Received: 19 September 2024 | Revised: 28 October 2024 | Accepted: 11 November 2024

DOI: 10.1002/mp.17542

RESEARCH ARTICLE

MEDICAL PHYSICS

# Strip and boundary detection multi-task learning network for segmentation of meibomian glands

**Weifang Zhu**[1] | **Dengfeng Liu**[1] | **Xinyu Zhuang**[2] | **Tian Gong**[1] | **Fei Shi**[1] | **Dehui Xiang**[1] | **Tao Peng**[3] | **Xiaofeng Zhang**[2] | **Xinjian Chen**[1,4]

[1]MIPAV Lab, School of Electronic and Information Engineering, Soochow University, Suzhou, China

[2]Department of Ophthalmology, The Fourth Affiliated Hospital of Soochow University, Suzhou, China

[3]School of Future Science and Engineering, Soochow University, Suzhou, China

[4]State Key Laboratory of Radiation Medicine and Protection, Soochow University, Suzhou, China

**Correspondence**
Xiaofeng Zhang, Department of Ophthalmology, the Fourth Affiliated Hospital of Soochow University, Suzhou, China.
Email: zhangxiaofeng@suda.edu.cn

Xinjian Chen, MIPAV Lab, School of Electronic and Information Engineering, Soochow University, Suzhou, China.
Email: xjchen@suda.edu.cn

**Funding information**
National Natural Science Foundation of China, Grant/Award Numbers: 62371326, U20A20170, 62271337, 62371328; Natural Science Foundation of Jiangsu Province of China, Grant/Award Number: BK20231310; National Key Research and Development Program of China, Grant/Award Number: 2018YFA0701700; China Postdoctoral Science Foundation, Grant/Award Number: 2023M742568; Priority Academic Program Development of Jiangsu Higher Education Institutions, Grant/Award Number: YX11900123

## Abstract

**Background:** Automatic segmentation of meibomian glands in near-infrared meibography images is basis of morphological parameter analysis, which plays a crucial role in facilitating the diagnosis of meibomian gland dysfunction (MGD). The special strip shape and the adhesion between glands make the automatic segmentation of meibomian glands very challenging.

**Purpose:** A strip and boundary detection multi-task learning network (SBD-MTLNet) based on encoder-decoder structure is proposed to realize the automatic segmentation of meibomian glands.

**Methods:** A strip mixed attention module (SMAM) is proposed to enhance the network's ability to recognize the strip shape of glands. To alleviate the problem of adhesion between glands, a boundary detection auxiliary network (BDA-Net) is proposed, which introduces boundary features to assist gland segmentation. A self-adaptive interactive information fusion module (SIIFM) based on reverse attention mechanism is proposed to realize information complementation between meibomian gland segmentation and boundary detection tasks. The proposed SBD-MTLNet has been evaluated on an in-house dataset (453 images) and a public dataset MGD-1K (1000 images). Due to the limited number of images, a five-fold cross validation strategy is adopted.

**Results:** Average dice coefficient of the proposed SBD-MTLNet reaches 81.08% and 84.32% on the in-house dataset and the public one, respectively. Comprehensive experimental results demonstrate the effectiveness the proposed SBD-MTLNet, outperforming other state-of-the-art methods.

**Conclusions:** The proposed SBD-MTLNet can focus more on the shape characteristics of the meibomian glands and the boundary contour information between the adjacent glands via multi-task learning strategy. The segmentation results of the proposed method can be used for the quantitative morphological characteristics analysis of meibomian glands, which has potential for the auxiliary diagnosis of MGD in clinic.

**KEYWORDS**
attention mechanism, meibography image segmentation, meibomian gland, multi-task learning

Weifang Zhu, Dengfeng Liu, and Xinyu Zhuang contributed equally to this work.

  

# 1 INTRODUCTION

Meibomian glands are the sebaceous glands located within the eyelid margin and embedded in the tarsal plate of the eyelid, which are composed of numerous elastic and pliable fibers, densely connective tissue and abundant meibomian gland acini. The meibomian gland acini secrete a lipid mixture known as meibum. Meinum is a complex mixture of diverse polar and nonpolar lipids, along with 90 distinct proteins, and it plays an important role in the maintenance of eye homeostasis and health.[1–4] Meibomian gland dysfunction (MGD) is a kind of chronic and diffuse meibomian gland disorder that typically manifests as obstruction of the meibomian gland or alterations in the quantity or quality of meibum secreted by the gland.[5] MGD is one of the most common causes of dry eye disease (DED), characterized by the morphological alterations of the meibomian gland.[6,7] The precise automatic segmentation of meibomian glands is the key to the quantitative morphological analysis of meibomian gland, such as tortuosity, length, width and count, which is crucial for the identification of MGD.

Machine learning based methods have been applied in the segmentation of meibomian glands, including semi-automatic segmentation methods based on software such as ImageJ, as well as automatic segmentation methods based on traditional image processing, to enable clinicians to obtain information about the meibomian gland accurately and efficiently.[8–11] Arita et al. developed an objective meibomian gland area evaluation method using ImageJ software, in which the meibomian gland and atrophy regions were manually marked on near-infrared images, and the atrophy degree of the meibomian gland was subsequently calculated.[8] Koh et al. utilized fuzzy C-means clustering based method to enhance the contrast of near-infrared meibomian images, extract glands and inter-gland lines based on Gaussian smoothing and morphological processing, and detect the gland width using SIFT-Shannon entropy. Based on the features including the average arclength of the gland and inter-gland lines and gland width, a linear support vector machine (SVM) was trained to separate healthy and unhealthy glands.[9] Koprowski et al. proposed an automatic meibomian gland detection method, which included median filtering and normalization based image pre-processing, and morphological opening, Gaussian filtering, skeletonization, distance transform and watersheds based image processing. The algorithm was implemented with MATLAB platform with image processing toolbox and evaluated on 172 lower and upper eyelid images from 86 eyes.[10] Xiao et al. developed an automated multi-parametric algorithm for the objective analysis of meibography images, including segmentation of the tarsal conjunctiva area as the region of interest (ROI), segmentation and identification of meibomian glands within the ROI, and quantitative multi-parametric analysis.[11] These traditional machine learning based methods have some limitations: (1) Semi-automatic segmentation methods based on software like ImageJ, may lead to significant variances between different operators, which is subjective and inefficient. (2) Automatic segmentation methods based on traditional machine learning techniques typically require manually designed features, which is time-consuming and experiential.

In recent years, deep learning especially convolutional neural network (CNN) based methods have been applied for the automatic segmentation and function evaluation of meibomian glands.[12–14,16–21] Wang et al. used pyramid scene parsing network (PSPNet) to segment the entire eyelid region and meibomian gland atrophy area and calculate the percent of atrophy.[12] However, this method can not segment the individual meibomian gland and analyze the morphological features further. Khan et al. developed a conditional generative adversarial network (CGAN) based method for the automated meibomian gland segmentation, which was trained and tested on 112 meibography images (90 for training) and achieved average aggregated Jaccard index of 0.664.[13] Swiderska et al. developed Deeplabv3+ with ImageNet pre-trained Inception-ResNetv2 as backbone and Mask R-CNN with ResNet-50 as feature extractor to segment meibomian gland, respectively. These two CNN based methods were evaluated on 1616 meibography images (697 upper eyelids and 919 lower eyelids) from 282 individuals and achieved average aggregated Jaccard index of 0.4718 and 0.4476 respectively.[14]

Since U-Net[15] was successfully proposed for medical image segmentation in 2015, many U-Net and its variants based methods have also been proposed for the automatic meibomian gland segmentation.[16–21] Prabhu et al. firstly applied U-Net for the automatic segmentation of meibomian glands in meibography images, which was evaluated on 400 images from prototype hand-held imager and 400 ones from Oculus Keratograph 5M. Meibomian metrics including gland-drop, tortuosity, width, length and number of glands were quantitatively analyzed.[16] Setu et al. applied transfer learning and U-Net with ImageNet pre-trained Inception-ResNetv2 as backbone to segment meibomian glands and extract the attributes of individual meibomian gland, including gland number, length, width and tortuosity, achieving an objective evaluation of morphological parameters.[17] This method can work well in most cases without gland adhesions. Dai et al. proposed an enhanced Mini U-Net for the automatic segmentation of meibomian glands.[18] However, this model was trained only on 60 right-eye images (40 for training and 20 for validation), which may result in skewed distribution and overfitting problems. Saha et al. developed an encoder-decoder architecture with skip connections and ResNet34 based encoder model to segment meibomian glands and eyelids. The area and ratio of glands were quantitatively analyzed and the meiboscore was estimated. They also contributed a public infrared meibography image dataset MGD-1k.[19] Wang et al. built a U-Net based method combined with VAE-GAN for automatically segmenting tarsus and meibomian gland areas in meibography images, which was evaluated on 1434 meibography images from 366 patients diagnosed with DED (70% for training, 20% for validation and 10% for test) and achieved area under the ROC curve (AUC) values of 0.985 and 0.938, respectively.[20] Our previous work proposed a U-shaped information fusion network (IF-Net) to segment meibomian gland area (eyelid region) and meibomian gland atrophy area,[21] without precise individual meibomian gland segmentation.

In view of the strip-shaped characteristics of meibomian glands and the phenomenon of adhesions between glands, a U-shape encoder-decoder structure based strip and boundary detection multi-task learning network (SBD-MTLNet) is proposed in this paper, which improves the segmentation performance by introducing boundary features to assist the gland segmentation. The main contributions of this paper can be summarized as follows:

1. A strip mixed attention module (SMAM) is proposed, which combines local and global spatial features from vertical and horizontal dimension to improve the recognition of strip-shaped glands.
2. To relieve the problem of adhesion between glands, a boundary detection auxiliary network (BDA-Net) is proposed to effectively utilize the boundary information.
3. A self-adaptive interactive information fusion module (SIIFM) based on reverse attention mechanism is proposed to realize information complementation between meibomian gland segmentation and boundary detection tasks.
4. A multi-task learning strategy is adopted to jointly learn meibomian gland segmentation and boundary detection to improve the gland segmentation performance.
5. A multi-scale attention module (MSAM) is embedded at the top of the encoder path to extract rich multi-scale information.

# 2 | METHODS

## 2.1 | Overview of the proposed SBD-MTLNet

Figure 1 shows the overall architecture of the proposed encoder-decoder structure based SBD-MTLNet, which includes one feature encoder path and dual feature decoder paths, equipped with multi-scale attention module (MSAM) and strip mixed attention module (SMAM) at the top of encoder path, and boundary detection auxiliary network (BDA-Net) and self-adaptive interactive information fusion modules (SIIFM) in the decoder paths, respectively.

## 2.2 | Feature encoder and decoders

The meibomian gland segmentation and boundary detection tasks share the same feature encoder path. After the original input passes through a $3 \times 3$ convolution, the feature map will be sent into four encoder layers, each of which includes one max-pooling, two $3 \times 3$ convolutions, batch normalization and Relu activation. The dual decoder paths consist of two branches: one is the main branch of meibomian gland segmentation composed of four bilinear interpolation upsampling based decoder layers and SIIFM modules, and the other is the boundary detection auxiliary network (BDA-Net) composed of four layers of boundary detection modules (BDM).

## 2.3 | Strip mixed attention module

The irregular strip-shaped structures of the meibomian glands make the segmentation challenging. To address this issue, a strip mixed attention module (SMAM) is proposed, which mainly includes a strip mixed pooling block (SMPB) and a channel attention combination block (CACB). The structure of SMAM is shown in Figure 2.

### 2.3.1 | Strip mixed pooling block

As shown in Figures 1 and 2, SMAM takes the output feature of the fourth encoder layer $E_4$ as input $F_{SMAM,in}$ and sends it into two blocks including SMPB and CACB. To extract local features at different scales, the upper branch of SMPB adjusts the input feature to grid sizes of $2 \times 2$ and $7 \times 7$ respectively, through global grid max pooling (GGMP). After that, local information is extracted through $1 \times 1$ and $3 \times 3$ convolutions respectively. Bilinear interpolation upsampling is then utilized to restore the resolution of features, getting $M_{2\times2} = UP(GGMP_{2\times2}\ (F_{SMAM,in}))$ and $M_{7\times7} = UP(GGMP_{7\times7}(F_{SMAM,in}))$ respectively, in which $UP(\cdot)$ represents upsampling. The input $F_{SMAM,in}$ is adjusted by a $3 \times 3$ convolution and fused with $M_{2\times2}$ and $M_{7\times7}$ to get $F_M = Conv_{3\times3}\ (F_{SMAM,in}) + M_{2\times2} + M_{7\times7}$, which can extract local spatial characteristics across different scales.

The lower branch of SMPB adjusts the input feature through global grid average pooling (GGAP) from the horizontal and vertical directions to the region sizes of $1 \times W$ and $H \times 1$, respectively. Then, with strip-shaped kernels, $1 \times 3$ and $3 \times 1$ convolutions are carried out for feature extraction so that the network can focus

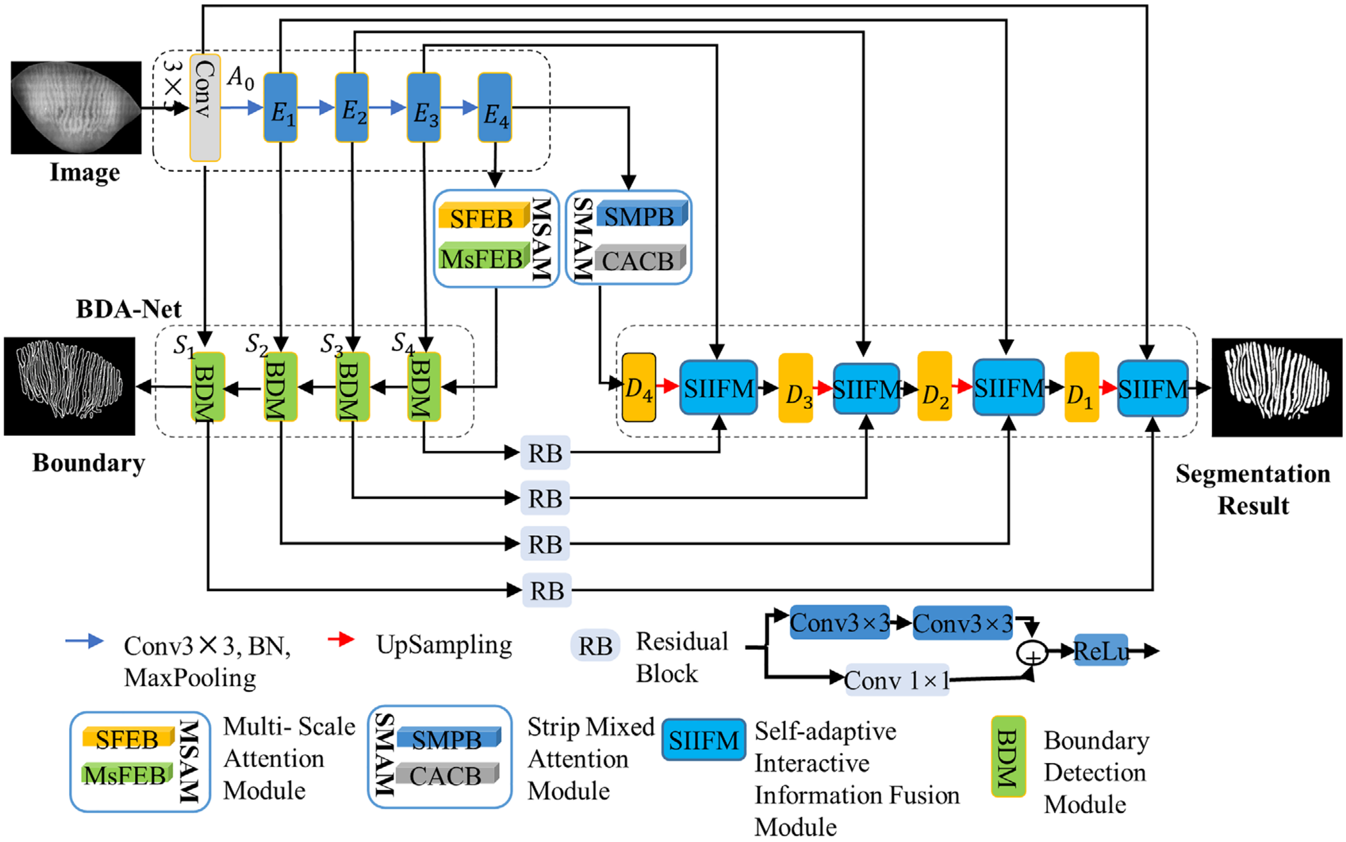


**FIGURE 1** The overall structure of the proposed SBD-MTLNet.

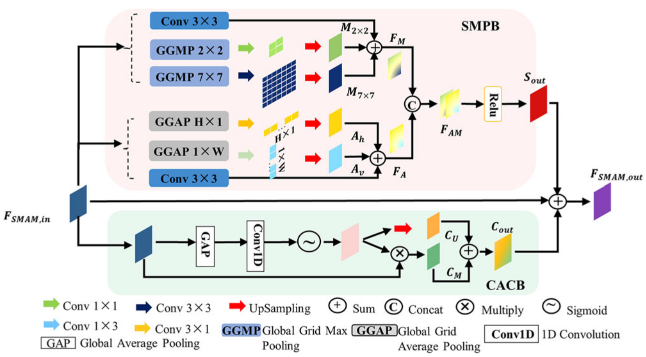


**FIGURE 2** The strip mixed attention module (SMAM).

on the horizontal and vertical strip-shaped areas. Bilinear interpolation upsampling is used to restore the resolution and get $A_h = UP(GGAP_{H\times 1}(F_{SMAM,in}))$ and $A_v = UP(GGAP_{1\times W}(F_{SMAM,in}))$, which preserves the global features of the strip-shaped region and the local features within the grid. The input feature $F_{SMAM,in}$ is also adjusted by a 3 × 3 convolution and fused with $A_h$ and $A_v$ to obtain the output feature $F_A = Conv_{3\times 3}\,(F_{SMAM,in}) + A_h + A_v$.

The outputs of upper branch and lower branch of SMPB $F_M$ and $F_A$ are concatenated and activated with ReLu function to get the output of the SMPB $S_{out} = Relu(Cat(F_M, F_A))$.

### 2.3.2 | Channel attention combination block

Inspired by efficient channel attention (ECA) module proposed in ECA-Net,[22] a channel attention combination block (CACB) is designed to improve the channel utilization and enhance the channel correlation information. As shown in Figure 2, a global average pooling (GAP) is applied to the input feature $F_{SMAM,in}$ to obtain the aggregated feature. Then, a fast one dimensional convolution and a sigmoid activation are used to generate channel matrix. The input feature $F_{SMAM,in}$ is then multiplied with the channel matrix to get $C_M = \sigma(Conv_{1D}\,(GAP(F_{SMAM,in}))) \times F_{SMAM,in}$, in which $\sigma(\cdot)$ represents the sigmoid activation function and $Conv_{1D}(\cdot)$ represents one dimensional convolution. $C_M$ can capture local cross-channel interaction information efficiently. To reduce the useless channel information, the channel matrix is upsampled by bilinear interpolation to get $C_U = Up(\sigma(Conv_{1D}\,(GAP(F_{SMAM,in}))))$. $C_M$ and $C_U$ are summed to get the output of CACB $C_{out}$.

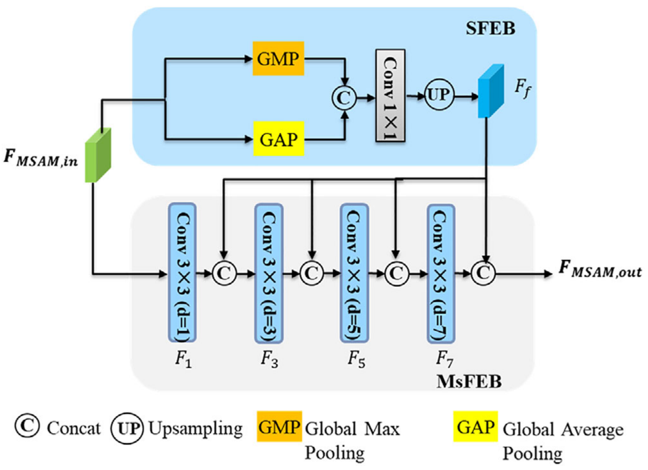


**FIGURE 3** The multi-scale attention module (MSAM).

Finally, the output of SMPB $S_{out}$, the output of CACB $C_{out}$ and the input feature $F_{SMAM,in}$ are summed to get the final output of SMAM $F_{SMAM,out}$.

## 2.4 | Multi-scale attention module

Because the sizes of meibomian glands vary widely, traditional convolution operations may cause incomplete segmentation of large size targets, while the absence of local context information for boundary feature of small size targets may lead to inaccurate segmentation. Therefore, how to effectively extract multi-scale context information to adapt to different sizes and shapes of meibomian glands is an urgent problem to be solved. In the previous semantic segmentation networks, PSPNet[23] proposed pyramid pooling module, in which pooling operations with different sizes were adopted to obtain multi-scale features. CE-Net[24] used several convolution branches with different receptive fields to improve the model's ability to obtain multi-scale information. CPFNet[25] proposed scale-aware pyramid fusion module, in which dilated convolutions with different dilation rates (1, 2 and 4) were performed to capture different scale features and a scale-aware module was designed to fuse these features. Inspired by previous work, in order to adapt to the variable size and morphology of meibomian glands, we design a multi-scale attention module (MSAM) to fully extract the multi-scale information of the glands.

Figure 3 shows the structure of MSAM, which consists of two parts: multi-scale feature extraction block (MsFEB) and spatial feature extraction block (SFEB). MsFEB utilizes four cascaded dilated convolutions with different dilation rates (1, 3, 5 and 7) to enlarge the receptive field of network and extracts features layer by layer ($F_1$, $F_3$, $F_5$ and $F_7$) to obtain multi-scale information. Although multi-scale features can be obtained by dilated convolutions, the spatial information is still insufficient. Therefore, the SFEB is designed to enhance the network's attention to the target at different scales. SFEB combines local and global spatial features by applying global maximum pooling (GMP) and global average pooling (GAP) to the input feature $F_{MSAM,in}$ respectively, which will be concatenated, refined by a $1 \times 1$ convolution, and upsampled by bilinear interpolation to get spatial feature $F_f$. $F_f$ is concatenated and fused with $F_1$, $F_3$, $F_5$ and $F_7$ layer by layer to get the output multi-scale spatial feature of MSAM $F_{MSAM,out}$.

## 2.5 | Boundary detection auxiliary network

In most encoder-decoder based CNN segmentation networks, consecutive pooling and convolution operations in the encoder path may lead to the loss of details, which may result in insufficient discriminative features in the decoder path and further lead to unsatisfied segmentation results, especially in the boundary region of the target. For the segmentation of meibomian glands, this will lead to gland adhesion in the segmentation results due to the small spacing between glands and their irregular shapes. In order to alleviate the adhesion between glands, it is necessary to strengthen the gland boundary detection. Chen et al proposed reverse attention to guide side-output residual learning in a top-down manner, which erased the current prediction and can make the salient object detection network discover the missing object parts and residual details (such as boundary) effectively and quickly, resulting in significant performance improvement.[26] Xia et al proposed an edge-reinforced network for 3D vessel-like structure segmentation, in which a reverse edge attention module was designed and embedded between adjacent layers of the encoder to discover edge features by increasing the weight of voxels on the border and better preserve the spatial edge information.[27] Kim et al proposed an uncertainty augmented context attention network (UACANet) for polyp segmentation, in which uncertain area with ambiguous saliency score (around 0.5) that is highly related to the boundary information was computed and combined with foreground and background area for context attention.[28] Inspired by these previous researches, we propose a boundary detection auxiliary network (BDA-Net), in which reverse attention mechanism is adopted and uncertain boundary region is computed to supervise the boundary characteristics of glands and assist in the segmentation of meibomian glands, alleviating the adhesion between glands. As shown in Figure 1, the BDA-Net consists of four cascaded boundary detection modules (BDM). The output of the BDA-Net is the predicted boundary results, which will be compared with the ground truth of gland boundary to calculate the loss function.

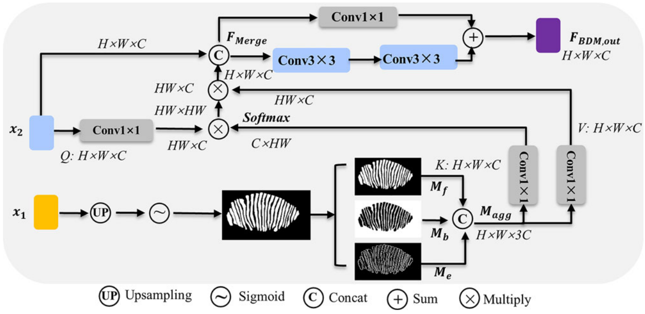


**FIGURE 4** The boundary detection module (BDM).

The structure of the BDM is shown in Figure 4. Take the output feature of MSAM or the output feature of the previous stage of BDM as input $x_1$, and take the output feature of the corresponding encoder layer as input $x_2$. Input $x_1$ is upsampled by bilinear interpolation to recover the same size as input $x_2$. After sigmoid activation of $x_1$, the feature maps for foreground $M_f$, background $M_b$ and uncertain boundary region $M_e$ are calculated by Equations (1)–(3) respectively,

$$M_f = \sigma(x_1) \tag{1}$$

$$M_b = 1 - \sigma(x_1) \tag{2}$$

$$M_e = 0.5 - abs(\sigma(x_1) - 0.5) \tag{3}$$

where $abs(\cdot)$ means absolute value operation.

$M_f$, $M_b$ and $M_e$ are then concatenated to get the aggregated feature map $M_{agg}$. After two $1 \times 1$ convolution operations on $M_{agg}$ and one $1 \times 1$ convolution on $x_2$, the Key ($K$), Value ($V$) and Query ($Q$) maps are got, respectively. Then, the interactive attention between $x_2$ and $M_{agg}$ is calculated, which is concatenated with $x_2$ to get the merged feature $F_{Merge}$. $F_{Merge}$ introduces uncertain region in attention mechanism and enhances the boundary region in $x_2$, guiding the network to focus on the boundary region. Finally, $F_{Merge}$ is refined by a residual block composed of one $1 \times 1$ convolution and two $3 \times 3$ convolutions to get the output of BDM $F_{BDM,out}$. These operations can be described by Equations (4)–(7),

$$M_{agg} = Cat(M_f,\ M_b,\ M_e) \tag{4}$$

$$K = Conv_{1\times1}\left(M_{agg}\right),\ V = Conv_{1\times1}\left(M_{agg}\right),$$
$$Q = Conv_{1\times1}\ (x_2) \tag{5}$$

$$F_{Merge} = Cat\left[Softmax\left(K \times Q^T\right) \times V,\ x_2\right] \tag{6}$$

$$F_{BDM,out} = Conv_{1\times1}\left(F_{Merge}\right)$$
$$+Conv_{3\times3}\left(Conv_{3\times3}\left(F_{Merge}\right)\right) \tag{7}$$

where K, V, Q represents Key, Value and Query respectively.

## 2.6 | Self-adaptive interactive information fusion module

To improve the gland segmentation performance, the multi-task learning strategy consisting of strip mixed attention module (SMAM) for strip-shaped gland recognition and BDA-Net for gland adhesion alleviation is adopted to jointly learn meibomian gland segmentation and boundary detection. Because there is information gap between two tasks, it is not appropriated to use direct summation or concatenation between two tasks for information fusion, which can ultimately affect the information recovery in the decoder path. To mine interactive information and achieve information complementarity between two tasks, the self-adaptive interactive information fusion module (SIIFM) is proposed, which is embedded in the decoder path and can effectively utilize boundary information to assist the segmentation of meibomian glands.

The structure of the SIIFM is shown in Figure 5. Boundary feature $S_i (i = 1, 2, 3, 4)$ from BDA-Net is activated by sigmoid function and inversed to generate the reverse attention weight matrix $S_{i,r}$ via Equation (8). $S_{i,r}$ is multiplied by the corresponding segmentation feature $E_{i-1}/\ A_0 (i = 2, 3, 4)$ from encoder path s then added to $S_i$ to get the enhanced boundary feature $S_E$ via Equation (9), whose boundary information is richer due to the adaptive fusion of the segmentation information. Similarly, the boundary information is adaptively integrated into the segmentation feature $E_{i-1}/A_0$ based on reverse attention to get the smoother segmentation feature $E_S$ by using Equations (10) and (12). Enhanced boundary feature $S_E$, smoother segmentation feature $E_S$ and decoder feature $D_i$ from the higher stage of decoder are concatenated to get $F_{ES}$. Through a $1 \times 1$ convolution and a residual block (RB) operation on $F_{ES}$, the output of the SIIFM $F_{SIIFM,out}$ is got by using Equation (13).

$$S_{i,r} = [1 - \sigma(S_i)] \tag{8}$$

$$S_E = S_{i,r} \times E_{i-1}/A_0 + S_i \tag{9}$$

$$E_{i-1,r}/\ A_{0,r} = [1 - \sigma(\ E_{i-1}/\ A_0)] \tag{10}$$

$$E_S = E_{i-1,r}\ /\ A_{0,r} \times S_i\ + E_{i-1}/A_0 \tag{11}$$

$$F_{ES} = Cat\,(S_E\ , E_S\ , D_i) \tag{12}$$

$$F_{SIIFM,out} = Relu\,(Conv_{1\times1}\,(Conv_{1\times1}\,(F_{ES}))$$
$$+\ Conv_{3\times3}\,(Conv_{3\times3}\,(Conv_{1\times1}\,(F_{ES})))) \tag{13}$$

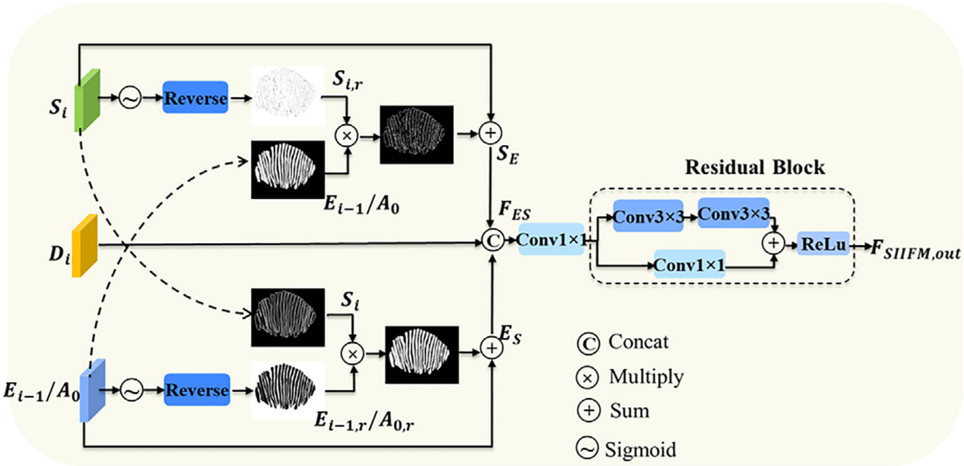


**FIGURE 5** The self-adaptive interactive information fusion module (SIIFM).

where *Relu* represents the ReLu activation function.

## 2.7 | Loss function

Joint loss function is designed for the multi-tasking learning strategy of joint gland segmentation and boundary detection. For gland segmentation task, Dice loss and binary cross entropy (BCE) loss are adopted as joint segmentation loss function $L_{Segmentation}$. BCE loss is adopted as the loss function of boundary detection task $L_{Boundary}$. Therefore, the total joint loss function $L_{Total}$ is as follows,

$$L_{Boundary} = L_{BCE} \tag{14}$$

$$L_{Segmentation} = L_{BCE} + \alpha L_{Dice} \tag{15}$$

$$L_{Total} = L_{Segmentation} + \beta L_{Boundary} \tag{16}$$

where $\alpha$ represents the tradeoff between binary cross entropy loss $L_{BCE}$ and Dice loss $L_{Dice}$, which is set to 1 in our experiments. $\beta$ represents the weight between the loss of $L_{Boundary}$ in the boundary detection task and the loss of $L_{Segmentation}$ in the gland segmentation task, which is set to 1 in this paper.

# 3 | EXPERIMENT SETTINGS

## 3.1 | Datasets

The proposed SBD-MTLNet has been evaluated on an in-house dataset and a public dataset MGD-1K.[19] The in-house dataset contains of 505 near-infrared meibography images from 300 patients, including 259 upper eyelids and 246 lower eyelids, with an original resolution of 1024 × 1280. These images were provided by the Fourth Affiliated Hospital of Soochow University.

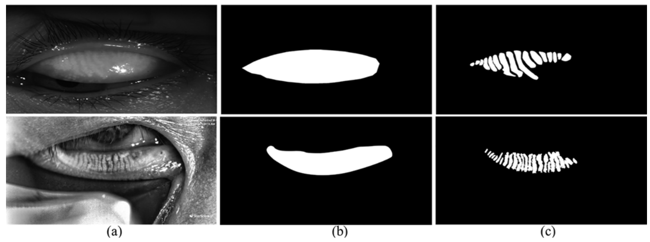


**FIGURE 6** Two near-infrared meibography images and ground truth of eyelids and meibomian glands. (a) Original near-infrared meibography images; (b) ground truth of eyelids; (c) ground truth of meibomian glands.

After excluding images with poor quality, 453 images (249 upper eyelids and 204 lower eyelids) were adopted in this study. This study adhered to the tenets of Declaration of Helsinki and was approved by the Institutional Review Board of the Fourth Affiliated Hospital of Soochow University (No. 20241036). Informed consent was obtained from all subjects. All the ground truth annotations were manually marked using ITK-SNAP[29] software by three annotators under the supervision of ophthalmologists. The public dataset MGD-1K includes 1000 meibography images with precise meibomian gland and eyelid segmentation ground truth from 320 patients, including 467 images of upper eyelids and 533 images of lower eyelids, imaged by the LipiView II ocular surface interferometer. Figure 6 shows two near-infrared meibography images and the corresponding ground truth of eyelids and meibomian glands, in which the first line is from the in-house dataset and the second line is from public MGD-1K. Due to the limited number of images, a five-fold cross validation strategy is adopted for the evaluation of the proposed method.

## 3.2 | Implementation details

The implementation of the proposed SBD-MTLNet is based on the public PyTorch platform and NVIDIA

RTX 3090 GPU with 24GB memory. "poly" learning rate policy is adopted to schedule learning rate $lr = basic\_lr \times (1 - \frac{iter}{total\_iter})^{power}$, where basic learning rate *basic_lr* is set to 0.001, *power* is set to 0.9, *iter* and *total_iter* represent the current iteration number and the total number of iterations. Stochastic gradient descent (SGD) algorithm with momentum of 0.9 and weight decay of 0.001 is used to optimize the network. The batch size and the number of epochs are set to 4 and180 respectively. In addition, considering the strip shape of the meibomian glands and the irregularities of the growth direction of glands, besides the rigid transformations such as horizontal flip, vertical flip and random rotation (–45 degrees to 45 degrees) are adopted as the online data augmentation strategies, non-rigid data augmentation strategies such as elastic transformation and disturbance are also adopted to increase the diversity of data. In all the ablation and comparison experiments, the eyelid ROI extracted based on the eyelid region segmentation results of our previous work IF-Net is adopted as the input of the network.

### 3.3 | Evaluation metrics

In order to objectively evaluate the performance of the proposed SBD-MTLNet, five indexes including Dice coefficient (Dice), intersection-over-union (IoU), sensitivity (Sen), specificity (Spe) and accuracy (Acc) are adopted.

## 4 | RESULTS

### 4.1 | Ablation experiments

Six ablation experiments about the proposed SMAM, SIIFM, BDA-Net and MSAM modules have been conducted on two datasets, including: ① Baseline (U-Net after channel halving), ② Baseline with SMAM (Baseline+SMAM), ③ Baseline with BDA-Net (Baseline+BDA-Net), ④ Baseline with BDA-Net and MSAM (Baseline+BDA-Net+MSAM), ⑤ Baseline with SMAM, BDA-Net and MSAM (Baseline+SMAM+BDA-Net+MSAM), ⑥ Baseline with SMAM, BDA-Net, MSAM and SIIFM (Baseline+SMAM+BDA-Net+MSAM+SIIFM, the proposed SBD-MTLNet). Tables 1 and 2 show the results of the ablation experiments on two datasets respectively.

It can be seen from Tables 1 and 2 that the Baseline with SMAM (experiment ②, Baseline+SMAM) has improved in all indicators compared with the Baseline on both datasets, which indicates that SMAM makes the network pay more attention to the information of the strip region through strip mixed pooling including global grid max pooling and global gird average pooling, and the channel attention combination block (CACB) improves the channel utilization and enhances the channel correlation information. It can be seen from experiment ③ (Baseline+BDANet) that boundary features including edge, texture and shape are effective to assist the meibomian gland segmentation. Compared with the Baseline, Dice increases by 1.12% and 0.65% on two datasets respectively and IoU increases by 2.01% and 0.93% respectively.

The results of experiment ④ (Baseline+BDANet+MSAM) show that the addition of MSAM improves most of the indexes further. The main reasons include: (1) The multi-scale feature extraction block (MsFEB) in the MSAM improves the network's ability to extract multi-scale features through cascaded dilated convolutions with different dilation rates, increasing the receptive field of network and enabling the network to capture more detailed information. (2) The spatial feature extraction block (SFEB) enhances the network's attention to the target at different scales by combining the local and global spatial features. Experiment ⑤ (Baseline+SMAM+BDANet+MSAM) combines the meibomian gland segmentation and gland boundary detection through multi-task learning strategy. The experimental results both on in-house dataset and public MGD-1K dataset show that the segmentation performance has been further improved based on experiment ④.

In Experiment ⑤ (Baseline+SMAM+BDANet+MSAM), due to the feature fusion in the decoder path adopts pixel-wise addition, a lot of interference information is introduced while the boundary features are integrated into the segmentation features. Therefore, the "information purification" of the fused features is crucial to further improve the segmentation performance. In the experiment ⑥(Baseline+SMAM+BDANet+MSAM+SIIFM, the proposed SBD-MTLNet), the self-adaptive interactive information fusion module (SIIFM) is inserted into the decoder path, which can selectively fuse valid information by weighted sum of segmentation features and boundary features. As the major indexes have increased, the results both in Tables 1 and 2 show that the performance of the proposed model has been further improved after the "information purification" of the SIIFM.

### 4.2 | Comparison experiments

In order to comprehensively evaluate the performance of SBD-MTLNet, comparison experiments have been conducted between the proposed SBD-MTLNet and several state-of-the-art segmentation networks including U-Net,[15] attention-UNet,[30] UNet++,[31] CPFNet,[25] CS$^2$-Net,[32] SP-Net,[33] CE-Net,[24] IF-Net,[21] GD-Net[34] and MLFGNet.[35] Tables 3 and 4 show the results of

**TABLE 1** Ablation experiment results on in-house dataset (mean ± standard deviation).

| No | SBD-MTLNet SMAM | BDA-Net | MSAM | SIIFM | Dice (%) | IoU (%) | Sen (%) | Spe (%) | Acc (%) |
|---|---|---|---|---|---|---|---|---|---|
| ① | | | | | 79.36 ± 1.38 | 64.79 ± 1.64 | 81.62 ± 1.58 | 93.43 ± 0.39 | 90.24 ± 0.48 |
| ② | √ | | | | 80.72 ± 0.81 | 67.35 ± 1.11 | 82.67 ± 1.36 | 93.67 ± 0.48 | 91.27 ± 0.47 |
| ③ | | √ | | | 80.48 ± 0.89 | 66.80 ± 0.99 | 82.00 ± 1.12 | 93.61 ± 0.60 | 91.20 ± 0.49 |
| ④ | | √ | √ | | 80.66 ± 0.81 | 66.91 ± 1.21 | 82.39 ± 1.26 | 93.53 ± 0.61 | 91.21 ± 0.52 |
| ⑤ | √ | √ | √ | | 80.88 ± 0.72 | 67.81 ± 0.98 | **83.11 ± 0.86** | 93.59 ± 0.61 | **91.33 ± 0.48** |
| ⑥ | √ | √ | √ | √ | **81.08 ± 0.63** | **67.93 ± 0.83** | 81.88 ± 0.90 | **93.86 ± 0.60** | 91.29 ± 0.51 |

*Note*: The bold values represent the best performances.

**TABLE 2** Ablation experiment results on public MGD-1K dataset (mean ± standard deviation).

| No | SBD-MTLNet SMAM | BDA-Net | MSAM | SIIFM | Dice (%) | IoU (%) | Sen (%) | Spe (%) | Acc (%) |
|---|---|---|---|---|---|---|---|---|---|
| ① | | | | | 83.39 ± 0.14 | 71.84 ± 0.20 | 86.54 ± 0.71 | 94.65 ± 0.21 | 92.88 ± 0.11 |
| ② | √ | | | | 84.14 ± 0.20 | 72.92 ± 0.29 | 86.67 ± 0.81 | 95.06 ± 0.43 | 93.23 ± 0.21 |
| ③ | | √ | | | 84.04 ± 0.21 | 72.77 ± 0.28 | 86.66 ± 0.60 | 94.97 ± 0.30 | 93.18 ± 0.18 |
| ④ | | √ | √ | | 84.10 ± 0.20 | 72.88 ± 0.26 | 86.67 ± 0.40 | 95.05 ± 0.33 | 93.23 ± 0.19 |
| ⑤ | √ | √ | √ | | 84.18 ± 0.14 | 72.98 ± 0.33 | 86.76 ± 0.69 | **95.06 ± 0.07** | 93.25 ± 0.16 |
| ⑥ | √ | √ | √ | √ | **84.32 ± 0.18** | **73.18 ± 0.26** | **87.40 ± 0.81** | 94.90 ± 0.29 | **93.27 ± 0.17** |

*Note*: The bold values represent the best performances.

**TABLE 3** Comparison results on in-house dataset (mean ± standard deviation, the p-value of the t-test is about dice coefficient).

| Methods | Dice (%) | IoU (%) | Sen (%) | Spe (%) | Acc (%) | *p*-value |
|---|---|---|---|---|---|---|
| U-Net[15] | 79.49 ± 1.11 | 65.19 ± 1.42 | 80.98 ± 2.47 | 93.77 ± 0.75 | 90.74 ± 0.47 | <0.001 |
| Attention-UNet[30] | 79.99 ± 0.89 | 66.20 ± 1.11 | 81.39 ± 0.94 | 93.42 ± 0.61 | 90.97 ± 0.51 | <0.001 |
| UNet++[31] | 79.80 ± 0.77 | 66.28 ± 1.07 | 81.03 ± 1.19 | 93.52 ± 0.81 | 91.03 ± 0.56 | <0.001 |
| CPFNet[25] | 77.31 ± 0.79 | 63.02 ± 1.01 | 79.66 ± 1.81 | 92.62 ± 0.87 | 89.88 ± 0.59 | <0.001 |
| CS$^2$-Net[32] | 80.60 ± 0.72 | 67.11 ± 1.07 | 82.12 ± 0.94 | 93.68 ± 0.63 | 91.27 ± 1.07 | 0.007 |
| SP-Net[33] | 80.52 ± 0.79 | 67.14 ± 1.02 | 82.22 ± 1.18 | 93.78 ± 0.57 | **91.33 ± 0.47** | <0.001 |
| CE-Net[24] | 78.55 ± 0.74 | 64.54 ± 1.09 | 80.48 ± 1.22 | 93.05 ± 0.69 | 90.42 ± 0.55 | <0.001 |
| IF-Net[21] | 80.39 ± 0.81 | 66.71 ± 1.14 | 82.31 ± 1.34 | 93.64 ± 0.51 | 91.28 ± 0.53 | <0.001 |
| GD-Net[34] | 79.82 ± 0.87 | 66.52 ± 1.23 | 81.74 ± 1.40 | 93.43 ± 0.71 | 91.01 ± 0.57 | <0.001 |
| MLFGNet[35] | 80.19 ± 0.80 | 67.34 ± 1.08 | **83.79 ± 0.60** | 93.13 ± 0.63 | 91.18 ± 0.56 | <0.001 |
| SBD-MTLNet | **81.08 ± 0.63** | **67.93 ± 0.83** | 81.88 ± 0.90 | **93.87 ± 0.60** | 91.29 ± 0.51 | – |

*Note*: The bold values represent the best performances.

the comparison experiments on in-house dataset and public MGD-1K dataset respectively. Paired two-side T-test between the proposed SBD-MTLNet and other methods about Dice coefficient has been adopted for the statistical significance analysis.

As shown in Tables 3 and 4, our proposed SBD-MTLNet achieves the best performances about Dice and IoU indexes on both in-house and public datasets, reaching average Dice 81.08% and 84.32%, and average IoU 67.93% and 73.18%, respectively. All the p-values of T-test on Dice index are less than 0.01, showing that the proposed SBD-MTLNet significantly outperforms all other state-of-the-art networks. Attention-UNet[30] adopted the gated attention mechanism to suppress irrelevant areas in the input and highlight the significant features of specific local areas. However, this method does not introduce multi-scale information, which is very important for meibomian glands with various sizes. UNet++[31] combined long and short connections, which can extract features at different levels and integrate them by superposition. The addition of multiple shallow U-Net like structures reduced the scale difference during feature integration, while increasing the number of parameters. Similar

**TABLE 4** Comparison results on public MGD-1K dataset (mean ± standard deviation, the p-value of the t-test is about dice coefficient).

| Methods | Dice (%) | IoU (%) | Sen (%) | Spe (%) | Acc (%) | *p*-value |
|---|---|---|---|---|---|---|
| U-Net[15] | 83.29 ± 0.25 | 71.70 ± 0.35 | 86.11 ± 0.44 | 94.74 ± 0.16 | 92.88 ± 0.16 | <0.001 |
| Attention-UNet[30] | 83.95 ± 0.16 | 72.65 ± 0.24 | 86.26 ± 0.84 | 95.09 ± 0.21 | 93.18 ± 0.13 | <0.001 |
| UNet++[31] | 83.70 ± 0.13 | 72.29 ± 0.17 | 86.87 ± 0.61 | 94.70 ± 0.23 | 93.00 ± 0.16 | <0.001 |
| CPFNet[25] | 82.32 ± 0.27 | 70.31 ± 0.36 | 85.46 ± 0.60 | 94.38 ± 0.27 | 92.42 ± 0.19 | <0.001 |
| CS$^2$-Net[32] | 84.12 ± 0.18 | 72.89 ± 0.25 | 86.20 ± 0.59 | **95.21 ± 0.24** | 93.26 ± 0.17 | <0.001 |
| SP-Net[33] | 84.13 ± 0.21 | 72.91 ± 0.31 | 86.44 ± 0.46 | 95.14 ± 0.34 | 93.25 ± 0.21 | 0.001 |
| CE-Net[24] | 83.32 ± 0.20 | 71.73 ± 0.28 | 86.16 ± 0.54 | 94.71 ± 0.28 | 92.86 ± 0.19 | <0.001 |
| IF-Net[21] | 83.98 ± 0.20 | 72.68 ± 0.28 | 87.20 ± 0.55 | 94.75 ± 0.32 | 93.10 ± 0.19 | <0.001 |
| GD-Net[34] | 83.76 ± 0.23 | 72.38 ± 0.32 | 86.19 ± 0.70 | 94.98 ± 0.20 | 93.09 ± 0.16 | <0.001 |
| MLFGNet[35] | 84.04 ± 0.15 | 72.76 ± 0.20 | 86.34 ± 0.72 | 95.11 ± 0.13 | 93.21 ± 0.11 | <0.001 |
| SBD-MTLNet | **84.32 ± 0.18** | **73.18 ± 0.26** | **87.40 ± 0.81** | 94.90 ± 0.29 | **93.27 ± 0.17** | – |

*Note*: The bold values represent the best performances.

as Attention-UNet, UNet++ also lacks the multi-scale information. CE-Net [24] used several convolution branches with different receptive fields to obtain multi-scale information. Although the addition of multi-scale information can improve the model's attention to objects with different sizes, CE-Net does not pay attention to the boundary characteristics of glands. CPFNet[24] proposed a scale-aware pyramid fusion module, in which multi-scale information captured by three parallel dilated convolution layers with shared weights was dynamically fused by two scale-aware modules. However, CPFNet does not perform well in our meibomian gland segmentation task, maybe it could not capture the boundary information well. CS$^2$-Net[32] introduced asymmetric convolution into attention module to help the network focus on strip shape objects. However, it falls short in capturing gland boundary information. SP-Net[33] utilized strip pooling strategy in spatial pooling and directly applied interpolation operation to recover feature after pooling, thereby focusing solely on the strip region and reducing the influence of irrelevant information. While it can enhance the network's attention to the strip region, it may ignore channel information. GD-Net[34] adopted a multi-task learning strategy in which two different segmentation tasks share the encoder feature information and pixel-wise addition was applied to the output features from the two tasks in the decoder feature recovery stage, which may not be sufficient to fill in the information gap. Moreover, the directly pixel-wise addition may also introduce irrelevant information, adversely affecting the quality of the decoder feature. MLFGNet[35] designed multi-scale progressive guidance (MFPG) module, local feature guided attention (LFGA) module and multi-scale deep supervision (MDS) module for the curve-lined nerve fiber segmentation, which could not handle the gland adhesion problem well. The proposed SBD-MTLNet adopts the multi-task learning strategy to jointly learn meibomian gland segmentation and boundary detection, in which SMAM is designed for strip-shaped gland recognition and BDA-Net is for gland adhesion alleviation and both of them work well.

Figure 7 shows some meibomian gland segmentation results of different methods, in which the segmentation results of the proposed SBD-MTLNet are closest to the ground truth with less false positives and negatives. The yellow-boxed regions in Figure 7 are zoomed in and shown in Figure 8. As can be seen from Figure 8, the propose SBD-MTLNet can effectively alleviate the adhesion of the glands and improve the segmentation performance.

## 4.3 | Tortuosity analysis of meibomian gland

Based on the segmentation results of meibomian glands, the tortuosity analysis of the gland is performed, which can provide important information for the diagnosis of MGD. Previous studies show that the tortuosity of the meibomian gland is positively correlated with the grading of MGD. There are different definitions of the tortuosity.[11,17,36,37] In Ref. [17], as shown in Figure 9(a), point A and B represent the endpoints of the center line of the meibomian gland. Let path length represent the distance between A and B over the center line, and chord length represent the Euclidian distance between A and B. The tortuosity of the meibomian gland is defined as the ratio between the path length and the chord length.

In Ref. [36], the tortuosity is defined as follows,

$$MG\ Tortuosity = \frac{MG\ perimeter}{2 \times H} - 1 \tag{17}$$

in which MG perimeter represents the number of pixels at the edge of the meibomian gland, $H$ represent the height of the minimum external rectangle of the MG, as shown in Figure 9(a).

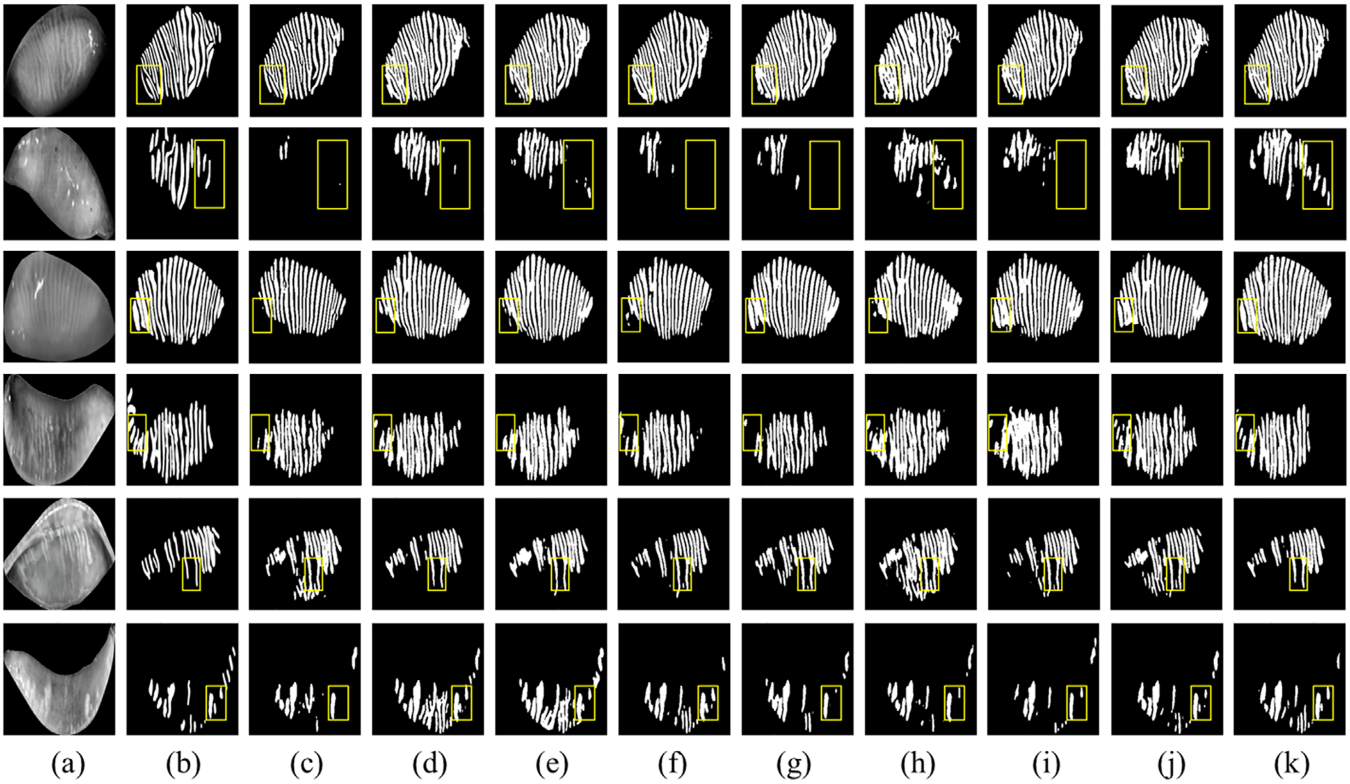


**FIGURE 7** Segmentation results of different methods. (a) Original image; (b) ground truth; (c) UNet; (d) CS2-Net; (e) IF-Net; (f) SP-Net; (g) attention-UNet; (h) CE-Net; (i) GD-Net; (j) MLFGNet; (k) the proposed SBD-MTLNet.

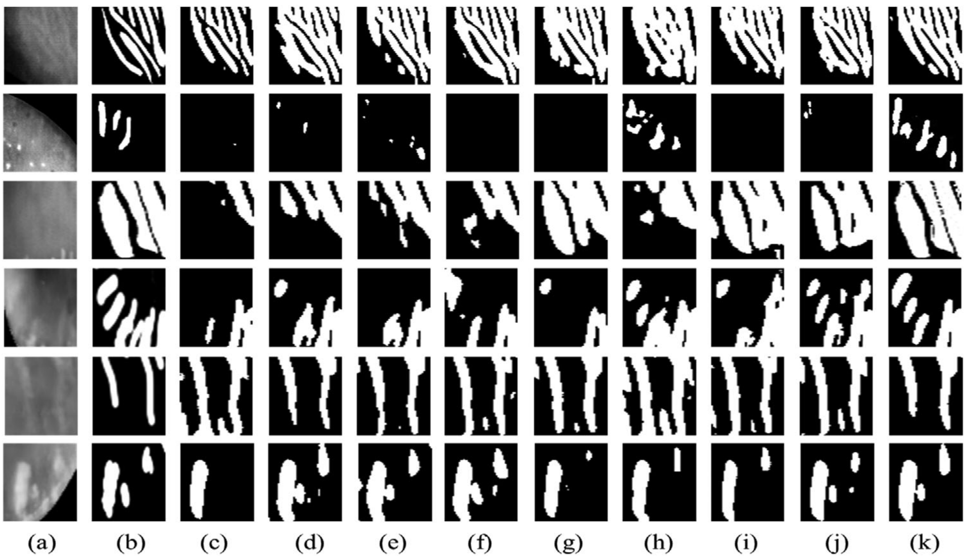


**FIGURE 8** Zoom-in display of the yellow-boxed regions in Figure 7. (a) Original image; (b) ground truth; (c) UNet; (d) CS2-Net; (e) IF-Net; (f) SP-Net; (g) attention-UNet; (h) CE-Net; (i) GD-Net; (j) MLFGNet; (k) the proposed SBD-MTLNet.

In Ref. [37], gland tortuosity index (TI) is defined as the ratio of the cavity area to the corresponding external convex hull area of the gland, which represents the concave-convex defect of the gland. The greater the value of the tortuosity index, the more serious the concave-convex defect of gland. As shown in Figure 9(b), gland tortuosity index (TI) can be calculated as,

$$TI = \frac{S_1}{S_2 + S_1} \times 100\% \tag{18}$$

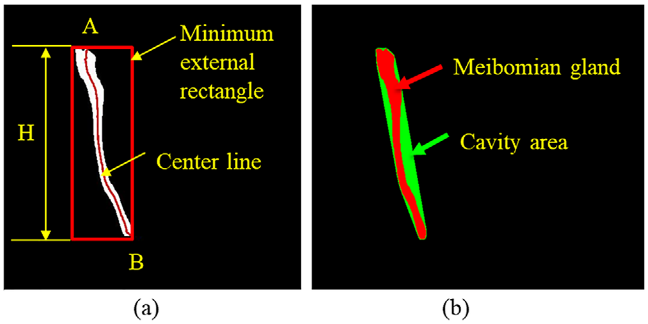


**FIGURE 9** Different definitions of meibomian gland tortuosity index. (a) Lengths for calculation of tortuosity; (b) Areas for calculation of tortuosity, meibomian glands (red area), cavities (green area), convex hull (red + green area).

**TABLE 5** The criterion for meibomian gland distortion screening.

| The tortuosity index of meibomian gland (%) | Whether distorted (1/0) |
|---|---|
| <45 | 0 |
| ≥45 | 1 |

where, $S_1$ represents area of the cavity region of the meibomian gland (green area in Figure 9(b)), $S_2$ represents area of the meibomian gland (red area in Figure 9(b)), and $S_1 + S_2$ represents the area of the corresponding external convex hull of the gland.

In this paper, we adopt Equation (18) to calculate the gland tortuosity index (TI). Inspired by the study of Arita et al.,[38] we adopt a criterion for the gland distortion screening, that is if the tortuosity index (TI) exceeds 45%, it will be considered as distorted (shown in Table 5). This distortion screening criterion can reflect not only the distortion degree of the meibomian gland, but also the gland segmentation performance. Based on this, meibomian gland distortion screening experiments on both in-house dataset (249 upper eyelids) and public dataset (1000 images) are conducted, in which the TI index calculation and distortion screening are performed on the gland segmentation ground truth and the segmented glands by different methods one by one. Table 6 shows the distortion screening accuracy on both datasets, in which the proposed SBD-MTLNet outperforms other state-of-the-art methods, reflecting the effectiveness of its segmentation performance from another point of view.

**TABLE 6** Comparison results of distortion screening on in-house dataset (249 upper eyelids) and public dataset.

| Methods | Acc (%) In-house dataset (249 upper eyelids) | Acc (%) Public dataset |
|---|---|---|
| U-Net | 86.44 | 91.00 |
| Attention-UNet | 87.35 | 90.60 |
| UNet++ | 84.98 | 91.20 |
| CPFNet | 86.55 | 91.40 |
| CE-Net | 85.93 | 90.00 |
| IF-Net | 87.75 | 90.10 |
| SP-Net | 85.53 | 91.00 |
| CS$^2$-Net | 85.74 | 90.10 |
| GD-Net | 84.69 | 91.00 |
| MLFGNet | 85.61 | 91.30 |
| SBD-MTLNet | **89.55** | **92.00** |

*Note*: The bold values represent the best performances.

## 5 | DISCUSSION AND CONCLUSIONS

In this paper, we propose a strip and boundary detection multi-task learning network SBD-MTLNet to realize the automatic segmentation of meibomian glands in near-infrared meibography images. SMAM is proposed to combine local and global spatial features from vertical and horizontal dimensions to improve the network's recognition capacity of strip-shaped glands. With the multi-task learning strategy, BDA-Net is proposed to effectively utilize the boundary information and relieve the problem of adhesion between glands. SIIM based on reverse attention mechanism is proposed to realize information complementation between meibomian gland segmentation and boundary detection tasks. MSAM is embedded at the top of the encoder path to extract rich multi-scale information. The proposed SBD-MTLNet has been evaluated on both in-house dataset and public dataset. Ablation experiments show that all the proposed modules and multi-task learning strategy contribute to the final segmentation performance. Compared with strip shape object segmentation networks such as CS$^2$-Net,[32] SP-Net[33] and MLFGNet,[35] method with multi-task learning strategy such as GD-Net,[34] and other state-of-the-art U-shape segmentation networks,[15,21,24,25,30,31] the proposed SBD-MTLNet achieves superior performance. The comprehensive experimental results show the effectiveness of the proposed SBD-MTLNet.

There are some limitations in this work. First, because only 1 × 3 and 3 × 1 convolutions are used as strip-shaped kernels in the proposed SMAM, it may lead the network to focus more on the horizontal and vertical strip-shaped areas and result in poor performance on meibomian glands with other angles. We will focus on orientation agnostic or angle optimization method inspired by angel correct module (ACM)[39] for the further improvement of segmentation performance. Second, as can be seen from Tables 1–4, the performances of the proposed SBD-MTLNet on the public MGD-1K dataset (1000 images) are better than on the in-house dataset (453 images). As can be seen from Figure 6, the image contrast of the in-house dataset is lower than that of the public one, which is a general phenomenon between these two datasets. In addition, the bright reflection interference of the in-house dataset is more severe than that of the public one. These two image quality factors are the leading causes of the poorer performance on the in-house dataset. Additionally, the insufficient number of meibography images may limit the performance of the proposed method. In future work, we plan to collect more near-infrared meibography images from multiple medical centers and focus on the processing of images with poor quality to further improve the segmentation performance and the generalization of the proposed SBD-MTLNet. Based on the segmentation results, we will focus on analyzing the correlation between the atrophy degree of meibomian gland and its morphological characteristics including tortuosity, length, width and count, etc. We are also focusing on the meibomian gland segmentation and tortuosity analysis based MGD grading in our follow-up work, which has important potential for the auxiliary diagnosis of MGD and DED in clinic.

## ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China (62371326, U20A20170, 62271337, 62371328), the Natural Science Foundation of Jiangsu Province of China (BK20231310), the National Key Research and Development Program of China (2018YFA0701700), China Postdoctoral Science Foundation (2023M742568), and a project funded by the Priority Academic Program Development of Jiangsu Higher Education Institutions (YX11900123).

## CONFLICT OF INTEREST STATEMENT

The authors have no relevant conflicts of interest to disclose.

## DATA AVAILABILITY STATEMENT

The in-house dataset underlying the results presented in this paper is not publicly available at this time but may be obtained from the authors upon reasonable request.

## REFERENCES

1. Butovich I. Meibomian glands, meibum, and meibogenesis. *Exp Eye Res*. 2017;163:2-16.
2. Verrma S, Moreno I, Trapp M, et al. Meibomian gland development: where, when and how? *Differentiation*. 2023;132:41-50.
3. Pucker A, Haworth K. The presence and significance of polar meibum and tear lipids. *Ocular Surf*. 2015;13(1):26-42.
4. Wang J, Call M, Mongan M, et al. Meibomian gland morphogenesis requires developmental eyelid closure and lid fusion. *Ocular Surf*. 2017;15(4):704-712.
5. Sabeti S, Kheirkhah A, Yin J, et al. Management of meibomian gland dysfunction: a review. *Surv Ophthalmol*. 2020;65(2):205-217.
6. Wang Y, Li S, Chen X, et al. Meibomian gland absence related dry eye in ectodysplasin a mutant mice. *Am J Pathol*. 2016;186(1):32-42.
7. Chhadva P, Goldhardt R, Galor A. Meibomian gland disease: the role of gland dysfunction in dry eye disease. *Ophthalmology*. 2017;124(11S):S20-S26.
8. Arita R, Suehiro J, Haraguchi T, et al. Objective image analysis of the meibomian gland area. *Br J Ophthalmol*. 2014;98(6):746-755.
9. Koh Y, Celik T, Lee H, et al. Detection of meibomian glands and classification of meibography images. *J Biomed Opt*. 2012;17(8):086008.
10. Koprowski R, Wilczyński S, Olczyk P, et al. A quantitative method for assessing the quality of meibomian glands. *Comp Biol Med*. 2016;75:130-138.
11. Xiao P, Luo Z, Deng Y, et al. An automated and multiparametric algorithm for objective analysis of meibography images. *Quant Imag Med Surg*. 2021;11(4):1586-1599.
12. Wang J, Yeh T, Chakraborty R, et al. A deep learning approach for meibomian gland atrophy evaluation in meibography images. *Transl Vis Sci Technol*. 2019;8(6):37.
13. Khan Z, Umar A, Shirazi S, et al. Image based analysis of meibomian gland dysfunction using conditional generative advesarial neural network. *BMJ Open Ophthalmol*. 2021;6:e000436.
14. Swiderska K, Blackie C, Maldonado-Codina C, et al. A deep learning approach for meibomian gland appearance evaluation. *Ophthalmol Sci*. 2023;3:100334.
15. Ronneberger O, Fischer P, Brox T, U-Net: convolutional networks for biomedical image segmentation. Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention. 2015:234-241.
16. Prabhu S, Chakiat A, Shashank S, et al. Deep learning segmentation and quantification of meibomian glands. *Biomed Signal Proc Control*. 2020;57:101776.
17. Setu M, Horstmann J, Schmidt S, et al. Deep learning-based automatic meibomian gland segmentation and morphology assessment in infrared meibography. *Sci Rep*. 2021;11(1):7649.
18. Dai Q, Liu X, Lin X, et al. A novel meibomian gland morphology analytic system based on a convolutional neural network. *IEEE Access*. 2021;9:23083-23094.
19. Saha R, Chowdhury A, Na K, et al. Automated quantification of meibomian gland dropout in infrared meibography using deep learning. *Ocular Surf*. 2022;26:283-294.
20. Wang Y, Shi F, Wei S, et al. A deep learning model for evaluating meibomian glands morphology from meibography. *J Clin Med*. 2023;12:1053.
21. Liu D, Zhu W, Zhuang X, et al. IF-Net: information fusion network for meibomian gland area and atrophy area segmentation. Proceedings of Medical Imaging 2023: Image Processing, SPIE. 2023:124643X.
22. Wang Q, Wu B, Zhu P, et al. ECA-Net: efficient channel attention for deep convolutional neural networks. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:11531-11539.
23. Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. 2017:6230-6239.
24. Gu Z, Cheng J, Fu H, et al. CE-Net: context encoder network for 2D medical image segmentation. *IEEE Trans Med Imag*. 2019;38(10):2281-2292.
25. Feng S, Zhao H, Shi F, et al. CPFNet: context pyramid fusion network for medical image segmentation. *IEEE Trans Med Imag*. 2020;39(10):3008-3018.
26. Chen S, Tan X, Wang B, et al. Reverse attention for salient object detection. Proceedings of European Conference on Computer Vision. 2018:236-252.
27. Xia L, Zhang H, Wu Y, et al. 3D vessel-like structure segmentation in medical images by an edge-reinforced network. *Med Image Anal*. 2022;82:102581.
28. Kim T, Lee H, Kim D, UACANet: uncertainty augmented context attention for polyp segmentation. Proceedings of 29th ACM International Conference on Multimedia. 2021:2167-2175.
29. Yushkevich P, Piven J, Hazlett H, et al. User-guided 3D active contour segmentation of anatomical structures: significantly

improved efficiency and reliability. *Neuroimage*. 2006;31(3):1116-1128.
30. Oktay O, Schlemper J, Folgoc L, et al. Attention U-Net: learning where to look for the pancreas. Proceedings of International Conference on Medical Imaging with Deep Learning. 2018:1-10.
31. Zhou Z, Siddiquee M, Tajbakhsh N, et al. UNet++: a nested U-Net architecture for medical image segmentation. 2018:3-11.
32. Mou L, Zhao Y, Fu H, et al. $CS^2$-Net: deep learning segmentation of curvilinear structures in medical imaging. *Med Image Anal*. 2020;67:101874.
33. Hou Q, Zhang L, Cheng M, et al. Strip pooling: rethinking spatial pooling for scene parsing. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020:4003-4012.
34. Yao C, Wang M, Zhu W, et al. Joint segmentation of multi-class hyper-reflective foci in retinal optical coherence tomography images. *IEEE Trans Biomed Eng*. 2021;69(4):1349-1350.
35. Tang W, Chen X, Yuan J, et al. Multi-scale and local feature guidance network for corneal nerve fiber segmentation. *Phys Med Biol*. 2023;68:095026.
36. Lin X, Fu Y, Li L, et al. A novel quantitative index of meibomian gland dysfunction, the meibomian gland tortuosity. *Transl Vis Sci Technol*. 2020;9:34.
37. Wang J, Li S, Yeh T, et al. Quantifying meibomian gland morphology using artificial intelligence. *Opt Vis Sci*. 2021;98(9):1094-1103.
38. Arita R, Itoh K, Maeda S, et al. *Association of contact lens-related allergic conjunctivitis with changes in the morphology of meibomian glands*. Japanese Ophthalmological Society; 2012:14-19.
39. Xu H, Liu X, Xu H, et al. Rethinking boundary discontinuity problem for oriented object detection. Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024:17406-17415.

**How to cite this article:** Zhu W, Liu D, Zhuang X, et al. Strip and boundary detection multi-task learning network for segmentation of meibomian glands. *Med Phys*. 2025;52:1615–1628.
https://doi.org/10.1002/mp.17542