The ROYAL COLLEGE *of*
OPHTHALMOLOGISTS

Check for updates

ARTICLE

Joint optic disk and cup segmentation for glaucoma screening using a region-based deep learning network

Feng Li[1,5], Wenjie Xiang[1,5], Lijuan Zhang[2,5], Wenzhe Pan[1], Xuedian Zhang[1,3], Minshan Jiang [1✉] and Haidong Zou[4]

OBJECTIVES: To develop and validate an end-to-end region-based deep convolutional neural network (R-DCNN) to jointly segment the optic disc (OD) and optic cup (OC) in retinal fundus images for precise cup-to-disc ratio (CDR) measurement and glaucoma screening.
METHODS: In total, 2440 retinal fundus images were retrospectively obtained from 2033 participants. An R-DCNN was presented for joint OD and OC segmentation, where the OD and OC segmentation problems were formulated into object detection problems. We compared R-DCNN's segmentation performance on our in-house dataset with that of four ophthalmologists while performing quantitative, qualitative and generalization analyses on the publicly available both DRISHIT-GS and RIM-ONE v3 datasets. The Dice similarity coefficient (DC), Jaccard coefficient (JC), overlapping error (E), sensitivity (SE), specificity (SP) and area under the curve (AUC) were measured.
RESULTS: On our in-house dataset, the proposed model achieved a 98.51% DC and a 97.07% JC for OD segmentation, and a 97.63% DC and a 95.39% JC for OC segmentation, achieving a performance level comparable to that of the ophthalmologists. On the DRISHTI-GS dataset, our approach achieved 97.23% and 94.17% results in DC and JC results for OD segmentation, respectively, while it achieved a 94.56% DC and an 89.92% JC for OC segmentation. Additionally, on the RIM-ONE v3 dataset, our model generated DC and JC values of 96.89% and 91.32% on the OD segmentation task, respectively, whereas the DC and JC values acquired for OC segmentation were 88.94% and 78.21%, respectively.
CONCLUSION: The proposed approach achieved very encouraging performance on the OD and OC segmentation tasks, as well as in glaucoma screening. It has the potential to serve as a useful tool for computer-assisted glaucoma screening.

Eye (2023) 37:1080–1087; https://doi.org/10.1038/s41433-022-02055-w

INTRODUCTION

Glaucoma is a chronic and painless eye disorder characterized by progressive degeneration of the optic nerves, which can result in irreversible vision loss or even permanent blindness [1, 2]. The early screening and diagnosis of glaucoma are beneficial for preserving vision and quality of life. Optic nerve head (ONH) assessment is used as one of the most clinically significant glaucoma screening techniques, where the cup-to-disc ratio (CDR) serves as the most representative measurement indicator [3–5]. Accurate optic disc (OD) and optic cup (OC) segmentation is the premise for precise CDR computation and a guarantee for correct glaucoma screening and diagnosis. In general, CDR is obtained by means of manually delineating the borders of the OD/OC or manually correcting the contours produced by segmentation algorithms [6–8]. Nevertheless, the manual contouring of OD/OC borders is laborious, expensive, and subjected to personal experiences. As a result of the obscure border information of OD/OC, the CDR value for the same subject is often subject to substantial inter- and intra-observer variability. Accordingly, an automatic segmentation technique without human intervention or manual drawing that can jointly segment OD and OC from retinal fundus images to achieve precise CDR measurement is highly desired.

OD and OC segmentation are admittedly challenging tasks as a result of large appearance variations, small target regions, low-contrast boundaries, blood vessel occlusions, pathological changes, and variable imaging conditions. It was reported in [9–12] that the OD and OC boundaries were approximated as elliptical shapes with vertical diameters and horizontal diameters in most cases, respectively. In terms of a simple yet effective assumption, several ellipse fitting methods [10–20] have been developed for OD and OC segmentation. However, the segmentation results predicted by these approaches needed to be postprocessed through an ellipse fitting procedure to generate ellipses for the OD and OC regions.

To mitigate that limitation, great efforts have been devoted to automatically segmenting OD and OC from retinal fundus images, producing statistical shape models [21–24], multiview and multi-modal approaches [25–27], and superpixel-based methods [28–31]. Notwithstanding some specific evaluation criteria, these methods rely heavily on the utilization of handcrafted OD and OC features, which lack sufficient discriminative representations. In practice, it is difficult to design good handcrafted features, which leads to degraded segmentation performance relative to that achieved by the powerful layer-by-layer feature learning abilities

---

[1]School of Optical-Electrical and Computer Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]School of Electrical and Electronic Engineering, Shanghai Institute of Technology, Shanghai 201418, China. [3]School of Medical Imaging, Shanghai University of Medicine and Health Sciences, Shanghai 201318, China. [4]Department of Ophthalmology, Shanghai First People's Hospital, Shanghai 200080, China. [5]These authors contributed equally: Feng Li, Wenjie Xiang, Lijuan Zhang. ✉email: jiangmsc@gmail.com

of deep learning networks. In recent years, deep learning, especially convolutional neural networks, which yield highly discriminative feature representations and promote the development of many computer vision tasks, has yielded heightened performance in image segmentation tasks. In particular, the fully convolutional network (FCN) [32–34], modified U-Net [35, 36], M-Net [12], and cup disc encoder decoder network (CDED-Net) [37] completely changed the traditional image segmentation field and have provided excellent OD and OC segmentation results. Nevertheless, most of these methods still regard a segmentation problem as the two separate problems, while some prior knowledge is not fully used (e.g., the OC is located in the OD, and the OD and OC are approximately ellipses). In addition, many algorithms require a great deal of time to postprocess the segmentation results with other strategies, thereby abandoning the end-to-end learning paradigm.

Inspired by these abovementioned methods, we followed the core assumption that OD and OC can be approximated by ellipses and used the impressive object detection ability of deep learning from the computer vision community to investigate the challenging OD and OC segmentation problems. In this study, we concentrated on the joint OD and OC segmentation problems, and developed and validated an end-to-end region-based deep convolutional neural network (R-DCNN). We formulated the OD and OC segmentation problems into object detection problems and designed a disc proposal network (DPN) and a cup proposal network (CPN) to yield minimal candidate bounding boxes for OD and OC in a sequential way. Additionally, we adopted ResNet34 as the backbone while utilizing dense atrous convolution (DAC) to extract more dense feature maps which were shared for OD and OC segmentation. Considering that the OC is located at the center of the OD, we applied a disc attention module with an attention mechanism to cascade the DPN and CPN so that the corresponding OD region could be cropped to further guide the OC localization, which made the OD and OC positively influence each other. The boundaries of OD and OC could be determined by calculating an inscribed ellipse within the corresponding bounding boxes, and then the CDR could be calculated for glaucoma screening. We evaluated the performance of the proposed approach against that of four ophthalmologists on our in-house test dataset while conducting a comparison with various types of recently proposed mainstream segmentation approaches on the publicly available DRISHTI-GS and RIM-ONE v3 datasets. Furthermore, we performed a qualitative evaluation of our model and investigated its generalization ability to different datasets. Extensive experiments clearly demonstrated that our method exceeded the state-of-the-art approaches with respect to the OD and OC segmentation and glaucoma detection tasks.

METHODS
Datasets
With the aim of algorithm development, 2440 retinal fundus images from 2033 participants were retrospectively retrieved from Shanghai First People's Hospital. A further description of the datasets is provided in the "Datasets" subsection of the eMethods section in the Supplemental Material. This research was conducted under the Declaration of Helsinki as revised in 2013 and approved by the local ethics review and institutional review boards. As a result of the retrospective and anonymized nature of this study, written consent was waived by the institutional review board. In addition, two publicly available datasets (DRISHTI-GS and RIM-ONE v3) were also used for training and testing.

Retinal fundus image annotation
Four ophthalmologists with an average of 7 years of experience in this field (ranging from 5 to 8 years) were invited from Shanghai First People's Hospital to manually annotate the OD and OC and perform image labeling (glaucomatous/non-glaucomatous labels). We also investigated the statistical agreement between the ophthalmologists regarding the examination of identical samples. The details are described in the "Retinal

fundus image annotation" subsection of the eMethods section in the Supplement. The resulting dataset was divided into three sets at the patient level: an in-house training set, an in-house validation set, and an in-house test set in an approximate 7:2:1 ratio. Each set was stratified in such a way that all datasets contained equal proportions of glaucomatous (~40%) and non-glaucomatous (~60%) cases, as listed in eTable 1 in the Supplement. During glaucoma screening, the discrimination between glaucomatous and non-glaucomatous cases made by the proposed R-DCNN was based on the estimated vertical CDR. If it exceeded a threshold value of 0.5, glaucoma was suspected, and a larger CDR value indicated a greater risk of glaucoma.

Image preprocessing and data augmentation
As a pre-processing, we used contrast limited adaptive histogram equalization (CLAHE) [38] for the original images within these datasets, which equalized their contrasts by changing the colors of image areas and interpolating the results across them. Then, we cropped an 800×800 region of interest (ROI) in each retinal fundus image according to the OD localization algorithm developed in [39]. Taking the limited number of images in these datasets into account and to prevent overfitting, we applied data augmentation. All images were subjected to horizontal and vertical flipping and rotation operations of 90, 180 and 270 degrees.

Development of the R-DCNN architecture
The proposed R-DCNN was composed of ResNet34 for basic feature extraction, a DPN for OD segmentation, a CPN for OC segmentation, and a disc attention module with an attention mechanism for connecting the DPN with CPN and cropping the OD regions from feature maps. To extract more dense feature representations from the original fundus images and preserve more spatial information, we inserted DAC into the ResNet34 network, the output feature maps of which were shared among the DPN, CPN and disc attention module. The DPN and CPN could produce candidate bounding boxes for OD and OC, respectively. Considering the prior information that the OC was located at the center of the OD, we further applied the disc attention module with attention mechanism to link the DPN with CPN, where the corresponding OD area was trimmed to assist in guiding the OC detection. The whole framework is presented in Fig. 1. Initially, the fundus images were fed into DAC-ResNet34 for feature extraction, where the DAC block is illustrated in eFig. 1 in the Supplement. Then, the DPN and CPN were designed to segment OD and OC, respectively. In the DPN/CPN, feature maps were input into a region proposal network (RPN) and cropped by the ROI pooling mechanism in accordance with the coordinates of the candidate bounding boxes. These cropped feature maps were further fed into the classifier, which could offer the final predictions about the candidate disc/cup regions with the highest probabilities. Finally, a disc attention module with an attention mechanism for fusing different feature maps from different stages was designed to chain the DPN and CPN and trim the corresponding OD region for the CPN. Further detailed descriptions of the R-DCNN architecture and the model training process are included in the eMethods section in the Supplement.

Statistical analysis
In this work, the Dice similarity coefficient (DC), Jaccard coefficient (JC), overlapping error (E), sensitivity (SE) and specificity (SP) were utilized to evaluate the OD and OC segmentation results. Here, these metrics were defined as follows:

$$Dice(DC) = \frac{2 \times TP}{2 \times TP + FP + FN} \quad (1)$$

$$Jaccard(JC) = \frac{TP}{TP + FP + FN} \quad (2)$$

$$E = 1 - \frac{Area(S \cap G)}{Area(S \cup G)} \quad (3)$$

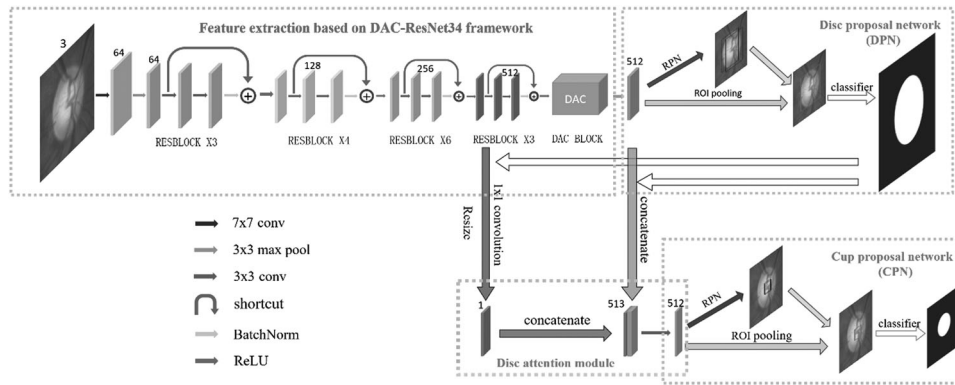$$SE = \frac{TP}{TP + FN} \quad (4)$$

$$SP = \frac{TN}{TN + FP} \quad (5)$$

Fig. 1   The overall architecture of our proposed R-DCNN. *RPN* Region proposal network, *DAC* Dense atrous convolution, *DPN* Disc proposal network, *CPN* Cup proposal network, *ROI* Region of interest.

where *TP*, *FP*, *TN* and *FN* denote true positives, false positives, true negatives, and false negatives, respectively. *S* and *G* refer to the segmentation area and ground truth, respectively.

In addition, we calculated the *CDR* as follows:

$$CDR = \frac{VCD}{VDD} \tag{6}$$

where *VCD* and *VDD* represent the vertical cup diameter and disc diameter, respectively.

Afterwards, we further evaluated the performance of the proposed R-DCNN with respect to glaucoma screening via receiver operating characteristic (ROC) curves and the area under the curve (AUC). Additionally, we performed Student's t test to measure the significance levels of the improvements yielded by the major components in our model.

RESULTS
Quantitative results
To see whether the segmentation results of the proposed R-DCNN were comparable to those hand-annotated by ophthalmologists, we conducted a consistency comparison between the R-DCNN and the four ophthalmologists on our in-house test dataset, as listed in eTable 2 in the Supplement. The results indicated that the stability of the R-DCNN was slightly weaker than those of the four different ophthalmologists. Nevertheless, the DC and JC for OD segmentation between the R-DCNN and each ophthalmologist were 98.51% and 97.07% on average and 97.63% and 95.39% on average for OC segmentation, respectively. These results were higher than the averages of 97.76%/96.41% and 97.05%/95.09% among the interophthalmologist values.

We also compared the OD and OC segmentation performance of our model with that of the most competitive segmentation methods on the DRISHTI-GS dataset and RIM-ONE v3 dataset. The compared methods included a regression-based method [21], a superpixel-based method [30], a modified U-Net [36], M-Net [12], the context encoder network (CE-Net) [40], and CDED-Net [37], as reported in Table 1 below, as well as in eFig. 2 and eFig. 3 in the Supplement. As seen, the proposed approach achieved the most outstanding results on both datasets for the OD and OC segmentation tasks when compared to the other existing state-of-the-art methods [12, 21, 24, 30, 36, 37, 40]. This finding demonstrated the benefits of constructing the OD and OC segmentation tasks as bounding box detection problems in our work. In addition, we attributed this improvement to the disc attention module with an attention mechanism that was used to chain the DPN and CPN, which enabled the OD and OC results to influence each other positively.

We measured the accuracy of the proposed segmentation model on both glaucomatous and non-glaucomatous cases. For OD and OC segmentation, no significant accuracy differences were observed. The details are reported in the "Segmentation

performance of the proposed model on both glaucomatous and non-glaucomatous cases" subsection of the eResults section in the Supplement. In addition, we also constructed a test dataset consisting of images of advanced glaucoma symptoms with notches and compared the OD and OC segmentation results obtained on these images with those of both ellipse fitting and anchor point segmentation (see the "Segmentation performance comparison of the proposed model with ellipse fitting methods" subsection of the eResults section in the Supplement). It was evident from these results that our R-DCNN performed better than the ellipse fitting methods with respect to OD and OC segmentation for images of advanced glaucoma symptoms with notches in terms of the DC and JC.

Qualitative results
The qualitative comparison results obtained on the OD and OC segmentation tasks with the superpixel-based method [30], modified U-Net method [36], and M-Net method [12] on DRISHTI-GS and RIM-ONE v3 are shown in Fig. 2a and b, respectively. As can be clearly observed, our proposed R-DCNN generated more accurate segmentation results than these cutting-edge methods. In particular, for the low-contrast sample images in the fifth row of Fig. 2a and the second row of Fig. 2b, our approach could still segment OD and OC with more accurate boundaries that were close to the ground truth. The success of our R-DCNN was mainly ascribed to the fact that DAC-ResNet34 extracted more feature representations of the objective region boundaries. Additionally, the disc attention module with an attention mechanism between the DPN and CPN also contributed a great deal to our approach's success, which could guide the CPN to locate the OC in terms of the output of the DPN.

Utilizing RIM-ONE v3 for training and DRISHTI-GS for testing
To further verify the generalizability of our approach, we also performed a cross-training experiment by using the RIM-ONE v3 dataset for training and the DRISHTI-GS dataset for testing. The results are listed in eTable 3 in the Supplement. Among these cross-training results, the best DC and JC scores of 96.43%/91.17% and 88.24%/77.93% were achieved by our approach on the OD and OC segmentation tasks, respectively. Through these results, we could observe that the proposed approach clearly improved the DC and JC metrics for OD and OC segmentation over those of the other state-of-the-art methods [12, 30, 36, 37, 40], demonstrating the good generalization ability of our method across unseen retinal fundus images.

Glaucoma screening
We further assessed the performance of the proposed R-DCNN for glaucoma screening on both the DRISHTI-GS dataset and RIM-ONE v3 dataset with the help of the computed CDR values. We report

Table 1. Results of the proposed method for OD and OC segmentation on DRISHTI-GS dataset (a) and RIM-ONE v3 dataset (b) respectively, compared with the existing methods.

	Author	Method	OD					OC				
			DC(%)	JC(%)	E(%)	SE(%)	SP(%)	DC(%)	JC(%)	E(%)	SE(%)	SP(%)
DRISHTI-GS dataset (a)	Sedai et al. [21]	Regression	95	–	–	–	–	85	–	–	–	–
	Sevastopolsky [36]	Modified U-net	90.43	83.50	10.8	91.56	99.69	85.21	75.15	28.5	0.8476	98.81
	Fu et al. [12]	M-net	96.78	93.86	–	97.11	99.91	86.18	77.30	–	88.22	98.62
	Gu et al. [40]	CE-net	96.42	93.23	–	97.59	99.9	88.18	80.06	–	88.19	99.09
	Tabassum [37]	CDED-Net	95.97	91.83	8.2	97.54	99.73	92.40	86.32	13.6	95.67	99.81
	Ours	Proposed	**97.23**	**94.17**	**7.7**	**97.62**	**99.91**	**94.56**	**89.92**	**10.3**	**96.89**	**99.89**
RIM-ONE v3 dataset (b)	Cheng et al. [30]	Superpixel	89.20	82.93	–	–	–	74.4	73.2	–	–	–
	Sevastopolsky [36]	Modified U-net	95	89	11	95.02	99.73	82	69	31	75.45	99.76
	Fu et al. [12]	M-Net	95.26	91.14	–	94.81	99.86	84.48	73	–	81.42	99.65
	Gu et al. [40]	CE-Net	95.27	91.15	8.7	95.02	99.86	84.35	74.24	16.5	83.52	99.70
	Tabassum [37]	CDED-Net	95.82	91.01	7.9	97.34	99.73	86.22	75.32	11.6	95.17	99.81
	Ours	Proposed	**96.89**	**91.32**	**7.5**	**97.78**	**99.90**	**88.94**	**78.21**	**9.1**	**96.68**	**99.83**

The bold entries represent the best results on the corresponding metrics.
DC Dice similarity coefficient, *JC* Jaccard, *E* Overlapping error, *SE* Sensitivity, *SP* Specificity.

the ROC curves with their AUC scores as the overall measures of screening strength, as presented in Fig. 3. It could be observed that the proposed approach generated obviously higher AUC values (0.968 and 0.941) than the other state-of-the-art methods [12, 36, 41] on both datasets. In terms of the AUC metric, our approach boosted the performance by 8.50%, 7.20% and 4.10% on the DRISHTI-GS dataset over the modified U-Net [36], M-Net [12], and fuzzy broad learning system (FBLS) [41], respectively, yet relative 7.70%, 6.50% and 3.50% improvements were achieved on the RIM-ONE v3 dataset, respectively. The reason for this finding may be that the more accurate CDR estimations were derived from better OD and OC segmentation effects, leading to better glaucoma screening results. Such encouraging results justified the efficacy of the proposed algorithm in glaucoma screening. In addition, it was noteworthy that the performance of our approach was consistent on both datasets, exhibiting its generalization capability on unseen data.

Ablation analysis
To investigate the effectiveness of the major components (including the DAC, DPN and CPN chaining, and disc attention module with an attention mechanism) in our R-DCNN on OD and OC segmentation, we carried out thorough ablation experiments, and the results are listed in Table 2.

Effectiveness of the DAC module. To explore the contribution of the DAC module, we compared the proposed DAC-ResNet34 network with the ResNet34 network. From the experimental results reported in Table 2, without the DAC module, the E values of OD and OC segmentation increased by approximately 0.4% and 0.5%, respectively. We obtained $p < 0.001$ for OD segmentation and $p = 0.037$ for OC segmentation. This suggested that the DAC module effectively improved the OD and OC segmentation performance of our deep learning framework.

Effectiveness of chaining the DPN and CPN modules. We directly fed the features extracted from DAC-ResNet34 into the DPN and CPN modules to investigate the importance of chaining both modules. From Table 2, we could observe that when regarding OD and OC as two completely independent objects in fundus images, the performances of OD and OC segmentation on E decreased by 0.9% and 1.0%, respectively. Compared with "Without chaining the DPN and CPN modules", we attained $p < 0.001$ for OD and OC segmentation. This illustrated that chaining the DPN and CPN modules could enable our model to consider the spatial connection between OD and OC and generate positive effects on each other, leading to the significant model performance improvements on the OD and OC segmentation tasks.

Effectiveness of the disc attention module. Finally, we chained the DPN and CPN modules and directly employed the cropped feature maps obtained from DAC-ResNet34 to investigate the contribution of the disc attention module. As shown in Table 2, it could be clearly observed that our disc attention module incorporating the corresponding OD areas from different stages of feature maps, dropped the error ratios of OD and OC segmentation by 0.1% and 0.6%, respectively. The p values ($p = 0.465$ for OD segmentation and $p < 0.001$ for OC segmentation) revealed that introducing the disc attention module with an attention mechanism could significantly decrease the error of OC segmentation, whereas it exhibited no obvious error reduction for OD segmentation.

DISCUSSION
In the task of OD segmentation, the segmentation results of our approach matched or exceeded those of other state-of-the-art methods [12, 21, 30, 36, 37, 40] on both the DRISHTI-GS dataset and RIM-ONE v3 dataset. This could mainly be because the
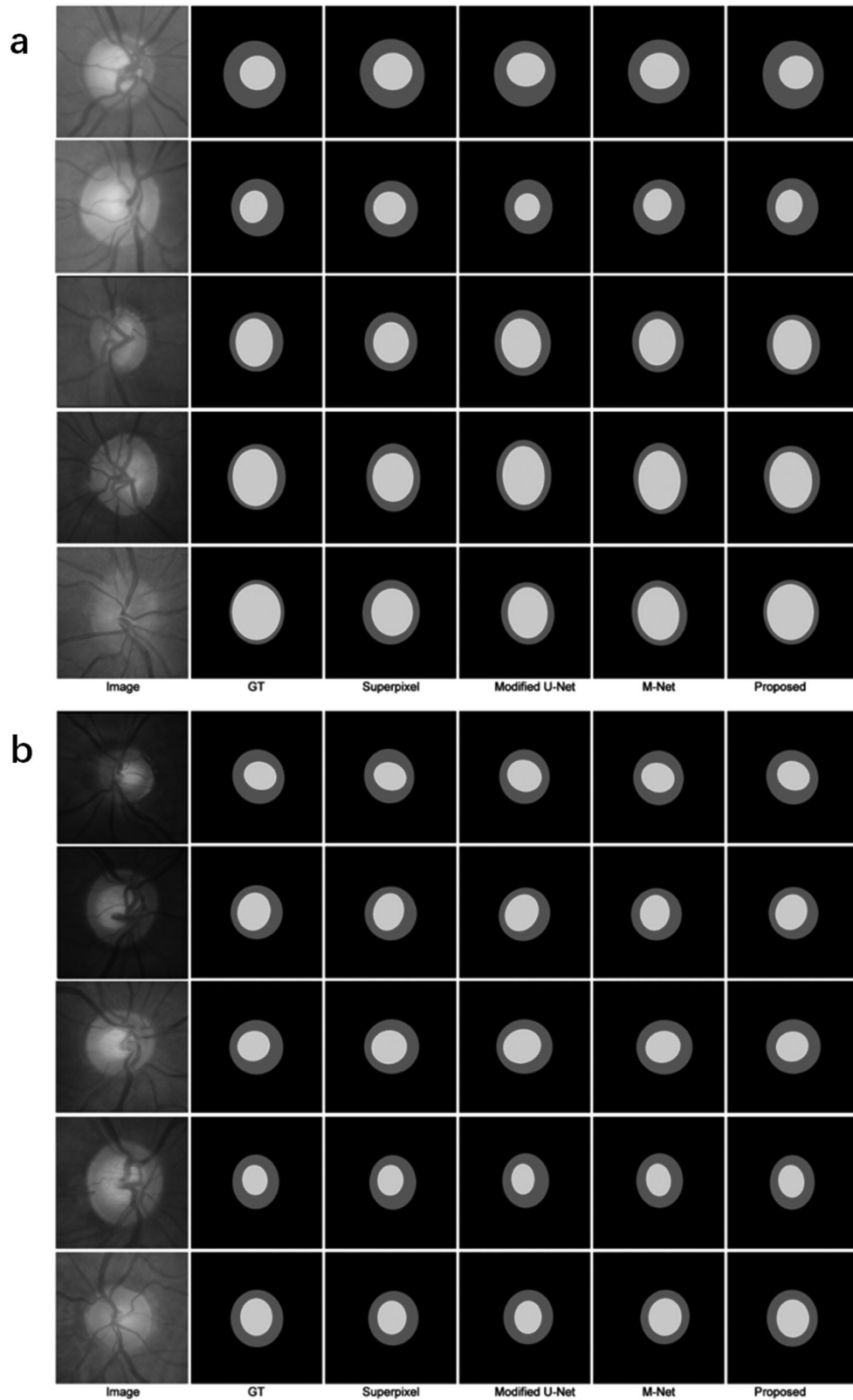
Fig. 2 **Qualitative comparison of different methods for OC and OD segmentation.** Qualitative segmentation results of our proposed approach in comparison with ground truth, superpixel [30], modified U-Net [36], and M-Net [12] on DRISHTI-GS dataset (**a**) and RIM-ONE v3 dataset (**b**), where the yellow and red regions denoted OC and OD segmentations, respectively. GT Ground truth.

ResNet34 network used for feature extraction loaded an ImageNet pretrained model as the initialized parameters during training, while DAC was incorporated into the ResNet34 network to enlarge the fields of view, densely extract deep feature maps and preserve more spatial information for objects of various sizes. Benefitting from the pretrained ResNet34 network and DAC, the hierarchical deep feature maps of DAC-ResNet34 effectively represented the complex hidden patterns for the benefit of OD segmentation.

On more challenging OC segmentation tasks, our approach significantly outperformed other cutting-edge methods [12, 21, 30, 36, 37, 40]. This approach mainly benefited from the bounding box detection process performed by the DPN, the CPN and the disc attention module with an attention mechanism between them. In our proposed approach, OD and OC were considered different objects in fundus images, and the original object segmentation problems were formulated into easier object detection issues. In addition, we only needed to calculate the inscribed ellipses without complex postprocessing procedures when the OD and OC bounding boxes were derived from the DPN and CPN. Between the DPN and CPN, the disc attention module with an attention mechanism for integrating the prior OC information was located at the center of the OD. This module enabled the OD and OC results to have a positive influence on each other, leading to a further performance boost for joint OD and OC segmentation.

Although it is a single decisive clinical indicator for glaucoma diagnosis, the CDR is the most common indicator used by ophthalmologists. In general, a higher CDR value indicates a higher risk of glaucoma. Accordingly, it is desirable to estimate more accurate CDRs for large-scale glaucoma screening. Our approach achieved better OD and OC segmentation effects than other methods [12, 21, 30, 36, 37, 40] so that more accurate CDRs

could be calculated. With the help of the accurately computed CDRs, our method achieved excellent performance in glaucoma screening and generated the best AUC scores, which were nearly 7.20% and 6.50% better than those acquired by M-Net [12] on both datasets, respectively. In addition, a cross-training experiment using the RIM-ONE v3 dataset for training and the DRISHTI-GS dataset for testing demonstrated the good generalization capacity of our approach across unseen datasets. Our method could potentially aid ophthalmologists in glaucoma screening programs involving fundus images worldwide and was both scientifically interesting and clinically impactful. A discussion on calculating neuroretinal rim areas and assessing retinal nerve fiber layer (RNFL) thickness is described in the "eDiscussion" section in the Supplement. Compared to the information obtained in a few minutes based on the optical coherence tomography (OCT) of the optic nerve, we also discussed the advantages of our approach in glaucoma screening in the "eDiscussion" section in the Supplement.

An ablation study regarding the DPN and CPN chaining indicated that our proposed DAC could make full use of the spatial connection between OD and OC to make their results influence each other in a positive fashion, which significantly decreased the errors in OD and OC segmentation. Through an ablation study on the disc attention module with an attention mechanism, we confirmed that this module, which associated the corresponding OD regions from multiple stages of feature maps, also reduced the error ratios in OD and OC segmentation. In particular, the reduction in the OC segmentation error was obvious. This was reasonable, as the disc attention module with an attention mechanism concatenated local information derived from large feature maps and global information derived from small feature maps for cropping the corresponding OD regions to further help complete OC segmentation. The ablation study on DAC showed that DAC contributed a great deal to OD segmentation as a result of extracting deeper feature maps and preserving more spatial information.

Despite generating promising results in the joint OD and OC segmentation, as well as glaucoma screening tasks, the proposed method has several limitations. First, due to the limited size of the in-house dataset, we only utilized image rotation and horizontal/vertical flipping to enlarge the dataset. In future work, we will apply a generative adversarial network (GAN) [42] or conditional variational autoencoders (CVAEs) [43] to synthesize more samples so that the segmentation performance can be further validated. Second, in the current study, the performance of our approach was investigated only using our in-house dataset and two publicly available datasets. In the future, it will be necessary to collect multiple datasets with corresponding annotations to help improve

Table 2. Ablation experiments for our experiment.

-	E_{disc}	p-value	E_{cup}	p-value
Without DAC module	8.1	<0.001	10.8	0.037
Without chaining DPN and CPN	8.6	<0.001	11.3	<0.001
Without disc attention module	7.8	0.465	10.9	<0.001
Proposed	**7.7**	–	**10.3**	–

The bold entries represent the best results on the corresponding metrics. *DAC* Dense atrous convolution, *DPN* Disc proposal network, *CPN* Cup proposal network
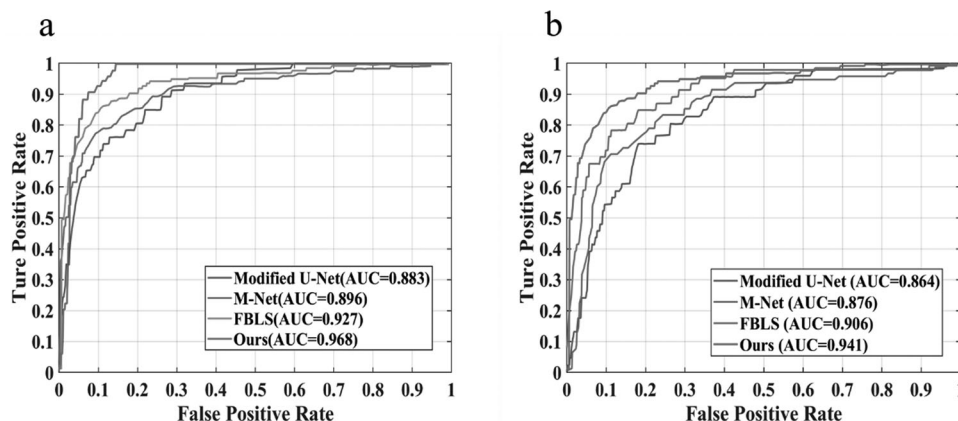

Fig. 3 ROC curves of the two datasets for glaucomma screening. The ROC curves with AUC scores for glaucoma screening on both DRISHTI-GS dataset (**a**) and RIM-ONE v3 dataset (**b**). ROC Receiver operating characteristic.

the overall network performance. Finally, we only carried out experiments on OD and OC segmentation tasks. In future work, we will extend our approach to other medical image segmentation tasks to verify its effectiveness.

CONCLUSIONS

In this paper, we developed a novel network called an R-DCNN to jointly segment OD and OC for precise CDR measurement and glaucoma screening. This involved constructing the original OD and OC segmentation problems into object detection issues. Through quantitative, qualitative and generalization analyses, the excellent performance of the proposed method on OD and OC segmentation and glaucoma screening tasks was demonstrated. The success of the designed approach provided a useful and efficient tool for computer-assisted glaucoma screening in clinical practice.

Summary
What was known before

- Traditional segmentation methods heavily rely on the utilization of hand-crafted features of OD and OC, and lead to degrade segmentation performance.
- Segmentation performance is often subject to substantial inter- and intra-observer variability, even among trained and professional experts.
- These methods required a great deal of time to post-process the segmentation results with other strategies, thereby abandoning the end-to-end learning.

What this study adds

- An end-to-end R-DCNN was proposed to jointly segment OD and OC for glaucoma screening.
- The proposed approach could automatically learn discriminative representations from raw input images.
- The proposed approach without post-processing achieved performance matching to ophthalmologists while exceeded other advanced methods in OD and OC segmentation, along with glaucoma screening.
- It provided a useful way to act as an efficient tool for computer-assisted glaucoma screening in clinical practice.

DATA AVAILABILITY
The export of human-related data is governed by the Ministry of Science and Technology of China (MOST) in accordance with the Regulations of the People's Republic of China on Administration of Human Genetic Resources (State Council No.717). Request for the non-profit use of the fundus images and related clinical information in the inhouse dataset should be sent to corresponding author. The joint application of the corresponding author together with the requester for the data sharing will be generated and submitted to MOST. The data will be provided to the requester after the approval from MOST. The DRISHTI-GS dataset and RIM-ONE v3 dataset are publicly available at https://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php and https://pan.baidu.com/s/1T9rvZKnKC4EF2MEY41lhpg?pwd=gy6m, respectively.

REFERENCES
1. Tham YC, Li X, Wong TY, Quigley HA, Aung T, Cheng CY. Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. Ophthalmology. 2014;121:2081–90. https://doi.org/10.1016/j.ophtha.2014.05.013.
2. Zhao RC, Chen XL, Liu XY, Chen ZL, Guo F, Li S. Direct cup-to-disc ratio estimation for glaucoma screening via. semi-supervised learning. IEEE J Biomed Health. 2020;24:1104–13. https://doi.org/10.1109/JBHI.2019.2934477.
3. Morgan JE, Sheen NJL, North RV, Choong Y, Ansari E. Digital imaging of the optic nerve head: monoscopic and stereoscopic analysis. Brit J Ophthalmol. 2005;89:879–84. https://doi.org/10.1136/bjo.2004.046169.
4. Thakur N, Juneja M. Survey on segmentation and classification approaches of optic cup and optic disc for diagnosis of glaucoma. Biomed Signal Proces. 2018;42:162–89. https://doi.org/10.1016/j.bspc.2018.01.014.
5. Almazroa A, Burman R, Raahemifar K, Lakshminarayanan V. Optic disc and optic cup segmentation methodologies for glaucoma image detection: a survey. J Ophthalmol. 2015;2015:1–28. https://doi.org/10.1155/2015/180972.
6. Mittapalli PS, Kande GB. Segmentation of optic disk and optic cup from digital fundus images for the assessment of glaucoma. Biomed Signal Proces. 2016;24:34–46. https://doi.org/10.1016/j.bspc.2015.09.003.
7. Edupuganti VG, Chawla A, Kale A. Automatic optic disk and cup segmentation of fundus images using deep learning. In: 25th IEEE International Conference on Image Processing (ICIP). 2018, pp. 2227–31. https://doi.org/10.1109/ICIP.2018.8451753.
8. Haleem MS, Han LX, van Hemert J, Li BH, Fleming A, Pasquale LR, et al. A novel adaptive deformable model for automated optic disc and cup segmentation to aid glaucoma diagnosis. J Med Syst. 2018;42:1–18. https://doi.org/10.1007/s10916-017-0859-4.
9. Jonas JB, Budde WM, Panda-Jonas S. Ophthalmoscopic evaluation of the optic nerve head. Surv Ophthalmol. 1999;43:293–320. https://doi.org/10.1016/S0039-6257(98)00049-6.
10. Cheng J, Liu J, Wing D, Wong K, Yin FS, Cheung C, et al. Automatic optic disc segmentation with peripapillary atrophy elimination. In: 33rd Annual International Conference of the IEEE Engineering-in-Medicine-and-Biology-Society (EMBS). 2011, 6224–7. https://doi.org/10.1109/IEMBS.2011.6091537.
11. Wang Y, Yu XS, Chi JN, Wu CD. Automatic segmentation of optic disc and cup in retinal fundus images using improved two-layer level set method. Math Probl Eng. 2019;2019:1–10. https://doi.org/10.1155/2019/4836296.
12. Fu HZ, Cheng J, Xu YW, Wong DWK, Liu J, Cao XC. Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Trans Med Imaging. 2018;37:1597–605. https://doi.org/10.1109/TMI.2018.2791488.
13. Bokhari STF, Sharif M, Yasmin M, Fernandes SL. Fundus image segmentation and feature extraction for the detection of glaucoma: a new approach. Curr Med Imaging. 2018;14:77–87. https://doi.org/10.2174/1573405613666170405145913.
14. Lowell J, Hunter A, Steel D, Basu A, Ryder R, Fletcher E, et al. Optic nerve head segmentation. IEEE Trans Med Imaging. 2004;23:256–64. https://doi.org/10.1109/TMI.2003.823261.
15. Chakravarty A, Sivaswamy J. Joint optic disc and cup boundary extraction from monocular fundus images. Comput Meth Prog Bio. 2017;147:51–61. https://doi.org/10.1016/j.cmpb.2017.06.004.
16. Zhang Z, Liu JA, Cherian NS, Sun Y, Lim JH, Wong WK, et al. Convex hull based neuro-retinal optic cup ellipse optimization in glaucoma diagnosis. In: Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society. 2009, pp 1441–4. https://doi.org/10.1109/IEMBS.2009.5332913.
17. Roychowdhury S, Koozekanani DD, Kuchinka SN, Parhi KK. Optic disc boundary and vessel origin segmentation of fundus images. IEEE J Biomed Health. 2016;20:1562–74. https://doi.org/10.1109/JBHI.2015.2473159.
18. Ayub J, Ahmad J, Muhammad J, Aziz L, Ayub S, Akram U, et al. Glaucoma detection through optic disc and cup segmentation using K-mean clustering. In: International Conference on Computing, Electronic and Electrical Engineering (ICE Cube). 2016, pp 143–7. https://doi.org/10.1109/ICECUBE.2016.7495212.
19. Jiang YM, Xia H, Xu YW, Cheng J, Fu HZ, Duan LX.et al. Optic disc and cup segmentation with blood vessel removal from fundus images for glaucoma detection. In: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2018;2018:862–5. https://doi.org/10.1109/EMBC.2018.8512400.
20. Wang ZY, Dong NQ, Rosario SD, Xu M, Xie PT, Xing EP. Ellipse detection of optic disc-and-cup boundary in fundus images. In: 16th IEEE International Symposium on Biomedical Imaging (ISBI). 2019, pp. 601–4. https://doi.org/10.1109/ISBI.2019.8759173.
21. Sedai S, Roy PK, Mahapatra D, Garnavi R. Segmentation of optic disc and optic cup in retinal fundus images using shape regression. In: 38th Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (EMBC). 2016, pp. 3260–4. https://doi.org/10.1109/EMBC.2016.7591424.
22. Sigut J, Nunez O, Fumero F, Gonzalez M, Arnay R. Contrast based circular approximation for accurate and robust optic disc segmentation in retinal images. Peerj. 2017;5:1–18. https://doi.org/10.7717/peerj.3763.
23. Zhao S, Gao ZF, Zhang HY, Xie YQ, Luo JW, Ghista D, et al. Robust segmentation of intima-media borders with different morphologies and dynamics during the cardiac cycle. IEEE J Biomed Health. 2018;22:1571–82. https://doi.org/10.1109/JBHI.2017.2776246.
24. Zhou ZH, Dai M, Wang TL, Zhao RZ. Prior distribution-based statistical active contour model. Multimed Tools Appl. 2019;78:35813–33. https://doi.org/10.1007/s11042-019-08101-2.

25. Joshi GD, Sivaswamy J, Krishnadas SR. Depth discontinuity-based cup segmentation from multiview color retinal images. IEEE Trans Bio Med Eng. 2012;59:1523–31. https://doi.org/10.1109/TBME.2012.2187293.

26. Miri MS, Lee K, Niemeijer M, Abràmoff MD, Kwon YH, Garvin MK. Multimodal segmentation of optic disc and cup from stereo fundus and SD-OCT images. Medical Imaging. 2013;8669. https://doi.org/10.1117/12.2007010.

27. Gao Y, Yu XS, Wu CD, Zhou W, Wang XN, Zhuang YM. Accurate optic disc and cup segmentation from retinal images using a multi-feature based approach for glaucoma assessment. Symmetry Basel. 2019;11:1–18.

28. Xu YW, Duan LX, Lin S, Chen XY, Wang DWK, Wong TY, et al. Optic cup segmentation for glaucoma detection using low-rank superpixel representation. In: 17th International Conference on Medical Image Computing and Computer-AssistedIntervention (MICCAI). 2014;8673:788–95. https://doi.org/10.1007/978-3-319-10404-1_98.

29. Tan NM, Xu YW, Goh WB, Liu J. Robust multi-scale superpixel classification for optic cup localization. Comput Med Imag Grap. 2015;40:182–93. https://doi.org/10.1016/j.compmedimag.2014.10.002.

30. Cheng J, Liu J, Xu YW, Yin FS, Wong DWK, Tan NM, et al. Superpixel classification based optic disc and optic cup segmentation for glaucoma screening. IEEE Trans Med Imaging. 2013;32:1019–32. https://doi.org/10.1109/TMI.2013.2247770.

31. Cheng J, Tao DC, Wong DWK, Liu J. Quadratic divergence regularized SVM for optic disc segmentation. Biomed Opt Express. 2017;8:2687–96. https://doi.org/10.1364/BOE.8.002687.

32. Qin PZ, Wang LY, Lv HB. Optic disc and cup segmentation based on deep learning. In: IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). 2019, pp. 1835-40. https://doi.org/10.1109/itnec.2019.8729455.

33. Shelhamer E, Long J, Darrell T. Fully convolutional networks for semantic segmentation. IEEE Trans Pattern Anal. 2017;39:640–51. https://doi.org/10.1109/TPAMI.2016.2572683.

34. Shankaranarayana SM, Ram K, Mitra K, Sivaprakasam M. Joint optic disc and cup segmentation using fully convolutional and adversarial networks. In: International Workshop on Fetal and Infant Image Analysis (FIFI) / 4th International Workshop on Ophthalmic Medical Image Analysis (OMIA). 2017;10554:168–76. https://doi.org/10.1007/978-3-319-67561-9_19.

35. Ronneberger O, Fischer P.Brox T, U-Net: Convolutional networks for biomedical image segmentation. In: Medical Image Computing and Computer Assisted Intervention Society (MICCAI). 2015;9351:234–41. https://doi.org/10.1007/978-3-319-24574-4_28.

36. Sevastopolsky A. Optic disc and cup segmentation methods for glaucoma detection with modification of U-Net convolutional neural network. Pattern Recognit Image Anal. 2017;27:618–24. https://doi.org/10.1134/S1054661817030269.

37. Tabassum M, Khan TM, Arsalan M, Naqvi SS, Ahmed M, Ahmed H, et al. CDED-Net: joint segmentation of optic disc and optic cup for glaucoma screening. IEEE Access. 2020;8:102733–47. https://doi.org/10.1109/ACCESS.2020.2998635.

38. Zuiderveld KJ. Contrast limited adaptive histogram equalization. Graphics Gems. 1994, pp. 474–85. https://doi.org/10.1016/B978-0-12-336156-1.50061-6.

39. Zou BJ, Chen CL, Zhu CZ, Duan XC, Chen ZL. Classified optic disc localization algorithm based on verification model. Comput Graph UK. 2018;70:281–7. https://doi.org/10.1016/j.cag.2017.07.031.

40. Gu ZW, Cheng J, Fu HZ, Zhou K, Hao HY, Zhao YT, et al. CE-Net: context encoder network for 2D medical image segmentation. IEEE Trans Med Imaging. 2019;38:2281–92. https://doi.org/10.1109/TMI.2019.2903562.

41. Ali R, Sheng B, Li P, Chen Y, Li HT, Yang P, et al. Optic disc and cup segmentation through fuzzy broad learning system for glaucoma screening. IEEE Trans Ind Inf. 2021;17:2476–87. https://doi.org/10.1109/TII.2020.3000204.

42. Goodfellow IJ, Pouget-Abadie J, Mirza M, Xu B, Warder-Farley D, Ozair S, et al. Generative adversarial nets. In: 28th Conference on Neural Information Processing Systems (NIPS). 2014;27:2672–80. https://doi.org/10.1145/3422622.

43. Zhao TC, Zhao R, Eskenazi M. Learning discourse-level diversity for neural dialog models using conditional variational autoencoders. In: 55th Annual Meeting of the Association-for-Computational-Linguistics (ACL). 2017;1:654–64. https://doi.org/10.18653/v1/P17-1061.

AUTHOR CONTRIBUTIONS
FL, and LJZ have made contributions to the writing. WJX and WZP implemented the deep learning algorithm. HDZ and MSJ acquired the data. XDZ analyzed and interpreted the data. All authors were involved in the study design, and approval of the final manuscript.

COMPETING INTERESTS
The authors declare no competing interests.

ADDITIONAL INFORMATION
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41433-022-02055-w.

Correspondence and requests for materials should be addressed to Minshan Jiang.

Reprints and permission information is available at http://www.nature.com/reprints

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