding features of each category learned by the deep learning model in a two-dimensional space. In the internal test dataset, the t-SNE technique showed that the CDACNN had better capability to separate the embedding features of each category than that of DenseNet in the regions of cornea and Conj_Cor (Fig. 5). Notably, as shown in column 1 of Fig. 5, the features of the region of Conj_Cor extracted by CDACNN completely discriminated normal cornea from cornea abnormalities (including keratitis and other cornea abnormalities). In contrast, for the DenseNet method, some normal cornea samples were mixed with the keratitis samples, which were not easily discriminated, as shown by the dotted square in column 1 of Fig. 5. Therefore, the features of the region of Conj_Cor extracted by the CDACNN method obtained satisfactory separability.

Performance comparison of CDACNN and DenseNet in the internal test dataset

We trained CDACNN and DenseNet methods based on three types of images (original slit-lamp image, the region of cornea, and the region of Conj_Cor) to obtain six different combinations for the classification of keratitis, other cornea abnormalities, and normal cornea. To achieve an optimal diagnosis strategy, we compared the performance of six combinations in the internal test dataset, as shown in Table 4, Figs. 6 and 7. Analyzing the confusion matrixes in Fig. 6, compared to the original images and the region of cornea, both CDACNN and DenseNet achieved better performance based on the region of Conj_Cor. Furthermore, the CDACNN method outperformed the DenseNet method on the region of Conj_Cor, in which the CDACNN misclassified only 11 slit-lamp images while the DenseNet misclassifies 17 images. It is worth mentioning that the number of misclassified keratitis by CDACNN was reduced by half (from 14 to 7) relative to the DenseNet method on the region of Conj_Cor, as shown in the first column of Fig. 6. Correspondingly, the ROC curves (Fig. 7) of these six combinations demonstrated that

One-vsRest	CDACNN			DenseNet		
Classification	Sensitivity	Specificity	Accuracy	Sensitivity	Specificity	Accuracy (95%
	(95% Cl)	(95% CI)	(95% CI)	(95% Cl)	(95% CI)	Cl)
Keratitis vs. oth	er + normal					
Conj_Cor	98.6%	99.2%	98.9%	97.1%	99.8%	98.5%
	(0.975–0.996)	(0.984–1.000)	(0.982–0.995)	(0.956–0.986)	(0.994–1.002)	(0.977–0.992)
Cornea	98.3%	98.2%	98.3%	98.1%	96.2%	97.2%
	(0.972–0.995)	(0.971–0994)	(0.975–0.991)	(0.969–0.993)	(0.946–0.979)	(0.961–0.982)
Original	97.7%	97.2%	97.5%	97.9%	96.4%	97.2%
	(0.964–0.991)	(0.958–0.987)	(0.965–0.984)	(0.967–0.992)	(0.948–0.980)	(0.961–0982)
Other vs. kerati	tis + normal					
Conj_Cor	96.9%	99.2%	98.9%	99.2%	98.1%	98.3%
	(0.940–0.999)	(0.986–0.998)	(0.982–0.995)	(0.977–1.007)	(0.972–0.990)	(0.975–0.991)
Cornea	91.5%	99.1%	98.1%	88.5%	98.7%	97.4%
	(0.868–0.963)	(0.984–0.997)	(0.972–0.989)	(0.830–0.940)	(0.980–0.995)	(0.964–0.984)
Original	85.4%	98.6%	96.9%	84.6%	98.5%	96.7%
	(0.793–0.915)	(0.978–0.994)	(0.958–0.979)	(0.784–0.908)	(0.977–0.993)	(0.955–0978)
Normal vs. kera	titis + other					
Conj_Cor	100%	100%	100%	99.5%	100%	99.8%
	(1.000-1.000)	(1.000-1.000)	(1.000-1.000)	(0.987–1.002)	(1.000–1.000)	(0.995–1.001)
Cornea	100%	99.7%	99.8%	98.1%	99.8%	99.2%
	(1.000–1.000)	(0.992–1.001)	(0.995–1.001)	(0.967–0.995)	(0.995–1.002)	(0.986–0.997)
Original	99.5%	99.0%	99.2%	98.7%	99.3%	99.1%
	(0.987–1.002)	(0.982–0.998)	(0.986–0.997)	(0.975–0.998)	(0.987–1.000)	(0.985–0.997)

Table 4 Performance of the CDACNN and DenseNet in the internal test dataset

Normal normal cornea, Others cornea with other abnormalities, CI confidence interval, Conj_Cor conjunctiva and cornea, CDACNN cost-sensitive deep attention convolutional neural network



Fig. 6 Confusion matrices of CDACNN and DenseNet methods in the internal test dataset. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. *Conj_Cor* conjunctiva and cornea, *CDACNN* cost-sensitive deep attention convolutional neural network



Fig. 7 Receiver operating characteristic curves of CDACNN and DenseNet in the internal test dataset. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. *Conj_Cor* conjunctiva and cornea, *CDACNN* cost-sensitive deep attention convolutional neural network

the CDACNN had the best performance on the region of Conj_Cor for the classification of keratitis, other cornea abnormalities, and normal cornea.

Furthermore, the details on the performance of these six combinations in the internal test datasets are displayed in Table 4. Compared to the DenseNet method, the CDACNN method obtained the best performance on the region of Conj_Cor for discriminating keratitis, other cornea abnormalities, and normal cornea. On the region of Conj_Cor, the CDACNN method discriminated keratitis from normal cornea and cornea with other abnormalities with an AUC of 0.998 (95% confidence interval CI 0.996–1.000), a sensitivity of 98.6% (95% CI 0.975–0.996), and a specificity of 99.2% (95% CI 0.984–1.000). The CDACNN discriminated cornea with other abnormalities from keratitis and normal cornea with an AUC of 0.997 (95% CI 0.995–0.999), a sensitivity of 96.9% (95% CI 0.940–0.999), and a specificity of 99.2% (95% CI 0.986–0.998). The CDACNN discriminated normal cornea from abnormal cornea (including keratitis and other cornea abnormalities) with an AUC of 1 (95% CI 1.000–1.000), a sensitivity of 1 (95% CI 1.000–1.000), and a specificity of 1 (95% CI 1.000–1.000).

Performance comparison of CDACNN and DenseNet in the external test datasets

To explore the generalization ability of the CDACNN method, we performed and compared the performance of the CDACNN and DenseNet on three types of images (original images and two regions of cornea and Conj_Cor) in three external test datasets (JEH, NOC, and ZEH). The ROC curves of different combination models are shown in Fig. 8. Additional file 1: Figs. S1–S6 statistically showed the separability of features via the t-SNE technique and the confusion matrixes of CDACNN and DenseNet methods in three external test datasets. The details on the classification performance of the two methods in the external datasets, including accuracy, specificity, and sensitivity with 95% CI, are shown in Additional file 1: Tables S1–S3.

On the region of Conj_Cor in the JEH test dataset, the AUCs of CDACNN for classification of keratitis, other cornea abnormalities, and normal cornea were 0.996 (95% CI 0.994–0.997), 0.985 (95% CI 0.979–0.990), and 0.998 (95% CI 0.997–0.999), respectively, with accuracies of 96% (95% CI 0.952–0.969), 96.1% (95% CI 0.952–0.969), and 97.5% (95% CI 0.969–0.982), respectively. On the region of Conj_Cor in the NOC dataset, the AUCs of CDACNN for classification of keratitis, other cornea abnormalities, and normal cornea were 0.991 (95% CI 0.988–0.994), 0.981 (95% CI 0.976–0.985), and 0.988 (95% CI 0.985–0.991), respectively, with accuracies of 95.8% (95% CI 0.951–0.966), 93.9% (95% CI, 0.930–0.948), and 95.3% (95% CI 0.945–0.960), respectively. On the region of Conj_Cor in the ZEH dataset, the AUCs of CDACNN for classification of keratitis, other cornea abnormalities, and normal cornea were 0.993 (95% CI 0.989–0.996),



Fig. 8 Receiver operating characteristic curves of CDACNN and DenseNet in the external test datasets. **a**. Jiangdong Eye Hospital (JEH) dataset. **b**. Ningbo Ophthalmic Center (NOC) dataset. **c**. Zhejiang Eye Hospital (ZEH) dataset. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. *Conj_Cor* conjunctiva and cornea, *CDACNN* cost-sensitive deep attention convolutional neural network

0.994 (95% CI 0.991–0.997), and 0.995 (95% CI 0.992–0.997), respectively, with accuracies of 95.3% (95% CI 0.939–0.966), 97% (95% CI 0.959–0.981), and 96.8% (95% CI 0.956–0.979), respectively. Detailed performance comparisons on three external datasets further verified the excellent performance and strong generalization ability of the CDACNN method.

Interpretability analysis of CDACNN

To explore the reasonability of the CDACNN method, visual heatmaps were generated using the gradient-weighted class activation mapping (Grad-CAM) [44] technique for highlighting the disease-related regions on which the diagnosis model focused most. The Grad-CAM is an explainable technique for CDACNN, which leveraged the gradients of any target concepts flowing into the last layer of the CDACNN to generate a localization map highlighting remarkable regions in the image for predicting the concept. Redder regions represent more significant features of the CDACNN method. For keratitis and cornea with other abnormalities, heatmaps effectively highlighted the region of lesions. For normal cornea, the heatmap showed highlighted the region of the cornea. Typical examples of the heatmaps for keratitis, cornea with other abnormalities, and normal cornea are presented in Fig. 9. Using the Grad-CAM, we illustrated the rationale of the CDACNN method for discriminating keratitis, cornea with other abnormalities, and normal cornea.

Discussion

In this study, we proposed an object localization method combined with cost-sensitive deep attention convolutional neural network (OL-CDACNN) for discriminating keratitis, other cornea abnormalities, and normal cornea. The effectiveness and efficiency of eight object localization methods for two regions of cornea and Conj_Cor were investigated in detail to obtain a clinically applicable localization method SSD. The performance of six combinable strategies using two classifiers (CDACNN and DenseNet) and



Fig. 9 Representative heatmaps for keratitis, other cornea abnormalities, and normal cornea. The original images (left) and the corresponding heatmaps (right) for each category were shown in pairs

three different types of images (original slit-lamp image, the region of cornea, and the region of Conj_Cor) were explored for keratitis diagnosis. Qualitative and quantitative experiments demonstrated that the CDACNN based on the region of Conj_Cor outperformed the other combinable strategies. The internal test dataset and three external test datasets comprehensively verified the effectiveness and generalizability ability of the OL-CDACNN method in automatic localization and classification for keratitis. Moreover, the t-SNE and Grad-CAM techniques provided an interpretable path for the diagnosis of keratitis.

Recently, several studies for the automated diagnosis of keratitis have been published. The detailed comparison of these studies is shown in Table 5. Kuo et al. [24] provided a promising tool for identifying early FK at rural area based on slit-lamp images. Redd et al. [26] and Ghosh et al. [27] utilized various CNNs to differentiate between BK and FK using slit-lamp images. BK and FK are only a subset of keratitis so that these models may fail to identify other types of keratitis. Gu et al. [25] developed a hierarchical deep learning network with multi-task and multi-label classifiers for distinguishing infectious keratitis, non-infectious keratitis, corneal dystrophy or degeneration, and corneal neoplasm. However, the original slit-lamp images used in these existing studies contained noisy regions that were irrelevant to keratitis, such as eyelashes and eyelids. These noises inevitably introduced redundant features for keratitis, which could potentially affect the recognition performance of the final classifier. Therefore, prior to performing keratitis classification, an object localization algorithm was employed in our study to effectively filter out excessive noise and obtain the region of interest for keratitis lesions. Also, our study covered all types of keratitis, including BK, FK, viral keratitis, and other corneal abnormalities, enabling the proposed deep learning model to be capable of application in more realistic clinical scenarios. For underdeveloped areas with a scarcity of experienced ophthalmologists, the application of the OL-CDACNN enabled timely screening of keratitis to ensure appropriate treatment.

When compared to the previous studies, this study introduced several important features in the theoretical analysis. First, the one-stage SSD method was employed to efficiently localize the region of Conj Cor, filtering out most of the noise surrounding the keratitis lesion. This localization step improved the performance of the subsequent feature extraction and classification. Second, the deep dense attention module was incorporated in the CDACNN method to extracthighly discriminative features of the lesions while suppressing irrelevant features. This attention mechanism enhanced the model's ability to focus on the most relevant information of keratitis. Third, the cost-sensitive method was integrated into the loss function of the CDACNN method, fully considering the minority class to ensure that the classifier achieved high sensitivity of keratitis diagnosis in the imbalanced slit-lamp dataset. These advancements in localization algorithm, feature extraction technique, and cost-sensitive method collectively contribute to the overall effectiveness and robustness of the proposed CDACNN method for diagnosing keratitis using slit-lamp images. The experimental results demonstrated that the OL-CDACNN achieved the best performance for detecting keratitis with an AUC of 0.998 (95% CI 0.996-1), a sensitivity of 98.6% (95% CI 0.975-0.996), and an accuracy of 98.9% (95% CI 0.982-0.995) in the internal test dataset. Even in external test datasets from three different clinical centers, the OL-CDACNN still achieved satisfactory

	OL-CDACNN	Kuo et al. [24]	Gu et al. [25]	Redd et al. [26]	Ghosh et al. [<mark>27</mark>]
What was solved	 Accurately locate the region of Conj_Cor associated with keratitis lesions from original slit-lamp images Provide automatic screening of keratitis, other cornea abnormalities, and normal cornea based on slit-lamp images, with a better performance than conventional cNNs Substantial slit-lamp images from multiple-centers verified the effective- ness and generalizability of the model 	 Provide a promising tool for improv- ing first-line medical care at rural area in early identification of FK 	 The DL algorithm may be useful for computer-assisted corneal disease diagnosis with excellent performance 	 Develop computer vision models for image-based differentiation of bacte- rial and fungal corneal ulcers 	 Apply DL algorithms for rapidly discriminating between FK and BK
Methods	SSD for accurately locating the region of Conj_Cor Deep attention modules and cost-sensive method were applied to enhance sitive method were applied to enhance features and address the imbalanced data problem	DenseNet for discerning FK	Inception-v3 with multi-task multi- label classification layer	MobileNet, DenseNet, ResNet, VGG were compared and applied to distin- guish BK and FK	CNN and ensemble learning for distin- guishing FK and BK
Outcomes (95% CI)					
Accuracy (%)	98.9-100	69.4(63.9–74.5)	I	I	83.0
Sensitivity (%)	96.9-100	71.1(62.1–78.6)	I	1	77.0(0.81-0.83)
Specificity (%)	99.2-100	68.4(61.1–74.9)	I	1	I
AUC	0.997–1	0.650	0.930(0.904-0.952)	0.86(0.78-0.93)	0.904
Dataset	12.407 slit-lamp images collected from Ningbo Eye Hospital, Jiangdong Eye Hospital, Ningbo Ophthalmic Center, and Zhejiang Eye Hospital	288 slit-lamp images collected from Kaohsiung Chang Gung Memorial Hospital	5325 ocular surface slit-lamp images collected from Shanghai Eye, Ear, Nose, and Troat Hospital and the Afliated Hospital of Guizhou Medical University	980 slit-lamp images obtained from one study site within the Aravind Eye Care System in South India	2167 slit-lamp anterior segment images from 194 patients collected from Oph- thalmology Department

performance for detecting keratitis with AUCs ranging from 0.991 to 0.996, sensitivities ranging from 94.5% to 98.1%, and accuracies ranging from 95.3% to 96%, indicating that the OL-CDACNN has strong generalization ability.

The efficiency of the SSD was higher than other localization methods. Although it took 2.27 h to train the SSD method, its average testing time was only 0.049 s. As the training procedure could be implemented in advance on the local GPU server, the testing time would not increase after the trained model was deployed in ophthalmic clinics. Therefore, the SSD can be applied to the real-time localization of two regions of cornea and Conj_Cor for keratitis diagnosis.

The performance of the CDACNN method was superior to that of the DenseNet method. Two mechanisms, channel attention, and spatial attention, were performed to enable the CDACNN method to enhance the expression of keratitis-related features. As shown in Fig. 5, the t-SNE technique demonstrated that the high features extracted by the CDACNN were more separable than those of the DenseNet method. Partially normal cornea samples and keratitis samples were not easy to distinguish because they were mixed in the DenseNet method. Furthermore, the cost-sensitive was adopted to facilitate the CDACNN to focus more on the minority category of other cornea abnormalities. For the region of Conj_Cor, compared with the DenseNet method, the total number of misclassifications of the CDACNN method is reduced from 18 to 11, especially the number of misclassifications for keratitis is reduced by half.

To make the output of the OL-CDACNN interpretable, the Grad-CAM method was employed to generate heatmaps to visualize where the OL-CDACNN paid most attention to the final diagnosis result. Six representative slit-lamp images were presented to illustrate the regions contributing to the outcome of the OL-CDACNN. For the keratitis and other cornea abnormalities, the highlighted heatmaps were colocalized with the lesion regions of cornea. For the normal cornea, almost the entire cornea region was highlighted. This interpretability exploration for the OL-CDACNN could further facilitate its application in real-world clinics as ophthalmologists can understand the reason for the final diagnosis inferred by the OL-CDACNN.

Our study has several limitations. First, although the OL-CDACNN method provided an effective strategy for identifying keratitis with high performance, its efficiency is slightly lower than that of conventional CNN methods. Second, this study only explored the automatic screening for keratitis, other cornea abnormalities, and normal cornea based on slit-lamp images, however, the automatic grading of keratitis was still underinvestigated. As more and more slit-lamp images of keratitis subtypes are collected and annotated, we will try to explore and apply the OL-CDACNN method to the grading of bacterial keratitis, fungal keratitis, and virus keratitis. Despite the above limitations, this study provides a practical strategy for the automatic diagnosis of keratitis with promising generalization ability verified in the multicenter dataset.

Conclusions

In this paper, we proposed a feasible OL-CDACNN strategy for automatic diagnosis of keratitis based on the object localization method combined with cost-sensitive deep attention convolutional neural network. Our OL-CDACNN had high effectiveness and efficiency for discriminating among keratitis, other cornea abnormalities, and normal cornea in both internal and external test datasets. Qualitative and quantitative experiments verified that the proposed method was superior to other conventional methods. The region of Conj_Cor was more suitable for OL-CDACNN to be applied to the automatic diagnosis of keratitis when compared to the region of cornea and the original slitlamp image. Interpretability experiments and multicenter validation indicated that the OL-CDACNN method had better rationality and generalization ability in clinical applications. This OL-CDACNN has the high potential to be applied to digital slit-lamp cameras, which would be a cost-effective and convenient procedure for the early detection of keratitis in clinics.

Abbreviations

OL-CDACNN	Object localization method combined with cost-sensitive deep attention convolutional neural network
SSD	Single shot multibox detector
Conj_Cor	Conjunctiva and cornea
t-SNE	T-distributed Stochastic Neighbor Embedding
Grad-CAM	Gradient-weighted class activation mapping
Al	Artificial intelligence
AP	Average precision
mAP	Mean average precision
ROC	Receiver operating curve
AUCs	Area under the receiver operating characteristic curves
CI	Confidence intervals
NEH	Ningbo Eye Hospital
JEH	Jiangdong Eye Hospital
NOC	Ningbo Ophthalmic Center
ZEH	Zhejiang Eye Hospital

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40537-023-00800-w.

Additional file 1: Figure S1. Visualization of the separability for the embedding features learned by the CDACNN and DenseNet in the Jiangdong Eye Hospital (JEH) test dataset via t-SNE. Different colored point clouds represent the different categories. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. Conj_Cor, conjunctiva and cornea. t-SNE, t-distributed stochastic neighbor embedding. CDACNN, cost-sensitive deep attention convolutional neural network. Figure S2. Confusion matrices of CDACNN and DenseNet methods in the Jiangdong Eve Hospital (JEH) test dataset "Normal" indicates normal cornea "Others" indicates cornea with other abnormalities Conj_Cor, conjunctiva and cornea. CDACNN, cost-sensitive deep attention convolutional neural network. Figure S3. Visualization of the separability for the embedding features learned by the CDACNN and DenseNet in the Ningbo Ophthalmic Center (NOC) test dataset via t-SNE. Different colored point clouds represent the different categories. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. Conj_Cor, conjunctiva and cornea. t-SNE, t-distributed stochastic neighbor embedding. CDACNN, cost-sensitive deep attention convolutional neural network. Figure S4. Confusion matrices of CDACNN and DenseNet methods in the Ningbo Ophthalmic Center (NOC) test dataset. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. Conj_Cor, conjunctiva and cornea. CDACNN, cost-sensitive deep attention convolutional neural network. Figure S5. Visualization of the separability for the embedding features learned by the CDACNN and DenseNet in the Zhejiang Eye Hospital (ZEH) test dataset via t-SNE. Different colored point clouds represent the different categories. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. Conj_Cor, conjunctiva and cornea. t-SNE, t-distributed stochastic neighbor embedding. CDACNN, cost-sensitive deep attention convolutional neural network. Figure S6. Confusion matrices of CDACNN and DenseNet methods in the Zhejiang Eye Hospital (ZEH) test dataset. "Normal" indicates normal cornea. "Others" indicates cornea with other abnormalities. Conj_Cor, conjunctiva and cornea, CDACNN, cost-sensitive deep attention convolutional neural network, Table S1, Performance of the CDACNN and DenseNet in the JEH test dataset. Table S2. Performance of the CDACNN and DenseNet in the NOC test dataset. Table S3. Performance of the CDACNN and DenseNet in the ZEH test dataset.

Acknowledgements

Not applicable.

Author contributions

All authors had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. Conception and design: JJ, WL, ZL, M.Z., and J.G. Funding obtainment: JJ, ZL, and JG. Provision of study data: ZL, RG, and JJ. Collection and assembly of data: JJ, ZL, JY, WL, MZ, and JL. Data analysis and interpretation: JJ, ZL, MP, LG, and CW. Manuscript writing: all authors. Final approval of the manuscript: all authors.

Funding

This study was funded by the National Natural Science Foundation of China (Grant No. 62276210, 82201148, 61775180), the Natural Science Basic Research Program of Shaanxi (Grant No. 2022JM-380), the Natural Science Foundation of Zhejiang Province (Grant No. LQ22H120002), the Medical Health Science and Technology Project of Zhejiang Province (Grant No. 2022RC069), the International Science and Technology Cooperation Program Project Shaanxi Province Key Research and Development Program (Grant No. 2020KWZ017), the Postgraduate Innovation Fund of Xi'an University of Posts and Telecommunications (Grant No. CXJJYL2021066), and the Ningbo Science & technology program (grant no. 2021S118).

Availability of data and materials

The code and example data used in this study can be accessed at GitHub (https://github.com/jiangjiewei/Keratitis-OL-CDACNN). The datasets generated and/or analyzed during the current study are available upon reasonable request from the corresponding author. Correspondence and requests for data materials should be addressed to Zhongwen Li, Jiamin Gong or Mingmin Zhu.

Declarations

Ethics approval and consent to participate

The Institution Review Board of NEH approved the study. Informed consent was exempted, due to the retrospective nature of the data acquisition and the use of deidentified images.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 25 August 2022 Accepted: 16 July 2023

Published online: 24 July 2023

References

- 1. Bourne RR, Stevens GA, White RA, Smith JL, Flaxman SR, Price H, et al. Causes of vision loss worldwide, 1990–2010: a systematic analysis. Lancet Glob Health. 2013;1(6):e339–49.
- Flaxman SR, Bourne RR, Resnikoff S, Ackland P, Braithwaite T, Cicinelli MV, et al. Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis. Lancet Glob Health. 2017;5(12):e1221–34.
- Burton MJ. Corneal blindness: prevention, treatment and rehabilitation. Community eye health. 2009;22(71):33.
 Varacalli G, Di Zazzo A, Mori T, Dohlman TH, Spelta S, Coassin M, et al. Challenges in Acanthamoeba keratitis: a
- Varacalii G, Di Zazzo A, Mori I, Doniman IH, Speita S, Coassin M, et al. Challenges in Acanthamoeda keratitis: a review. J Clin Med. 2021;10(5):942.
- Soifer M, Wisely CE, Carlson AN. In vivo confocal microscopy evaluation of microbial keratitis. JAMA Ophthalmol. 2021;139(11):1240–1.
- Austin A, Lietman T, Rose NJ. Update on the management of infectious keratitis. Ophthalmology. 2017;124(11):1678–89.
- Tuft S, Somerville TF, Li JPO, Neal T, De S, Horsburgh MJ, et al. Bacterial keratitis: identifying the areas of clinical uncertainty. Prog Retinal Eye Res. 2021. https://doi.org/10.1016/j.preteyeres.2021.101031.
- Lobo AM, Agelidis AM, Shukla D. Pathogenesis of herpes simplex keratitis: the host cell response and ocular surface sequelae to infection and inflammation. Ocul Surf. 2019;17(1):40–9.
- 9. Li W, Yang Y, Zhang K, Long E, He L, Zhang L, et al. Dense anatomical annotation of slit-lamp images improves the performance of deep learning for the diagnosis of ophthalmic disorders. Nat Biomed Eng. 2020;4(8):767–77.
- Weiss M, Molina R, Ofoegbuna C, Johnson DA, Kheirkhah A. A review of filamentary keratitis. Surv Ophthalmol. 2022;67(1):52–9.
- Singh RB, Das S, Chodosh J, Sharma N, Zegans ME, Kowalski RP, et al. Paradox of complex diversity: challenges in the diagnosis and management of bacterial keratitis. Prog Retinal Eye Res. 2021. https://doi.org/10.1016/j.preteyeres. 2021.101028.
- Hoffman JJ, Burton MJ, Leck A. Mycotic keratitis—a global threat from the filamentous fungi. J Fungi. 2021;7(4):273.
 Ung L, Bispo PJ, Shanbhag SS, Gilmore MS, Chodosh J. The persistent dilemma of microbial keratitis: global burden,
- diagnosis, and antimicrobial resistance. Surv Ophthalmol. 2019;64(3):255–71.
- 14. Kredics L, Narendran V, Shobana CS, Vágvölgyi C, Manikandan P, Group IHFKW. Filamentous fungal infections of the cornea: a global overview of epidemiology and drug sensitivity. Mycoses. 2015;58(4):243–60.
- Gulshan V, Rajan RP, Widner K, Wu D, Wubbels P, Rhodes T, et al. Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. JAMA Ophthalmol. 2019;137(9):987–93.
- Holmberg OG, Köhler ND, Martins T, Siedlecki J, Herold T, Keidel L, et al. Self-supervised retinal thickness prediction enables deep learning from unlabelled data to boost classification of diabetic retinopathy. Nat Mach Intell. 2020;2(11):719–26.
- Li Z, He Y, Keel S, Meng W, Chang RT, He M. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology. 2018;125(8):1199–206.
- Keel S, Wu J, Lee PY, Scheetz J, He M. Visualizing deep learning models for the detection of referable diabetic retinopathy and glaucoma. JAMA Ophthalmol. 2019;137(3):288–92.

- Jiang J, Liu X, Zhang K, Long E, Wang L, Li W, et al. Automatic diagnosis of imbalanced ophthalmic images using a cost-sensitive deep convolutional neural network. Biomed Eng Online. 2017;16(1):1–20.
- Jiang J, Wang L, Fu H, Long E, Sun Y, Li R, et al. Automatic classification of heterogeneous slit-illumination images using an ensemble of cost-sensitive convolutional neural networks. Ann Transl Med. 2021. https://doi.org/10.2103/ atm-20-6635.
- 21. Li Z, Guo C, Nie D, Lin D, Zhu Y, Chen C, et al. Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images. Commun Biol. 2020;3(1):1–10.
- 22. Yan Q, Weeks DE, Xin H, Swaroop A, Chew EY, Huang H, et al. Deep-learning-based prediction of late age-related macular degeneration progression. Nat Mach Intell. 2020;2(2):141–50.
- Li Z, Qiang W, Chen H, Pei M, Yu X, Wang L, et al. Artificial intelligence to detect malignant eyelid tumors from photographic images. NPJ Digit Med. 2022;5(1):1–9.
- 24. Kuo MT, Hsu BWY, Yin YK, Fang PC, Lai HY, Chen A, et al. A deep learning approach in diagnosing fungal keratitis based on corneal photographs. Sci Rep. 2020;10(1):1–8.
- Gu H, Guo Y, Gu L, Wei A, Xie S, Ye Z, et al. Deep learning for identifying corneal diseases from ocular surface slitlamp photographs. Sci Rep. 2020;10(1):1–11.
- 26. Redd TK, Prajna NV, Srinivasan M, Lalitha P, Krishnan T, Rajaraman R, et al. Image-based differentiation of bacterial and fungal keratitis using deep convolutional neural networks. Ophthalmol Sci. 2022;2(2): 100119.
- 27. Ghosh AK, Thammasudjarit R, Jongkhajornpong P, Attia J, Thakkinstian A. Deep learning for discrimination between fungal keratitis and bacterial keratitis: deepkeratitis. Cornea. 2022;41(5):616.
- Hung N, Shih AKY, Lin C, Kuo MT, Hwang YS, Wu WC, et al. Using slit-lamp images for deep learning-based identification of bacterial and fungal keratitis: model development and validation with different convolutional neural networks. Diagnostics. 2021;11(7):1246.
- 29. Tiwari M, Piech C, Baitemirova M, Prajna NV, Srinivasan M, Lalitha P, et al. Differentiation of active corneal infections from healed scars using deep learning. Ophthalmology. 2022;129(2):139–46.
- Huang G, Liu Z, Van Der Maaten L, Weinberger KQ, editors. Densely connected convolutional networks. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2017; 4700–4708.
- 31. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z, editors. Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016; 2818–2826.
- Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv. 2014. https:// doi.org/10.48550/arXiv.1409.1556.
- 33. Lindquist TD, Lindquist TP. Conjunctivitis: an overview and classification. Cornea, E-Book. 2021:358.
- Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu C Y, et al., editors. Ssd: Single shot multibox detector. 2016 European Conference on Computer Vision (ECCV). Springer; 2016; 21–37.
- Ren S, He K, Girshick R, Sun J. Faster r-cnn: Towards real-time object detection with region proposal networks. IEEE Trans. Pattern Anal. Mach. Intell. 2017; 1137–1149.
- Cai Z, Vasconcelos N, editors. Cascade r-cnn: Delving into high quality object detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016; 6154–6162.
- Li Y, Chen Y, Wang N, Zhang Z, editors. Scale-aware trident networks for object detection. IEEE International Conference on Computer Vision (ICCV). IEEE; 2019; 6054–6063.
- He K, Zhang X, Ren S, Sun J, editors. Deep residual learning for image recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE; 2016; 770–778.
- Lin T Y, Goyal P, Girshick R, He K, Dollár P, editors. Focal loss for dense object detection. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017; 2980–2988.
- 40. Woo S, Park J, Lee JY, Kweon IS. European Conference on Computer Vision (ECCV). Cham: Springer; 2018.
- Krawczyk B, Schaefer G, Woźniak M. A hybrid cost-sensitive ensemble for imbalanced breast thermogram classification. Artif Intell Med. 2015;65(3):219–27.
- 42. Paszke A, Gross S, Massa F, Lerer A, Bradbury J, Chanan G, et al. Pytorch: An imperative style, high-performance deep learning library. 2019 Advances in Neural Information Processing Systems (NIPS). MIT Press; 2019; 32
- 43. Van der Maaten L, Hinton G. Visualizing data using t-SNE. J Mach Learn Res. 2008;9(11).
- 44. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D, editors. Grad-cam: Visual explanations from deep networks via gradient-based localization. 2017 IEEE International Conference on Computer Vision (ICCV). IEEE; 2017; 618–626.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.