

An intelligent segmentation method for leakage points in central serous chorioretinopathy based on fluorescein angiography images

Jian-Guo Xu¹, Yong-Chi Liu¹, Fen Zhou², Jian-Xin Shen¹, Zhi-Peng Yan², Xin-Ya Hu³, Wei-Hua Yang³

¹College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, Jiangsu Province, China

²The Affiliated Eye Hospital of Nanjing Medical University, Nanjing 210029, Jiangsu Province, China

³Shenzhen Eye Hospital, Shenzhen Eye Medical Center, Southern Medical University, Shenzhen 518040, Guangdong Province, China

Correspondence to: Jian-Xin Shen. College of Mechanical & Electrical Engineering, Nanjing University of Aeronautics & Astronautics, Nanjing 210016, Jiangsu Province, China. cadatc@nuaa.edu.cn. Xin-Ya Hu and Wei-Hua Yang. Shenzhen Eye Hospital, Shenzhen Eye Medical Center, Southern Medical University, Shenzhen 518040, Guangdong Province, China. huxinya@sz-eyes.com; benben0606@139.com

Received: 2025-02-27 Accepted: 2025-08-26

Abstract

• **AIM:** To construct an intelligent segmentation scheme for precise localization of central serous chorioretinopathy (CSC) leakage points, thereby enabling ophthalmologists to deliver accurate laser treatment without navigational laser equipment.

• **METHODS:** A dataset with dual labels (point-level and pixel-level) was first established based on fundus fluorescein angiography (FFA) images of CSC and subsequently divided into training (102 images), validation (40 images), and test (40 images) datasets. An intelligent segmentation method was then developed, based on the You Only Look Once version 8 Pose Estimation (YOLOv8-Pose) model and segment anything model (SAM), to segment CSC leakage points. Next, the YOLOv8-Pose model was trained for 200 epochs, and the best-performing model was selected to form the optimal combination with SAM. Additionally, the classic five types of U-Net series models [*i.e.*, U-Net, recurrent residual U-Net (R2U-Net), attention U-Net (AttU-Net), recurrent residual attention U-Net (R2AttU-Net), and nested U-Net (UNet++)] were initialized with three

random seeds and trained for 200 epochs, resulting in a total of 15 baseline models for comparison. Finally, based on the metrics including Dice similarity coefficient (DICE), intersection over union (IoU), precision, recall, precision-recall (PR) curve, and receiver operating characteristic (ROC) curve, the proposed method was compared with baseline models through quantitative and qualitative experiments for leakage point segmentation, thereby demonstrating its effectiveness.

• **RESULTS:** With the increase of training epochs, the mAP50-95, Recall, and precision of the YOLOv8-Pose model showed a significant increase and tended to stabilize, and it achieved a preliminary localization success rate of 90% (*i.e.*, 36 images) for CSC leakage points in 40 test images. Using manually expert-annotated pixel-level labels as the ground truth, the proposed method achieved outcomes with a DICE of 57.13%, an IoU of 45.31%, a precision of 45.91%, a recall of 93.57%, an area under the PR curve (AUC-PR) of 0.78 and an area under the ROC curve (AUC-ROC) of 0.97, which enables more accurate segmentation of CSC leakage points.

• **CONCLUSION:** By combining the precise localization capability of the YOLOv8-Pose model with the robust and flexible segmentation ability of SAM, the proposed method not only demonstrates the effectiveness of the YOLOv8-Pose model in detecting keypoint coordinates of CSC leakage points from the perspective of application innovation but also establishes a novel approach for accurate segmentation of CSC leakage points through the “detect-then-segment” strategy, thereby providing a potential auxiliary means for the automatic and precise real-time localization of leakage points during traditional laser photocoagulation for CSC.

• **KEYWORDS:** You Only Look Once version 8 Pose Estimation; segment anything model; central serous chorioretinopathy; leakage point; segmentation

DOI:10.18240/ijo.2026.03.01

Citation: Xu JG, Liu YC, Zhou F, Shen JX, Yan ZP, Hu XY, Yang WH. An intelligent segmentation method for leakage points in central serous chorioretinopathy based on fluorescein angiography images. *Int J Ophthalmol* 2026;19(3):421-433

INTRODUCTION

Central serous chorioretinopathy (CSC) is a typical macular disease characterized by decreased visual acuity, central scotoma, and metamorphopsia. Currently, the primary treatment options for CSC include pharmacological therapy and laser therapy^[1-3]. Laser therapy encompasses traditional laser photocoagulation and micropulse laser photocoagulation. Specifically, traditional laser photocoagulation targets the leakage points within the edematous area of CSC, using laser ablation to seal the leakage points, thereby promoting the absorption of serous fluid and accelerating visual recovery. However, the segmentation of leakage points in fundus fluorescein angiography (FFA) images mainly depends on physicians' manual work. This process is time-consuming and prone to human error and poor reproducibility. Therefore, there is an urgent requirement for an automatic and accurate method to segment CSC leakage points from the background.

The rapid advancement of artificial intelligence (AI) technology has provided precise and efficient solutions for medical image processing. By leveraging AI to replace physicians' manual image analysis, medical resources are conserved, and the results are standardized, significantly reducing human errors. The successful application of convolutional neural networks in natural image processing has laid a solid foundation for their widespread use in medical image analysis. Today, convolutional neural networks have been extensively applied in various fields, including ophthalmology^[4-8], orthopedics^[9], and dentistry^[10]. In ophthalmology, these networks are leveraged for tasks such as lesion detection^[11-12], lesion segmentation^[13-14], and disease classification^[15-16]. According to recent reviews on CSC research^[17], deep learning techniques have primarily been applied to the segmentation of CSC edematous regions^[18-19], surgical guidance^[20-21], and postoperative efficacy evaluation^[22-23]. However, there is limited research on deep learning to segment CSC leakage points^[24]. Despite this, the extensive application and remarkable success of deep learning in lesion localization for other retinal diseases provide valuable insights for exploring CSC leakage point segmentation. For instance, Dai *et al*^[25] proposed a network named DeepDR for detecting lesions such as microaneurysms, cotton wool spots, hard exudates, and hemorrhages in diabetic retinopathy. Similarly, Tang *et al*^[26] introduced a supervised contrastive learning-based framework to optimize the detection of hard exudates in diabetic retinopathy, incorporating a discriminative edge detection module to enhance the accuracy of lesion

boundary identification. While deep learning has achieved significant success in tasks such as lesion detection and segmentation in diabetic retinopathy, these methods primarily target specific lesions like microaneurysms, cotton wool spots, and hard exudates. Consequently, they cannot be directly applied to the segmentation scenario of CSC leakage points, and a specialized approach tailored to the segmentation of CSC leakage points still needs further exploration.

In current research on retinal lesion segmentation, the integration of object detection algorithms with segmentation algorithms has been rarely explored. However, object detection algorithms can enhance segmentation efficiency by narrowing the scope, rapidly localizing lesions, improving accuracy, and reducing computational costs. Therefore, driven by this inspiration, a "detect-then-segment" based intelligent segmentation method for CSC leakage points is proposed by this study to address the challenges faced in manually locating CSC leakage points in current clinical practice. The rapid target detection capability of the You Only Look Once version 8 Pose Estimation (YOLOv8-Pose)^[27] model with the precise segmentation ability of the segment anything model (SAM)^[28] are combined to achieve accurate segmentation of CSC leakage points.

The YOLOv8 model used in this study was released by the Ultralytics team in 2023. It is a versatile model capable of performing object detection, image segmentation, image classification, and keypoint detection. The YOLOv8-Pose model, a branch of YOLOv8 specifically designed for keypoint detection tasks, achieves real-time keypoint detection while maintaining high accuracy. This model has been widely applied in various fields, including sports analysis, security surveillance, hand motion detection, and industrial inspection. For example, it has been applied for detecting and alerting elderly fall incidents^[29], license plate recognition in wide-angle scenarios^[30], correcting image tilt and distortion in industrial instruments^[31], and identifying keypoints and measuring the volume of pigs^[32]. These examples demonstrate that the YOLOv8-Pose model excels in human joint detection and tasks involving license plates, industrial instruments, and animal limb keypoints. Its efficiency, flexibility, and strong scalability make it a promising candidate for target detection tasks in medical imaging, offering new insights for the precise localization of lesions in retinal diseases.

The other model employed in this study is SAM, released by Meta AI in 2023. SAM is a versatile model capable of segmenting any given image. Since its introduction, it has been widely adopted in the medical field for various tasks, including the segmentation of multiple abdominal organs^[33], tumor segmentation across multiple imaging stages^[34], the separation of tumors from normal tissues, and nuclear segmentation^[35], polyp

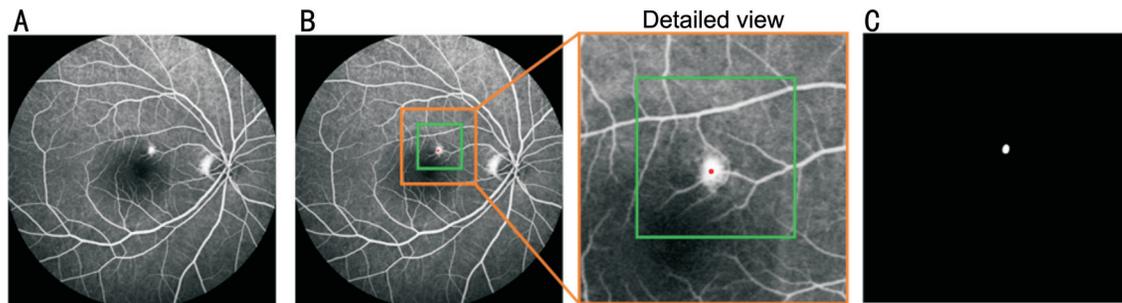


Figure 1 Images and labels of the dataset A: An FFA image of CSC; B: A point-level label for training the YOLOv8-Pose model, accompanied by a detailed view; C: A pixel-level label serves as the ground truth. CSC: Central serous chorioretinopathy; FFA: Fundus fluorescein angiography; YOLOv8-Pose: You Only Look Once version 8 Pose Estimation.

segmentation^[36], and skin cancer segmentation^[37]. Additionally, SAM has demonstrated high-performance segmentation even with limited medical data^[38]. These examples highlight SAM's significant advantages in medical image segmentation tasks, owing to its powerful generalization capabilities.

Given the unique advantages and successful applications of the YOLOv8-Pose model and SAM in the fields of natural image and medical image processing, this study attempts to construct an intelligent segmentation method for CSC leakage points supported by the “detect-then-segment” strategy, driven jointly by the two aforementioned models, enabling precise localization of CSC leakage points to assist ophthalmologists in accurately projecting laser treatment onto leakage points even without navigational laser equipment. The proposed method first uses the YOLOv8-Pose model to detect the keypoint coordinates of CSC leakage points and then employs SAM to accurately segment them. Specifically, the coordinates detected by the YOLOv8-Pose model provide precise prior guidance for SAM, enabling SAM to achieve accurate segmentation of CSC leakage points without additional training.

PARTICIPANTS AND METHODS

Ethical Approval This research received approval from the Ethics Committee of the Shenzhen Eye Hospital (the approval number: 2025KYPJ013) and the Medical Ethics Committee of the Affiliated Eye Hospital of Nanjing Medical University (the approval number: 2021006) and adhered to the principles outlined in the Declaration of Helsinki.

Acquisition of Data The Shenzhen Eye Hospital and the Affiliated Eye Hospital of Nanjing Medical University provided the dataset used in this experiment. These images were captured using a fundus camera after FFA was performed on patients with CSC. All fundus images were anonymized by removing patient information and then annotated by experts, serving as the ground truth. After excluding images with issues such as overexposure, eye movement, poor contrast, excessive blurring, or excessive diffusion of the contrast agent, 182 FFA images suitable for the experiment were obtained, forming the final dataset.

Preprocessing of Dataset

Label preparation To train the YOLOv8-Pose and baseline models in this research, point-level and pixel-level labels are prepared. Point-level labels are used to train the YOLOv8-Pose model. Each label uses a bounding box to enclose each leakage point, with a pixel inside the leakage area designated as a keypoint to indicate the rough position of the corresponding leakage point. Pixel-level labels are used as the ground truth to evaluate the leakage point segmentation performance of the proposed method and baseline models. Examples of the FFA images and the corresponding labels are shown below: Figure 1A displays an FFA image of CSC, where the white area in the center is a CSC leakage point. Figure 1B illustrates a point-level label used for training the YOLOv8-Pose model. The green bounding box encompasses the leakage point and its surrounding region, with the red keypoint indicating the rough position of the leakage point. Each leakage point is uniquely associated with a bounding box and a keypoint, and the green bounding box and the red keypoint are enlarged for better visibility. Figure 1C showcases the pixel-level label in which experts meticulously annotated the leakage point. The FFA images and corresponding labels are randomly divided into 102 images for training, 40 for validation, and 40 for test. The point-level and pixel-level labels are used for training the YOLOv8-Pose and baseline models, respectively.

When preparing labels for training the YOLOv8-Pose model, each leakage point is annotated with a bounding box and a keypoint in labelme software. The annotated data is saved in a JSON file, where the top-left and bottom-right corners of the bounding box and the keypoint coordinates are stored in key-value pairs. The original format of the annotated coordinates in the JSON file is shown in Figure 2A. Here, the red point marks the keypoint, indicating the rough position of the leakage point, while the green box outlines the bounding box required by the YOLOv8-Pose model. To facilitate the input of these labels into the YOLOv8-Pose model, the original coordinates are first converted into the center-format coordinates, as illustrated in Figure 2B. In this format, the yellow point depicts the center

position of the bounding box, with coordinates (x_c, y_c) , and the red point highlights the keypoint with coordinates (x_{kpt}, y_{kpt}) . Additionally, w and h denote the width and height of the bounding box, respectively.

Using an automatic script, the coordinates in the JSON files are extracted to obtain the original-format coordinates, denoted as $(x_1, y_1, x_2, y_2, x_{kpt}, y_{kpt})$, where (x_1, y_1) and (x_2, y_2) indicate the coordinates of the top-left and bottom-right corners of the bounding box, respectively, and (x_{kpt}, y_{kpt}) specifies the keypoint coordinates. These original-format coordinates are then converted into center-format coordinates $(x_c, y_c, w, h, x_{kpt}, y_{kpt})$ using Formula 1.

$$\begin{cases} x_c = (x_1 + x_2) / 2 \\ y_c = (y_1 + y_2) / 2 \\ w = x_2 - x_1 \\ h = y_2 - y_1 \end{cases} \quad (1)$$

The center-format coordinates are further transformed into YOLO label format coordinates $(0, n_x, n_y, n_w, n_h, n_x_{kpt}, n_y_{kpt})$ using Formula 2. Here, the prefix “n” denotes normalization, and $n_x, n_y, n_w, n_h, n_x_{kpt}$ and n_y_{kpt} represent the normalized x and y coordinates of the center point, the normalized width and height of the bounding box, and the normalized x and y coordinates of the keypoint, respectively. Additionally, w_{pic} and h_{pic} denote the width and height of the image. The value 0 signifies the class label, and all normalized values fall from 0 to 1.

$$\begin{cases} n_x = x_c / w_{pic} \\ n_y = y_c / h_{pic} \\ n_w = w / w_{pic} \\ n_h = h / h_{pic} \\ n_x_{kpt} = x_{kpt} / w_{pic} \\ n_y_{kpt} = y_{kpt} / h_{pic} \end{cases} \quad (2)$$

Data augmentation After randomly dividing the dataset into 102 training images, 40 validation images, and 40 test images, data augmentation is performed on the training and validation dataset images and their corresponding labels to train the YOLOv8-Pose model and the baseline models. The augmentation techniques included translation and rotation, as described in Formula 3. The center of rotation is set to the image center, and the rotation is implemented according to Formula 4. In the formulas, T stands for the translation operation, and (x, y) denotes the vector for image translation. Similarly, R denotes the rotation operation around the image center, with θ representing the rotation angle. Here, (x', y') marks the coordinates of a pixel point after rotation, (x, y) describes the coordinates of the pixel point before rotation, (x_{pc}, y_{pc}) indicates the coordinates of the image’s center position.

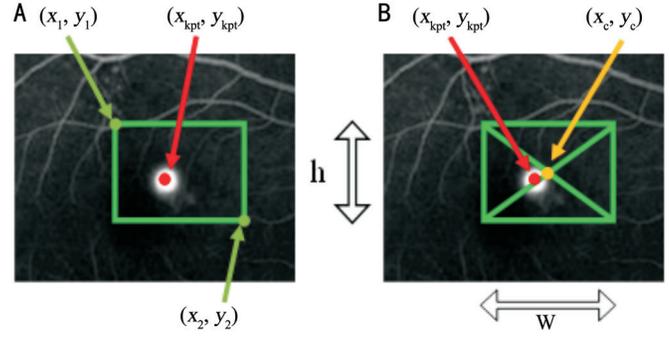


Figure 2 Coordinate format transformation A: The original format includes the coordinates of the bounding box (defined by its top-left and bottom-right corners) and the keypoint coordinates; B: The center format contains the coordinates of the bounding box (defined by its center point, width, and height) and the keypoint coordinates.

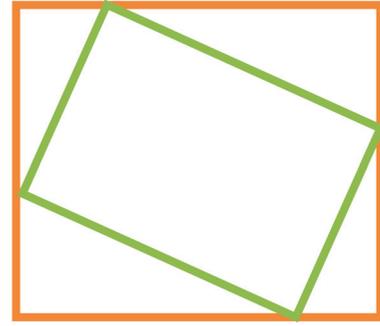


Figure 3 Construction of the new bounding box the solid green box is the rotated bounding box, and the orange box is the minimum axis-aligned bounding box, which serves as the new bounding box.

$$\begin{cases} T(x, y) \\ R(\theta) \end{cases} \quad (3)$$

$$\begin{cases} x' = (x - x_{pc}) \cdot \cos(\theta) - (y - y_{pc}) \cdot \sin(\theta) + x_{pc} \\ y' = (x - x_{pc}) \cdot \sin(\theta) + (y - y_{pc}) \cdot \cos(\theta) + y_{pc} \end{cases} \quad (4)$$

It is important to note that the YOLOv8-Pose model always requires the bounding box to be an Axis-Aligned Rectangle. After rotation, the sides of the new bounding box must remain aligned with the x and y axes, and the keypoint must be enclosed. To ensure this, a minimum bounding rectangle is calculated from the rotated bounding box and used as the new bounding box. This process is illustrated in Figure 3.

After rotation, the coordinates of the four corners of the bounding box are updated according to Formula 5. The minimum axis-aligned bounding box of the rotated rectangle is then determined using Formula 6. Here, (x_i, y_i) and (x'_i, y'_i) specify the coordinates of the four corners of the bounding box before and after rotation, respectively, where i ranges from 1 to 4. Additionally, (x_{\min}, y_{\min}) and (x_{\max}, y_{\max}) denote the coordinates of the top-left and bottom-right corners of the new bounding box.

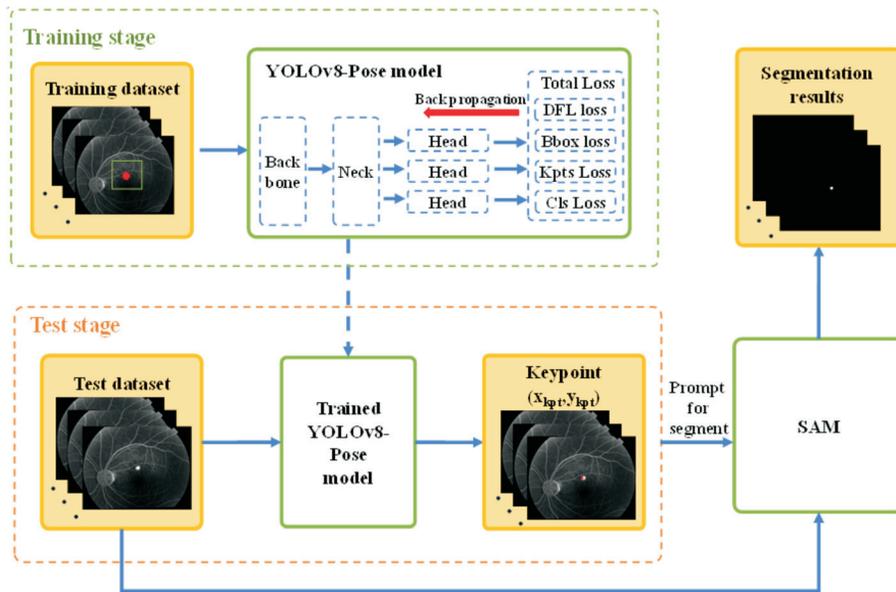


Figure 4 Overview of the proposed method SAM: Segment anything model; YOLOv8-Pose: You Only Look Once version 8 Pose Estimation.

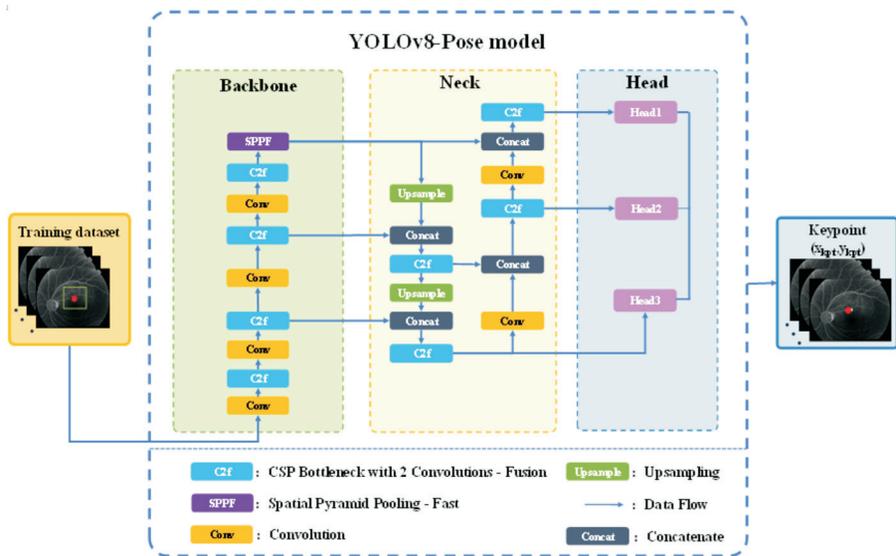


Figure 5 Workflow of the YOLOv8-Pose model^[27] YOLOv8-Pose: You Only Look Once version 8 Pose Estimation.

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_i - x_{pc} \\ y_i - y_{pc} \end{bmatrix} + \begin{bmatrix} x_{pc} \\ y_{pc} \end{bmatrix} \quad (5)$$

$$\begin{cases} x_{\min} = \min(x'_1, x'_2, x'_3, x'_4) \\ y_{\min} = \min(y'_1, y'_2, y'_3, y'_4) \\ x_{\max} = \max(x'_1, x'_2, x'_3, x'_4) \\ y_{\max} = \max(y'_1, y'_2, y'_3, y'_4) \end{cases} \quad (6)$$

The data augmentation method for the ground truth labels is identical to that applied to the FFA images and will not be reiterated here. Using the aforementioned data augmentation and coordinate transformation techniques, we augmented the two types of labels and their corresponding FFA images in the training and validation dataset, thereby preparing the data for subsequent model training. After augmentation, the training and validation datasets contain 2652 and 400 images, respectively, along with their corresponding two types of labels.

Implementation details As shown in Figure 4, this is a schematic diagram of the method proposed in this research. Deep learning techniques are adopted in this work, aiming to achieve accurate localization of these CSC leakage points. We trained and tested the YOLOv8-Pose model, utilizing it to preliminarily identify the rough position of the leakage points. These pixel coordinates of the leakage points are then used as prompts, which, along with the corresponding FFA images, are input into SAM for prompt-based segmentation. This approach ultimately enables the intelligent segmentation of CSC leakage points. The input-output workflow of the YOLOv8-Pose model is illustrated in Figure 5.

The YOLOv8-Pose model, a branch of the YOLOv8 model, is trained by the point-level labels, and constructs its loss function based on the ground-truth labels of the keypoints, bounding boxes, and other relevant labels, as well as its predicted labels,

to facilitate the update of its parameters. The model outputs include keypoint coordinates and bounding box coordinates. However, only the keypoint coordinates are utilized in this study. The model’s loss function consists of four components: the Bounding Box Loss L_{box} , the Keypoint Distance Loss L_{pose} , the Classification Loss L_{cls} , and the Distributive Focal Loss L_{dff} . The total loss is calculated as shown in Formula 7, where λ refers to the loss gain coefficients used to balance the contributions of four components. The default values for these gain coefficients are applied in this experiment: the Bounding Box Loss Gain λ_{box} is 7.5, the Class Loss Gain λ_{cls} is 0.5, the Pose Loss Gain λ_{pose} is 12, and the DFL Gain λ_{dff} is 1.5. The Bounding Box Loss L_{box} measures the difference between the predicted bounding box and the ground truth bounding box using the intersection over union (IoU) loss, optimizing the position of the bounding box. The Classification Loss L_{cls} calculates the difference between the predicted and ground truth classes using the binary cross-entropy loss function. Using the Euclidean distance, the Keypoint Distance Loss L_{pose} evaluates the discrepancy between the predicted and ground truth keypoints. The Distributive Focal Loss L_{dff} employs a cross-entropy-based approach to regress the predicted bounding box’s position to the probability distributions of the ground truth bounding box’s four boundaries (top, bottom, left, right). By leveraging these probability distributions, the model enhances its ability to localize targets. This innovative design allows the model to handle uncertainties more effectively in bounding box regression tasks, accurately capturing the boundary information of target objects.

$$L_{total} = \lambda_{box} \cdot L_{box} + \lambda_{cls} \cdot L_{cls} + \lambda_{pose} \cdot L_{pose} + \lambda_{dff} \cdot L_{dff} \quad (7)$$

The input-output workflow of SAM is presented in Figure 6. The keypoint coordinates predicted by the YOLOv8-Pose model and corresponding FFA images are loaded into SAM. These inputs are processed separately by the image encoder and the prompt encoder. Then, the processed features are combined in the mask decoder to generate the final segmentation results of the CSC leakage points.

SAM consists of three main parts: an image encoder, a prompt encoder, and a mask decoder. Further information can be found in this literature^[28]. The image encoder is based on the Vision Transformer architecture and utilizes a pre-trained Masked Autoencoder model. This model processes the input images through downsampling to generate image embeddings, which are foundational features for subsequent tasks. The prompt encoder transforms user-provided prompts of various forms into embeddings. These embeddings are then integrated with the image embeddings through self-attention and cross-attention mechanisms, producing more accurate feature representations. The mask decoder, a bidirectional

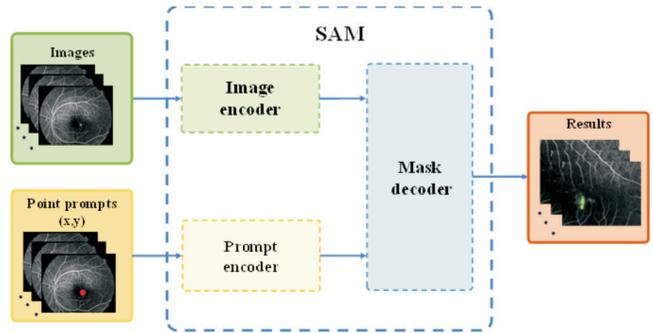


Figure 6 Prompt-driven workflow of SAM SAM: Segment anything model.

transformer decoder, takes the combined embeddings from the two encoders as input and yields the final segmentation mask. In general, the coordinates of the keypoint served as a prompt, as well as the corresponding FFA image, are utilized by SAM. After processing through the three modules described above, the segmentation results of the CSC leakage points are obtained.

The proposed method and the comparative methods are implemented on a Windows system using the PyTorch deep learning framework, with Python as the programming language implemented in the PyCharm IDE. The RTX 4080 GPU is employed to accelerate the training process. During the training stage, the YOLOv8-Pose model is trained for 200 epochs with a batch size of 16 and a default learning rate of 0.01. As mentioned earlier, the default values for the loss gains are used, and the keypoint shape parameter is set to (1, 2). Training for the baseline U-Net series models is conducted for 200 epochs with a batch size of 32 and a default learning rate of 0.001.

Evaluation metrics To evaluate the performance of the YOLOv8-Pose model in detecting the rough positions of leakage points in FFA images of the CSC, mean Average Precision at thresholds from 0.50 to 0.95 (mAP50-95), Recall, and Precision are used. To evaluate the final segmentation results of the models for the CSC leakage point segmentation task, we conducted a comprehensive performance comparison between the proposed method and baseline models using six metrics: area under the curve of receiver operating characteristic (AUC-ROC), AUC-precision recall (PR), IoU, Dice similarity coefficient (DICE), Precision, and Recall.

The mAP50-95 is defined by the mean Average Precision calculated over different thresholds ranging from 50% to 95%, as shown in Formula 8. A higher mAP50-95 value indicates that the model can accurately localize keypoints across different thresholds. Precision, which measures the proportion of correctly detected keypoints among all detected keypoints and reflects fewer false positives when higher, is defined as shown in Formula 9. Recall, which evaluates the proportion of correctly detected keypoints among all true

keypoints and demonstrates fewer missed true keypoints when higher, is defined as shown in Formula 10. The ROC curve reflects the model's ability to distinguish between positive and negative samples under different classification thresholds. The horizontal and vertical axes of the ROC curve stand for the False Positive Rate and True Positive Rate, respectively. Given that True Positive Rate is equivalent to Recall, the formula for True Positive Rate is not reiterated here, whereas the calculation of False Positive Rate is shown in Formula 11. The AUC-ROC measures the ability of a model to distinguish between positive and negative samples across all classification thresholds, with a higher AUC-ROC value indicating better classification performance. Additionally, given that CSC leakage points occupy a relatively small portion of the entire image, the PR curve, which is more suitable for imbalanced segmentation tasks, is adopted to characterize the model's classification performance under different thresholds. The PR curve uses Recall and Precision as its horizontal and vertical axes. DICE measures the similarity between the predicted segmentation results and the ground truth by focusing on the harmonic mean of precision and recall. The formula for DICE is shown in Formula 12, where X and Y define the sets of foreground pixels in the predicted segmentation results and the ground truth, respectively, and $|X \cap Y|$ denotes the overlapping region between the two sets, and $|X| + |Y|$ represents the total number of pixels in both sets combined. IoU indicates how well the predicted region aligns with the target by measuring the overlap between the predicted and ground truth regions. The formula for IoU is shown in Formula 13, where $|X \cup Y|$ is the total number of unique pixels in either the predicted or ground truth sets. Precision measures the accuracy of the model's predictions by calculating the proportion of correctly predicted positive pixels out of all pixels predicted as positive, reflecting the model's ability to avoid false alarms. Recall evaluates the completeness of the model's detection by calculating the proportion of correctly predicted positive pixels out of all actual positive pixels, indicating the model's ability to minimize missed detections. The formulas for recall and precision have been provided earlier.

$$mAP_{50-95} = \frac{1}{10} \sum_{k=1}^{10} AP_{IoU=0.5+0.05 \times (k-1)} \quad (8)$$

$$Precision = \frac{True\ Positives}{True\ Positives + False\ Positives} \quad (9)$$

$$Recall = \frac{True\ Positives}{True\ Positives + False\ Negatives} \quad (10)$$

$$FPR = \frac{False\ Positives}{False\ Positives + True\ Negatives} \quad (11)$$

$$DICE = \frac{2 \times |X \cap Y|}{|X| + |Y|} \quad (12)$$

Table 1 Abbreviations and full names of the baseline models

Abbreviation	Full name
U-Net ^[39]	U-Net
R2U-Net ^[40]	Recurrent residual U-Net
AttU-Net ^[41]	Attention U-Net
UNet++ ^[42]	Nested U-Net
R2AttU-Net ^[43]	Recurrent residual attention U-Net

$$IoU = \frac{|X \cap Y|}{|X \cup Y|} \quad (13)$$

Abbreviations of baseline models The abbreviations and corresponding full names of the five baseline models are listed in Table 1^[39-43], and these abbreviations are used throughout this paper.

RESULTS

This study employed the “detect-then-segment” strategy to perform an accurate segmentation of the leakage points in CSC, leveraging the advancement of both YOLOv8-Pose model and SAM. The experiment was set with 200 epochs for both YOLOv8-Pose model and baseline models. The experimental results are shown in Figures 7–12 and Table 2 shows the metrics comparing our method with the baseline models.

1) Figure 7 illustrates the training loss curve of the YOLOv8-Pose model, including both total loss and component loss, reflecting the loss variation of the model during the training stage. 2) Figure 8 demonstrates the keypoint detection performance of the YOLOv8-Pose model during the training stage, showing the model's ability to accurately detect the leakage points. 3) Figure 9 presents the ROC curve performance between our method and baseline models during the testing stage, illustrating the classification capability of the model for positive and negative samples. 4) Figure 10 displays the PR curve comparison of our method and baseline models during the testing stage, highlighting the model's ability to identify positive samples in case of sparse positive samples. 5) Figure 11 depicts the keypoints detected by the YOLOv8-Pose model during the testing stage, visually indicating the model's ability of precisely detecting the leakage points. 6) Figure 12 visualizes the final segmentation results of both the proposed method and baseline models during the testing stage, intuitively illustrating the superiority of the proposed method in the segmentation of leakage points in CSC. 7) Table 2 provides a quantitative comparison between the proposed method and the baseline models in terms of of DICE, IoU, precision, and recall, directly presenting the superior performance of the proposed method over the baseline models in segmenting leakage points in CSC during the testing stage.

DISCUSSION

Overall, the main contributions of this research are as

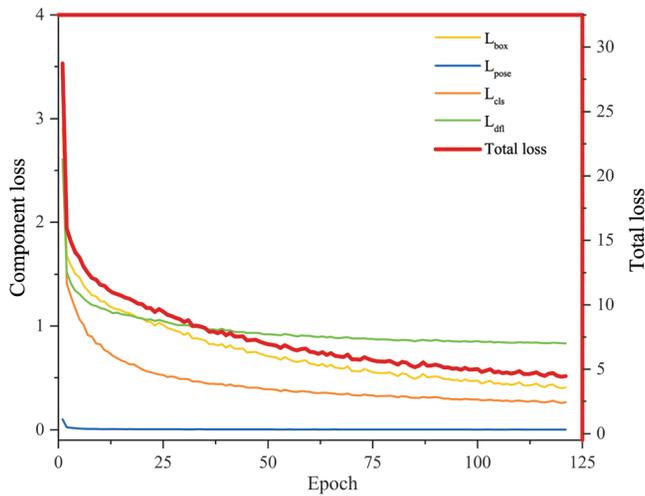


Figure 7 Training loss curves of the YOLOv8-Pose model: total loss and component loss YOLOv8-Pose: You Only Look Once version 8 Pose Estimation.

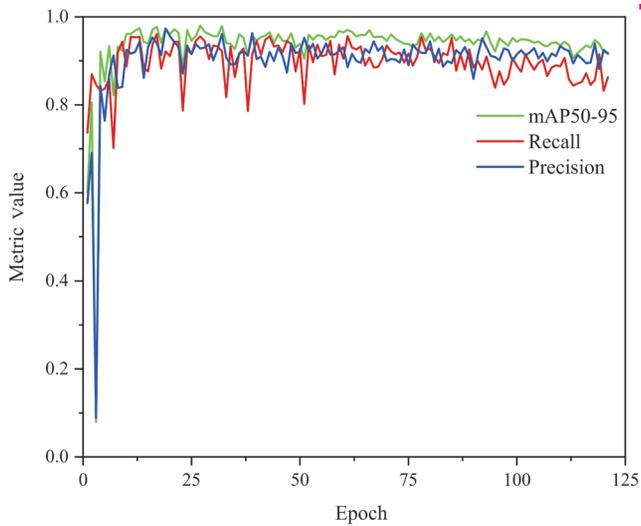


Figure 8 Keypoint detection performance: mAP50-95, recall, and precision mAP: Mean average precision.

follows: 1) A dedicated dataset, including FFA images and corresponding point-level and pixel-level labels for CSC leakage points, is constructed in this research, providing a robust foundation for the segmentation of CSC leakage points. 2) This study is the first to explore the effectiveness of combining the YOLOv8-Pose model with SAM for segmenting CSC leakage points, innovatively introducing the YOLOv8-Pose model into retinal lesion segmentation and substantially reducing computational costs. 3) The proposed method significantly improves the segmentation accuracy of CSC leakage points by adopting a “detect-then-segment” strategy, which leverages the advantages of the YOLOv8-Pose model and SAM. It expands the research scope for CSC leakage points segmentation and provides valuable insights for segmenting lesions in other retinal diseases. 4) Numerous experiments have been conducted to evaluate the performance of the YOLOv8-Pose model and SAM, in which five U-Net

Table 2 Performance comparison of DICE, IoU, precision, and recall: proposed method vs baseline models

Model	Seed	DICE (%)	IoU (%)	Precision (%)	Recall (%)
AttU-Net	1	14.20	10.61	30.20	16.67
	2	5.00	2.99	24.74	3.45
	3	8.60	6.38	16.50	11.35
	Average	9.27	6.66	23.81	10.49
UNet++	1	14.97	11.51	31.43	18.55
	2	14.95	11.43	20.38	19.76
	3	9.79	7.34	20.63	14.11
	Average	13.24	10.09	24.15	17.47
R2AttU-Net	1	11.59	8.25	13.03	14.63
	2	5.25	3.73	3.78	25.57
	3	5.04	3.17	12.18	3.36
	Average	7.29	5.05	9.66	14.52
R2U-Net	1	2.32	1.36	4.48	3.13
	2	9.24	6.47	8.81	17.67
	3	4.92	3.77	3.78	10.69
	Average	5.49	3.87	5.69	10.50
U-Net	1	19.28	14.17	35.08	29.30
	2	9.82	7.16	25.56	12.48
	3	14.50	10.66	21.58	19.45
	Average	14.53	10.66	27.40	20.41
Proposed method	-	57.13	45.31	45.91	93.57

DICE: Dice similarity coefficient; IoU: Intersection over union; AttU-Net: Attention U-Net; UNet++: Nested U-Net; R2AttU-Net: Recurrent residual attention U-Net; R2U-Net: Recurrent residual U-Net.

variants with three random seeds are used as baselines to compare with the proposed method, demonstrating the superiority of the “detect-then-segment” strategy.

Specifically, the YOLOv8-Pose model is first employed to detect the rough position (*i.e.*, the keypoints) of CSC leakage points in this study, and its performance was evaluated during the training stage using four loss curves, as shown in Figure 7. The total loss and the component loss gradually stabilize as the number of training epochs increases. Taking the total loss as an example, its value decreases from a relatively high initial value to a much lower value by the end of training. This indicates that the model continuously reduces the discrepancy between its predicted and actual values, and its parameters are adjusted from the initial random state toward an optimized direction. Furthermore, three metrics were used to evaluate the performance of keypoint detection in the model. As shown in Figure 8, with the increase of training epochs, the curves of the model’s mAP50-95, Recall, and Precision exhibit a trend of increasing from lower values and eventually stabilizing. Compared to the initial training stage, the values of these three metrics have significantly improved in the later stages of training, indicating that the model has achieved good performance on the validation dataset and further revealing that the model’s parameters have been effectively optimized during the training stage. These observations and analyses demonstrate that the YOLOv8-Pose model has effectively

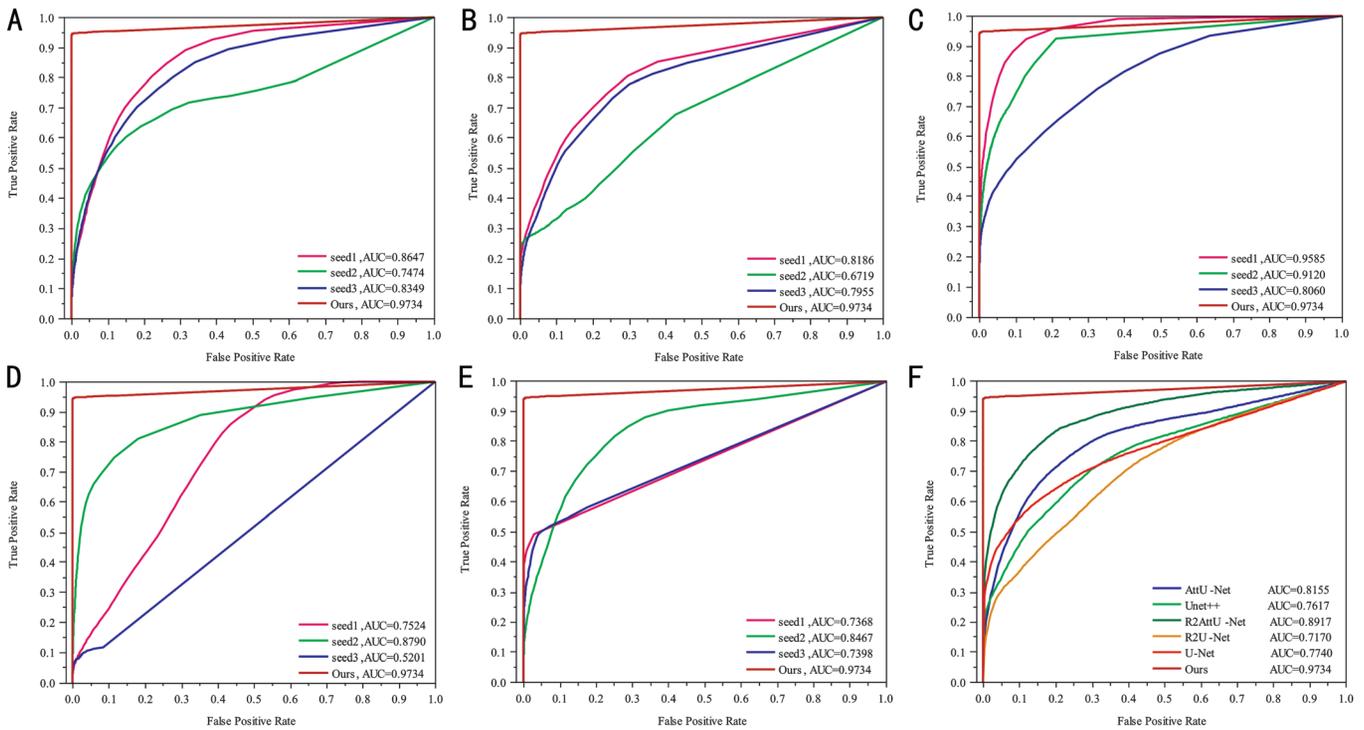


Figure 9 Comparison of ROC curve performance: proposed method vs baseline models A: AttU-Net vs proposed method; B: UNet++ vs proposed method; C: R2AttU-Net vs proposed method; D: R2U-Net vs proposed method; E: U-Net vs proposed method; F: Baseline models (averaged over 3 random seeds) vs proposed method. AUC: Area under curve; ROC: Receiver operating characteristic; AttU-Net: Attention U-Net; UNet++: Nested U-Net; R2AttU-Net: Recurrent residual attention U-Net; R2U-Net: Recurrent residual U-Net.

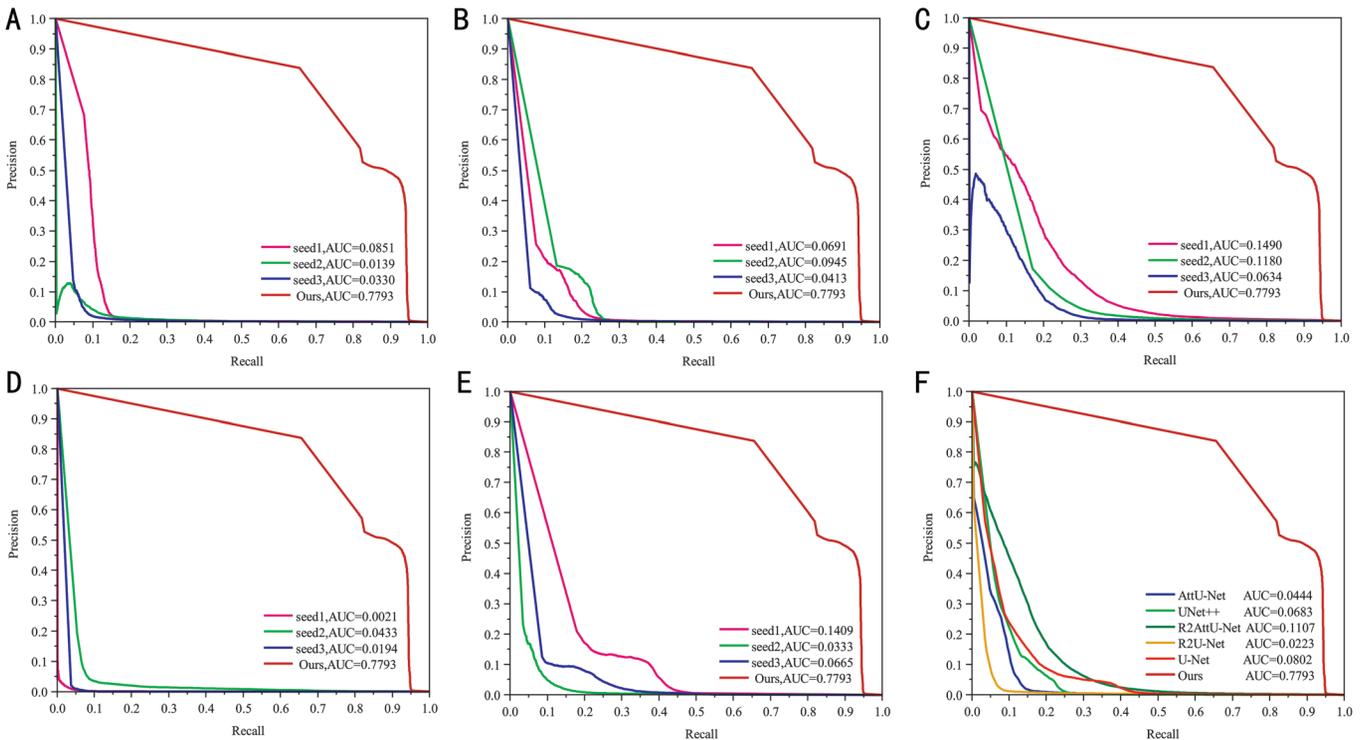


Figure 10 Comparison of PR curve performance: proposed method vs baseline models A: AttU-Net vs proposed method; B: UNet++ vs proposed method; C: R2AttU-Net vs proposed method; D: R2U-Net vs proposed method; E: U-Net vs proposed method; F: Baseline models (averaged over 3 random seeds) vs proposed method. AUC: Area under curve; PR: Precision-recall; AttU-Net: Attention U-Net; UNet++: Nested U-Net; R2AttU-Net: Recurrent residual attention U-Net; R2U-Net: Recurrent residual U-Net.

learned from the training dataset during the training stage, laying a solid foundation for the subsequent CSC leakage points segmentation task.

Moreover, we conducted a detailed evaluation of the proposed method's performance during the testing stage from both quantitative and qualitative perspectives. It is important to note

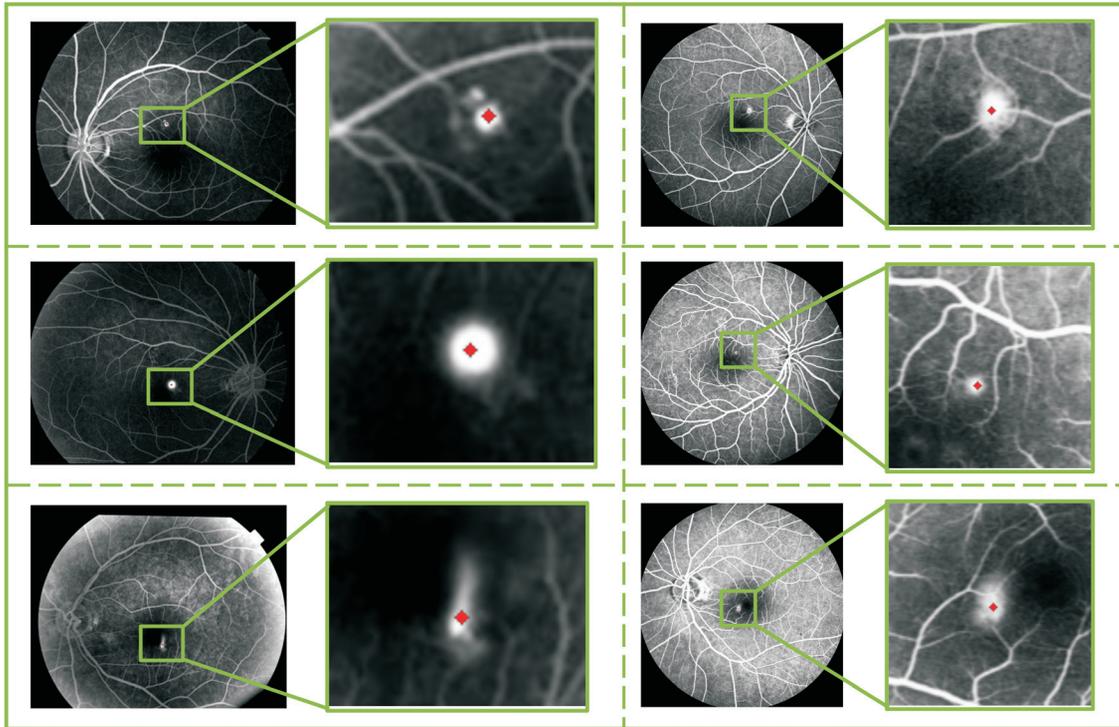


Figure 11 Examples of keypoints detected by the YOLOv8-Pose model: the red keypoints indicate the rough positions of the leakage points
YOLOv8-Pose: You Only Look Once version 8 Pose Estimation.

that the quantitative comparison of metrics was conducted on the 36 images from which leakage points were successfully detected by the YOLOv8-Pose model out of 40 test FFA images. SAM and the baseline models were then tested on the same 36 images for comparison. Each baseline model was trained for 200 epochs under three random seeds and tested. In contrast, the YOLOv8-Pose model in the proposed method was trained for 200 epochs and tested, with its output coordinates converted and used as prompts for SAM to perform the CSC leakage point segmentation task.

As shown in Figures 9 and 10, the proposed method achieves a slightly higher AUC-ROC value and a significantly higher AUC-PR value than baseline models' average performance across three random seeds. Concretely, the proposed method achieves an AUC-PR of 0.78 and an AUC-ROC of 0.97. In contrast, the baseline models achieve a highest average AUC-PR of 0.11 and a highest average AUC-ROC of 0.89 across three random seeds. It is worth noting that the AUC-PR has improved by 7.09 times compared to the best performance of the baseline models, demonstrating the significant advantage of our method in accurately identifying true positives. In addition, as shown in Table 2, our method significantly outperforms the baseline models across all evaluation metrics, demonstrating its effectiveness in segmenting CSC leakage points. For example, our method achieved a recall value of 93.57%, while the best baseline model only reached 20.41%. Regarding the DICE score, our method scored 57.13%, compared to the best baseline model's 14.53%, resulting in a lead of 42.6 percentage points.

To more clearly illustrate the CSC leakage point localization and segmentation effects of the proposed method on the test dataset, Figures 11 and 12 qualitatively show the test results of the proposed method. It can be observed that the YOLOv8-Pose model, after sufficient training, can successfully identify the rough positions of CSC leakage points in FFA images. This demonstrates the successful application of the model in the preliminary step of locating leakage points, providing precise coordinate prompts for the subsequent segmentation step. Figure 12 visually illustrates that the baseline models exhibit significant deviations in segmenting the CSC leakage points, thus leading to poor location outcomes. In contrast, SAM, guided by pixel coordinate prompts, achieves precise segmentation of CSC leakage points in FFA images.

Through observation and analyses from both the training and testing stages mentioned above, it is demonstrated that the method proposed in this study effectively completes the CSC leakage point segmentation task. This proves the effectiveness of the "detect-then-segment" strategy. Meanwhile, compared with the existing classic U-Net series models, this method also demonstrates certain advantages and competitiveness, which can be mainly attributed to the following factors: 1) By applying translation and rotation to the labels and corresponding FFA images, the YOLOv8-Pose model is endowed with ample training samples. This data augmentation lays a solid foundation for successfully identifying the leakage points. 2) The YOLOv8-Pose model is innovatively introduced in our method as the preliminary step to detect the

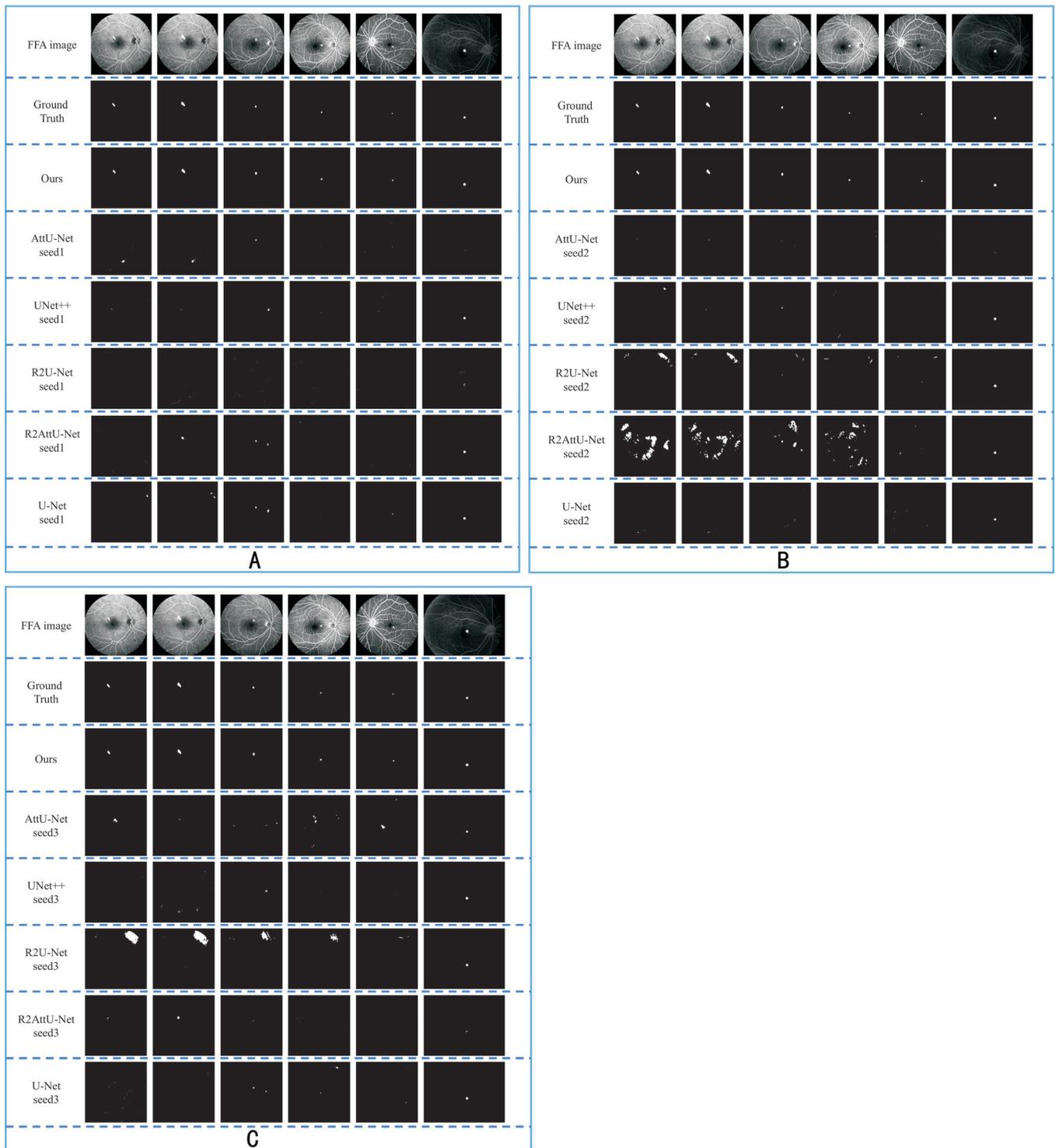


Figure 12 Performance comparison of the proposed method and baseline models A: Proposed method vs baseline model (random seed 1); B: Proposed method vs baseline model (random seed 2); C: Proposed method vs baseline model (random seed 3). FFA: Fundus fluorescein angiography; AttU-Net: Attention U-Net; UNet++: Nested U-Net; R2AttU-Net: Recurrent residual attention U-Net; R2U-Net: Recurrent residual U-Net.

rough positions of the CSC leakage points, providing SAM with precise prior guidance for the subsequent segmentation step. This “detect-then-segment” strategy effectively tackles background-dominated ineffective learning in FFA images, a common limitation in baseline models. 3) Due to the baseline models learning from the entire FFA images, a large number of leakage-like false positive areas can interfere with the models, and the smaller proportion of leakage points further reduces

their performance.

In general, the proposed method not only extends the segmentation strategies for leakage points in CSC but also offers valuable insights for the detection of lesions in other retinal diseases. Moreover, the method presented in this paper holds significant theoretical merit and has the potential to facilitate real-time localization of leakage points during traditional laser photocoagulation for CSC in ophthalmic

clinical practice. This capability may serve as an automated solution, thereby laying the foundation for enhancing the precision and efficiency of surgical interventions.

Despite its promising performance, the proposed method has some shortcomings that need to be improved: the preliminary localization of leakage points needs to be enhanced; and the segmentation accuracy for leakage points with blurred edges is not high. Therefore, in the future, we will attempt to address the limitations by improving the model architecture, refining training strategies, exploring advanced data augmentation techniques, and optimizing post-processing algorithms. Furthermore, we plan to integrate prior knowledge from the medical field, such as the morphological characteristics and distribution patterns of leakage points, into the model design to better capture these leakage points. Additionally, edge attention mechanisms will be incorporated to improve segmentation accuracy at the boundaries of leakage points.

In conclusion, to achieve precise localization of CSC leakage points and thus to assist ophthalmologists in accurate laser treatment delivery without navigational laser equipment, this study proposed an intelligent segmentation method for CSC leakage points based on a “detect-then-segment” strategy. First, a dataset including FFA images and corresponding point-level and pixel-level labels for CSC leakage points was constructed, laying the foundation for subsequent segmentation work. Second, the point-level labels were used to drive the YOLOv8-Pose model to effectively perform the preliminary localization of the leakage points, providing precise coordinate prompts for the accurate segmentation of CSC leakage points by SAM. Finally, the effectiveness of the proposed method was verified based on the pixel-level dataset. The experimental results and analysis from both the training and testing stages demonstrated the feasibility of the proposed method, conveying the potential value of the method in assisting doctors in accurately locating leakage points in traditional laser photocoagulation for CSC. It also offered a new direction for research on the automatic localization of lesions in other retinal diseases.

ACKNOWLEDGEMENTS

Foundations: Supported by the Shenzhen Science and Technology Program (No.JCYJ20240813152704006); the National Natural Science Foundation of China (No.62401259); the Fundamental Research Funds for the Central Universities (No.NZ2024036); the Postdoctoral Fellowship Program of CPSF (No.GZC20242228); High Performance Computing Platform of Nanjing University of Aeronautics and Astronautics.

Conflicts of Interest: Xu JG, None; Liu YC, None; Zhou F, None; Shen JX, None; Yan ZP, None; Hu XY, None; Yang WH, None.

REFERENCES

- 1 Park JB, Kim K, Kang MS, *et al.* Central serous chorioretinopathy: treatment. *Taiwan J Ophthalmol* 2022;12(4):394-408.
- 2 Sun YY, Zhao MW, Miao H, *et al.* Comparative efficacy of subthreshold micropulse laser versus conventional laser therapy in central serous chorioretinopathy. *Lasers Med Sci* 2025;40(1):225.
- 3 Feenstra HMA, van Dijk EHC, Cheung CMG, *et al.* Central serous chorioretinopathy: an evidence-based treatment guideline. *Prog Retin Eye Res* 2024;101:101236.
- 4 Zhu SJ, Zhan HD, Wu MN, *et al.* Research on classification method of high myopic maculopathy based on retinal fundus images and optimized ALFA-Mix active learning algorithm. *Int J Ophthalmol* 2023;16(7):995-1004.
- 5 Yang WH, Shao Y, Xu YW, *et al.* Guidelines on clinical research evaluation of artificial intelligence in ophthalmology (2023). *Int J Ophthalmol* 2023;16(9):1361-1372.
- 6 Wu JD, Fang HH, Zhu JY, *et al.* Multi-rater Prism: Learning self-calibrated medical image segmentation from multiple raters. *Sci Bull* 2024;69(18):2906-2919.
- 7 Jin K, Ye J. Artificial intelligence and deep learning in ophthalmology: current status and future perspectives. *Adv Ophthalmol Pract Res* 2022;2(3):100078.
- 8 Gong D, Li WT, Li XM, *et al.* Development and research status of intelligent ophthalmology in China. *Int J Ophthalmol* 2024;17(12):2308-2315.
- 9 Meena T, Roy S. Bone fracture detection using deep supervised learning from radiological images: a paradigm shift. *Diagnostics* 2022;12(10):2420.
- 10 Carrillo-Perez F, Pecho OE, Morales JC, *et al.* Applications of artificial intelligence in dentistry: a comprehensive review. *J Esthet Restor Dent* 2022;34(1):259-280.
- 11 Xu FB, Liu SP, Xiang YF, *et al.* Deep learning for detecting subretinal fluid and discerning macular status by fundus images in central serous chorioretinopathy. *Front Bioeng Biotechnol* 2021;9:651340.
- 12 Samagaio G, Estévez A, de Moura J, *et al.* Automatic identification of macular edema in optical coherence tomography images. *Proceedings of the 13th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*. January 27-29, 2018. Funchal, Madeira, Portugal. SCITEPRESS - Science and Technology Publications, 2018: 533-540.
- 13 Yoo TK, Kim BY, Jeong HK, *et al.* Simple code implementation for deep learning-based segmentation to evaluate central serous chorioretinopathy in fundus photography. *Trans Vis Sci Tech* 2022;11(2):22.
- 14 Husvogt L, Yaghy A, Camacho A, *et al.* Ensembling U-Nets for microaneurysm segmentation in optical coherence tomography angiography in patients with diabetic retinopathy. *Sci Rep* 2024;14(1):21520.
- 15 Aizaldeen Abdullah A, Aldhahab A, Al Abboodi HM. Review of eye diseases detection and classification using deep learning techniques. *BIO Web Conf* 2024;97:00012.

- 16 Miere A, Le Meur T, Bitton K, *et al.* Deep learning-based classification of inherited retinal diseases using fundus autofluorescence. *J Clin Med* 2020;9(10):3303.
- 17 Hassan SA, Akbar S, Rehman A, *et al.* Recent developments in detection of central serous retinopathy through imaging and artificial intelligence techniques—a review. *IEEE Access* 2021;9:168731-168748.
- 18 Goyanes E, de Moura J, Fernández-Vigo JI, *et al.* 3D features fusion for automated segmentation of fluid regions in CSCR patients: an OCT-based photodynamic therapy response analysis. *J Imag Inform Med* 2025;38(1):476-495.
- 19 Gende M, Castelo L, de Moura J, *et al.* Intra- and inter-expert validation of an automatic segmentation method for fluid regions associated with central serous chorioretinopathy in OCT images. *J Imag Inform Med* 2024;37(1):107-122.
- 20 Xu JG, Wan C, Yang WH, *et al.* A novel multi-modal fundus image fusion method for guiding the laser surgery of central serous chorioretinopathy. *Math Biosci Eng* 2021;18(4):4797-4816.
- 21 Zhang S, Yao J. Efficacy of subthreshold navigated laser photocoagulation and micropulse laser for chronic central serous chorioretinopathy. *Guoji Yanke Zazhi (Int Eye Sci)* 2022;22(10):1687-1692.
- 22 Hara C, Maruyama K, Wakabayashi T, *et al.* Choroidal vessel and stromal volumetric analysis after photodynamic therapy or focal laser for central serous chorioretinopathy. *Transl Vis Sci Technol* 2023;12(11):26.
- 23 Pfau M, van Dijk EHC, van Rijssen TJ, *et al.* Estimation of current and post-treatment retinal function in chronic central serous chorioretinopathy using artificial intelligence. *Sci Rep* 2021;11(1):20446.
- 24 Chen ML, Jin K, You K, *et al.* Automatic detection of leakage point in central serous chorioretinopathy of fundus fluorescein angiography based on time sequence deep learning. *Graefes Arch Clin Exp Ophthalmol* 2021;259(8):2401-2411.
- 25 Dai L, Wu L, Li HT, *et al.* A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nat Commun* 2021;12(1):3242.
- 26 Tang W, Cui KN, Chan RH. Optimized hard exudate detection with supervised contrastive learning. *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*. May 27-30, 2024, Athens, Greece. IEEE, 2024:1-5.
- 27 Jocher G, Chaurasia A, Qiu J. Ultralytics YOLOv8. 2023. <https://github.com/ultralytics/ultralytics>
- 28 Kirillov A, Mintun E, Ravi N, *et al.* Segment anything. *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. October 1-6, 2023, Paris, France. IEEE, 2024:3992-4003.
- 29 Ahamad SU, Ataei M, Devabhaktuni V, *et al.* Omobot: a low-cost mobile robot for autonomous search and fall detection. *2024 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*. July 15-19, 2024, Boston, MA, USA. IEEE, 2024:453-460.
- 30 Ji HH, Shi QR, Fan LL, *et al.* A license plate recognition method based on YOLOv8-pose and E-LPRNet. *2024 IEEE 13th Data Driven Control and Learning Systems Conference (DDCLS)*. May 17-19, 2024, Kaifeng, China. IEEE, 2024:1407-1411.
- 31 Zhang CL, Shi L, Bu N, *et al.* Automatic correction method of industrial instrument images based on YOLOv8 keypoint detection and perspective transformation. *Advanced Intelligent Computing Technology and Applications*. Singapore: Springer Nature Singapore 2024:3-11.
- 32 Yao YP, Chen X, Chen HJ, *et al.* Estimation of pig body measurements based on keypoint detection and multi-object tracking. *Journal of South China Agricultural University* 2024;45(5):722-729
- 33 Roy S, Wald T, Koehler G, *et al.* Sam. md: Zero-shot medical image segmentation capabilities of the segment anything model. 2023. arXiv:230405396.
- 34 Hu CF, Xia TY, Ju SH, *et al.* When SAM meets medical images: an investigation of segment anything model (SAM) on multi-phase liver tumor segmentation. 2023. arXiv:2304.08506.
- 35 Deng RN, Cui C, Liu Q, *et al.* Segment anything model (SAM) for digital pathology: assess zero-shot segmentation on whole slide imaging. 2023. arXiv:2304.04155.
- 36 Zhou T, Zhang YZ, Zhou Y, *et al.* Can SAM segment polyps 2023. arXiv:2304.07583.
- 37 Hu MZ, Li YH, Yang XF. SkinSAM: empowering skin cancer segmentation with segment anything model. 2023. arXiv:2304.13973.
- 38 Feng WJ, Zhu LT, Yu LQ. Cheap lunch for medical image segmentation by fine-tuning SAM on few exemplars. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries*. Cham: Springer Nature Switzerland. 2024:13-22.
- 39 Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. 2015. arXiv:1505.04597.
- 40 Alom MZ, Hasan M, Yakopcic C, *et al.* Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation. 2018. arXiv:1802.06955.
- 41 Oktay O, Schlemper J, Le Folgoc L, *et al.* Attention U-Net: learning where to look for the pancreas. 2018. arXiv:1804.03999.
- 42 Zhou ZW, Rahman Siddiquee MM, Tajbakhsh N, *et al.* UNet++: a nested U-Net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer International Publishing. 2018:3-11.
- 43 Qiang Z, Chen SY, Wang ZF. R2AU-Net: attention recurrent residual convolutional neural network for multimodal medical image segmentation. *Security and Communication Networks* 2021;2021:625688.