


Article

PFSegIris: Precise and Fast Segmentation Algorithm for Multi-Source Heterogeneous Iris

Lin Dong ^{1,2} , Yuanning Liu ^{1,2} and Xiaodong Zhu ^{1,2,*}

¹ College of Computer Science and Technology, Jilin University, Changchun 130012, China; donglin19@mails.jlu.edu.cn (L.D.); lyn@jlu.edu.cn (Y.L.)

² Key Laboratory of Symbolic Computation and Knowledge Engineering of Ministry of Education, Jilin University, Changchun 130012, China

* Correspondence: zhuxd@jlu.edu.cn; Tel.: +86-431-8515-9376

Abstract: Current segmentation methods have limitations for multi-source heterogeneous iris segmentation since differences of acquisition devices and acquisition environment conditions lead to images of greatly varying quality from different iris datasets. Thus, different segmentation algorithms are generally applied to distinct datasets. Meanwhile, deep-learning-based iris segmentation models occupy more space and take a long time. Therefore, a lightweight, precise, and fast segmentation network model, PFSegIris, aimed at the multi-source heterogeneous iris is proposed by us. First, the iris feature extraction modules designed were used to fully extract heterogeneous iris feature information, reducing the number of parameters, computation, and the loss of information. Then, an efficient parallel attention mechanism was introduced only once between the encoder and the decoder to capture semantic information, suppress noise interference, and enhance the discriminability of iris region pixels. Finally, we added a skip connection from low-level features to catch more detailed information. Experiments on four near-infrared datasets and three visible datasets show that the segmentation precision is better than that of existing algorithms, and the number of parameters and storage space are only 1.86 M and 0.007 GB, respectively. The average prediction time is less than 0.10 s. The proposed algorithm can segment multi-source heterogeneous iris images more precisely and quicker than other algorithms.

Keywords: iris segmentation; heterogeneous iris; fast segmentation; iris recognition



Citation: Dong, L.; Liu, Y.; Zhu, X. PFSegIris: Precise and Fast Segmentation Algorithm for Multi-Source Heterogeneous Iris. *Algorithms* **2021**, *14*, 261. <https://doi.org/10.3390/a14090261>

Academic Editor: Girija Chetty

Received: 9 August 2021

Accepted: 28 August 2021

Published: 30 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the advantages of uniqueness, stability, and security of iris texture, iris recognition [1] stands out among numerous biometric recognition technologies [2]. The iris recognition process includes iris image acquisition, iris preprocessing, and iris feature extraction and recognition [3,4]. Iris segmentation [5] is the accurate location of the iris region in the whole image, which plays a decisive role in the subsequent iris feature expression and recognition rate and is an important step in the entire iris recognition process. Due to the quality differences caused by factors such as heterogeneous spectrum (visible and near-infrared light) and different iris region sizes, the fast and accurate segmentation of multi-source heterogeneous iris images brings new challenges to iris recognition.

The main traditional iris segmentation methods are the calculus circle template method proposed by Daugman [6] and the Hough method based on edge detection [7]. Many subsequent algorithms have been improved and innovated based on these two methods. Traditional iris segmentation methods can only accurately segment the iris region under ideal conditions. However, conditions are often non-ideal in practice, e.g., due to noise interference from upper and lower eyelid occlusion, eyelash occlusion, and small grayscale changes in the outer boundary of the iris. Traditional iris segmentation methods are susceptible to these noises, leading to a great reduction in localization accuracy and thus affecting the recognition results.

With the development of deep learning in recent years, semantic segmentation methods based on deep neural networks have been widely used in iris segmentation. Semantic segmentation methods at present can be mainly divided into two kinds. One is based on region classification, represented by Mask-RCNN [8]. The working principle is to combine target detection and semantic segmentation. First, the target region is selected through target detection algorithms such as Fast-RCNN [9] and Faster-RCNN [10]. Then, a classifier such as SVM is used to classify the region to perform pixel classification. However, since global semantic information is not considered, segmentation accuracy will decrease in a slightly complex background. Another is semantic segmentation based on end-to-end form, avoiding the problems caused by the generation of the target region. A deep neural network is used to extract semantic information from images with annotation, learn and infer the pixel category of the original images based on the information, and classify each pixel in order to achieve the goal of semantic segmentation.

The existing end-to-end iris semantic segmentation methods are mostly based on several typical convolutional neural networks, such as FCN [11], SegNet [12], and U-Net [13]. Muhammad et al. [14] proposed a fully residual network model FRED-Net, which segmented the true iris boundary without the extra cost of denoising as a preprocessing step. Zhou et al. [15] proposed a neural network model, PI-Unet, for iris segmentation to segment heterogeneous iris images. Wang et al. [16] put forward a deep-learning-based iris segmentation method, IrisParseNet. A multi-task, joint approach was adopted to learn iris features in order to solve the noise problem of iris images. Li et al. [17] proposed a neural network, IRUNet, for robust iris segmentation in a non-cooperative environment to locate the inner and outer iris boundaries, respectively. You et al. [18] proposed an iris segmentation method, MFFIris-Unet, for heterogeneous noise, distinguishing noises and iris targets. In summary, the iris segmentation algorithms based on deep learning improve the shortcomings of the traditional iris segmentation algorithms to a certain extent and enhance segmentation accuracy. However, the current deep-learning-based iris segmentation network models incur a high cost in terms of a large parameter search space and a long segmentation time, have requirements for hardware devices, and perform poorly in multi-source heterogeneous iris segmentation.

Targeting these problems and motivated by the above observations, we propose a precise and fast segmentation network model, PFSegIris, for multi-source heterogeneous iris images that can accurately segment iris regions of different sizes; weaken the influence of different spectra and eyelid, eyelash occlusion noises; and enhance the discriminative ability of iris region pixels, thereby, having a better universality for iris images collected by different devices. First, the iris feature extraction module designed in this paper was embedded in each layer of PFSegIris to extract rich iris information, in which convolution kernels of different sizes obtain the information of iris regions of different sizes, reducing the amount of parameters and meanwhile taking accuracy into account. Then, an efficient parallel position and channel attention mechanism [19] was used, once in the middle, to focus on the iris region pixels of small targets at different spectra. Finally, a skip connection was added to make details such as edge features more precise. Images from one near-infrared iris database, JLU-6.0, and two visible iris databases, CASIA-Iris-Lamp-v4 and MICHE-I, were mixed for training. Several experiments were carried out for testing on four near-infrared iris databases, JLU-6.0 [20], JLU-7.0 [20], CASIA-Iris-Interval-v4 [21], and Mmu2 [22], and three visible iris databases, CASIA-Iris-Lamp-v4 [21], MICHE-I [23,24], and UBIRIS.V2 [25–27]. The test results show that the iris segmentation algorithm proposed can realize fast and precise segmentation for multi-source heterogeneous iris images.

Our main contributions can be summarized as follows:

1. Different from traditional methods and other iris segmentation algorithms based on deep learning, a more precise segmentation algorithm, PFSegIris, was designed to segment multi-source heterogeneous irises without any preprocessing or post-processing. We proved that accurate iris segmentation can be realized for images

with heterogeneous spectra (visible and near-infrared), different iris region sizes, and uneven quality.

2. While ensuring accuracy, PFSegIris is lightweight, significantly reducing the number of parameters, storage space, and calculations in comparison with other methods, and the average prediction time on seven heterogeneous databases was less than 0.10 s, which is a fast iris segmentation algorithm. In addition, the algorithm proposed can be applied to devices with low computing performance and storage capacity.

Furthermore, the remainder of this paper is organized as follows. Section 2 presents the proposed PFSegIris structure in detail. Experimental results and comparisons are described in Section 3. Finally, Section 4 is the conclusion.

2. Methods

The overview of the proposed PFSegIris is shown in Figure 1. The overall process was as follows:

1. An original iris image was directly input to the network model without any preprocessing. The encoder downsampled the image until it could be represented by its tiny features, which can fully constitute iris information of various sizes.
2. The parallel position-channel attention mechanism was used to enhance the feature representation of iris region and suppress the influence of noise on iris segmentation.
3. The decoder upsampled the iris image to its original dimensions and predicted the pixels of the iris region and the non-iris region.

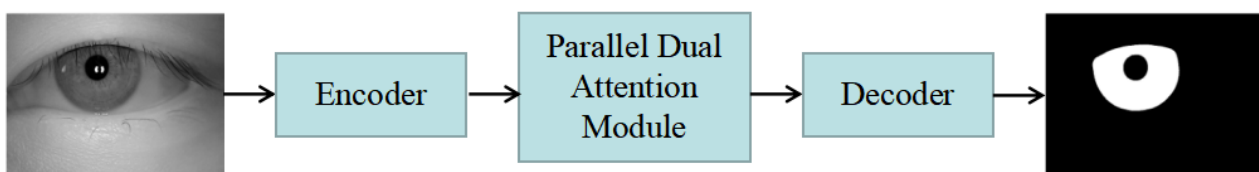


Figure 1. Overview of PFSegIris.

2.1. Encoder

The function of an encoder is to extract information about the location, shape, and size of iris regions in multi-source heterogeneous iris images. Only when the encoder is able to extract accurate and detailed iris features will the iris segmentation results be good. Therefore, designing a suitable encoder for iris segmentation is of great importance to the network model.

The encoder designed in this paper included four blocks, and each block included two different iris feature extraction modules: iris-feature module1 and iris-feature module2, which are shown in Figure 2a,b, respectively. The iris-feature modules include four layers: a mixed depthwise convolution, a pointwise convolution, a pointwise convolution, and a depthwise convolution, in which no activation function was added after the second and the fourth layers.

Depthwise convolutions with kernels of different sizes encode the high-dimensional features of the iris image, generating rich representations, which is superior to the traditional depthwise convolution suffering from the limitations of a single kernel size [28]. The 3×3 convolution kernels have a strong ability to perceive image details and can fully perceive the iris in a small area, while the 5×5 convolution kernels have a larger receptive field and can capture iris information in a larger area. Two pointwise convolutions were used to encode information between channels, with channel reduction followed by expansion. The depthwise convolution at the end can learn more abundant iris spatial information and, thereby, exhibits better segmentation performance. In order to obtain more information from the bottom layers and promote the propagation of gradient information cross layers, shortcuts were used to connect high-dimensional representations with a large number of channels and help the network avoid degradation as the depth increases. Table 1 shows

the structure details of the iris-feature module. The structure of the encoder was able to remarkably reduce the number of parameters and calculations of the model while taking the segmentation accuracy into account, which greatly reduced the computation cost.

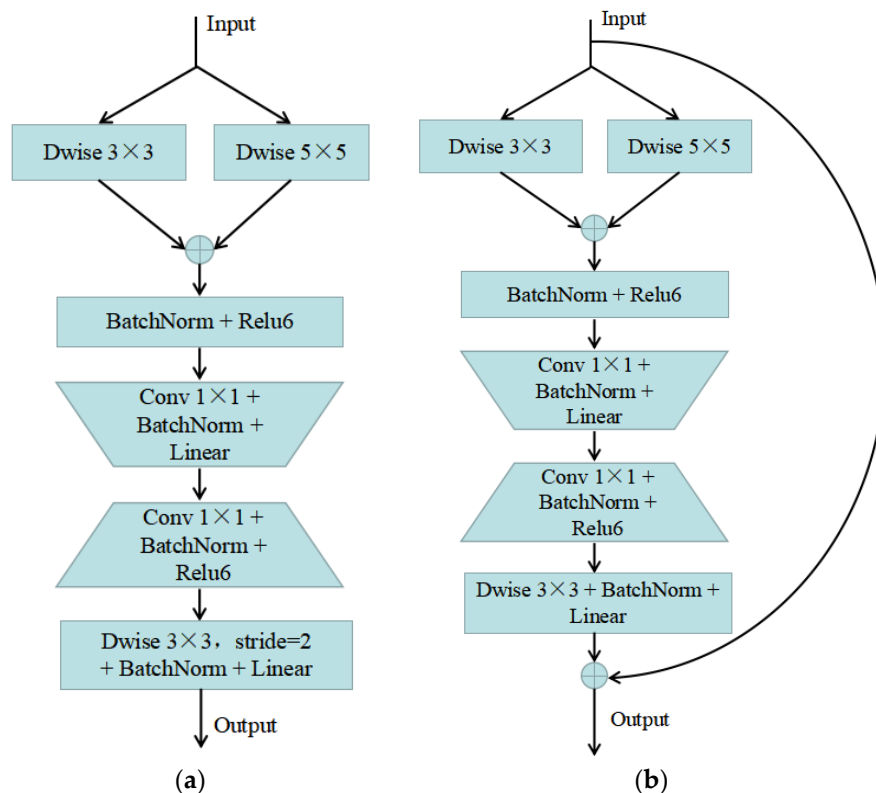


Figure 2. Two different iris feature extraction modules: (a) iris-feature module1; (b) iris-feature module2.

Table 1. The structure details of the iris-feature modules. t : channel reduction factor; s : stride; C : number of input channels; C' : number of output channels.

Input Dimension	Operator Type	Output Dimension
$H \times W \times C$	Dwise 3×3 Dwise 5×5	Add, BatchNorm, Relu6 $H \times W \times C$
$H \times W \times C$	Conv 1×1 , BatchNorm, Linear	$H \times W \times \frac{C}{t}$
$H \times W \times \frac{C}{t}$	Conv 1×1 , BatchNorm, Relu6	$H \times W \times C'$
$H \times W \times C'$	Dwise 3×3 , stride = s , BatchNorm, Linear	$\frac{H}{s} \times \frac{W}{s} \times C'$

2.2. Parallel Dual Attention Mechanism

There are always two problems in segmentation for a multi-source heterogeneous iris:

1. due to the influence of different factors such as spectra and noises, the prediction results of some pixels are likely to be affected; and
2. the sizes of the iris regions collected by different devices differ greatly, but the features of different scales should be treated equally.

PFSegIris makes further improvement on segmentation effects by adaptively integrating similar features at any scale from a global view through an efficient parallel dual attention mechanism. The parallel dual attention module consisted of a position attention module and a channel attention module, which assigned weights to the importance of information in both position and channel dimensions. The position attention module was used

to learn the spatial dependence of features, while the channel attention module was used to learn the internal correlation between channels. This form of dual attention improved the discrimination of pixels in small iris regions, suppressed the interference of irrelevant information such as spectra and noises, and helped reduce gradient disappearance while improving the capability and efficiency of iris feature representation.

2.2.1. Position Attention Module (PAM)

The position attention module is able to encode broader contextual information into local features so as to enhance the representation ability of iris features. Figure 3 shows the structure of the position attention module. The position attention module was calculated according to Equations (1) and (2).

$$s_{ji} = \frac{\exp(B_i \cdot C_j)}{\sum_{i=1}^N \exp(B_i \cdot C_j)} \quad (1)$$

$$E_j = \alpha \sum_{i=1}^N (s_{ji} D_i) + A_j \quad (2)$$

where A is the input of the module, which is fed into three convolution layers to generate three new feature maps B , C , and D , with the same dimension $C \times H \times W$. Then we reshaped B and C to dimension $C \times N$, $N = H \times W$. After that we performed a matrix multiplication between the transpose of B and C to obtain s with dimensions $N \times N$, and applied a SoftMax operation. s_{ji} is the matrix element of the j th row and the i th column, representing the impact of the i th position on the j th position. D was reshaped to dimension $C \times N$. We performed a matrix multiplication between D and s , reshaping the result to the dimensions $C \times H \times W$. A sum operation was added with the feature map A to get the output E , where α is a scale parameter, initialized as 0 and learning to assign more weight gradually in the training process [29].

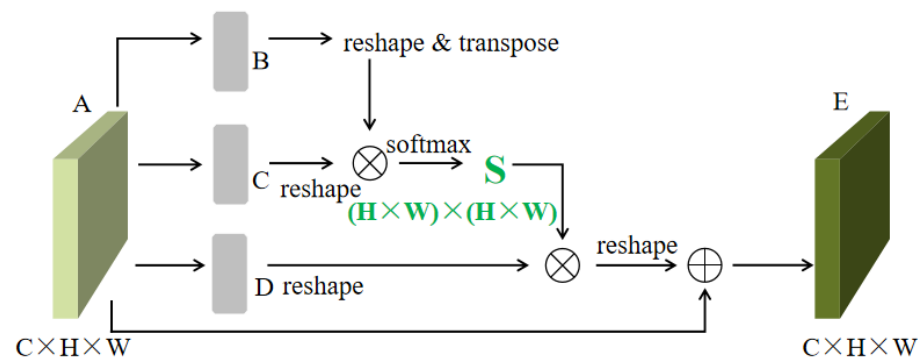


Figure 3. The structure of the position attention module.

It can be seen from Equations (1) and (2) that the resultant feature at each position was the weighted sum of features from all positions and original features. Therefore, the PAM had a global view of iris images and selectively fused the contexts based on the spatial attention maps. Similar semantic features were integrated and benefited each other, which promotes intra-class feature compactness and semantic consistency.

2.2.2. Channel Attention Module (CAM)

Each channel map of high-level iris features can be viewed as a response to a specific class, and different semantic responses are interrelated. By taking advantage of the interdependencies between channel maps, the interdependent feature maps can be enhanced, and the semantic-specific feature representation can be improved. Figure 4 shows the structure

of the channel attention module. The channel attention module was calculated according to Equations (3) and (4).

$$x_{ji} = \frac{\exp(A_i \cdot A_j)}{\sum_{i=1}^C \exp(A_i \cdot A_j)} \quad (3)$$

$$E_j = \beta \sum_{i=1}^C (x_{ji} A_i) + A_j \quad (4)$$

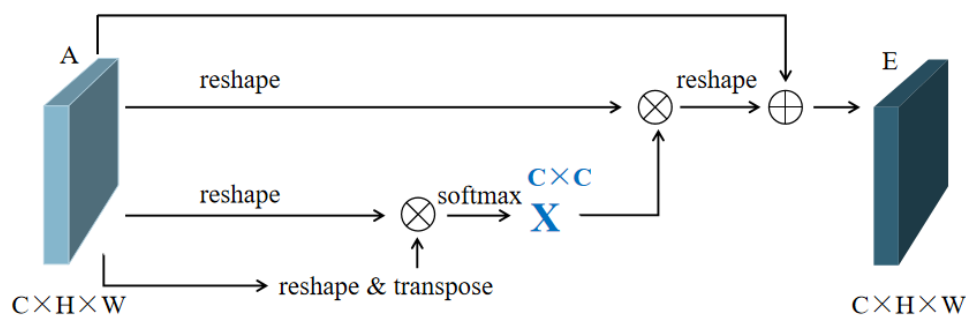


Figure 4. The structure of the channel attention module.

The difference from the PAM was that the calculation started directly on the original features so that the relationship between different channel maps could be maintained. The channel attention map dimension was obtained as $C \times C$ after applying a SoftMax layer. β was also a scale parameter, initialized to 0, and the optimal utilization degree was sought continuously with training. Equations (3) and (4) show that the output feature of each channel was the weighted data of all channel features and original features, which established the semantic dependencies between feature maps and helped to improve the distinguishability of iris features.

2.3. Decoder

The decoder's function is to convert the multi-source heterogeneous iris feature information extracted by the encoder into iris semantic information. The decoder performs operations such as upsampling and convolution on the output feature map of the encoder and finally classifies each pixel on the iris image using functions such as SoftMax in order to distinguish pixels in the iris region from those in the non-iris region. Three kinds of upsampling methods are commonly used in decoders for semantic segmentation, unpooling, interpolation, and transposed convolution. Unpooling is a non-linear upsampling method introduced in SegNet, which corresponds to the maxpooling operation used in the encoder for recording the location of the maximum value. Although the calculation time is reduced, a certain amount of memory space is occupied when storing the location information. Interpolation can reduce the amount of computation, but at the same time, the accuracy of segmentation decreases. Since the parameters in the transposed convolution can be learned, the number of parameters and the amount of computation will increase to a certain extent, but the segmentation effect will be more accurate.

Our decoder adopts four transposed convolution operations, which corresponds to the depthwise convolutions with stride of 2 used in the encoder four times for downsampling. The transposed convolution operation can improve iris segmentation accuracy, and it was combined with the structure of iris-feature module2. Finally, only one skip connection was used to fuse the low-level features to enrich the detailed information, further improving the iris edge segmentation result.

The network structure of PFSegIris is shown in Figure 5. Table 2 presents the network structure descriptions of PFSegIris.

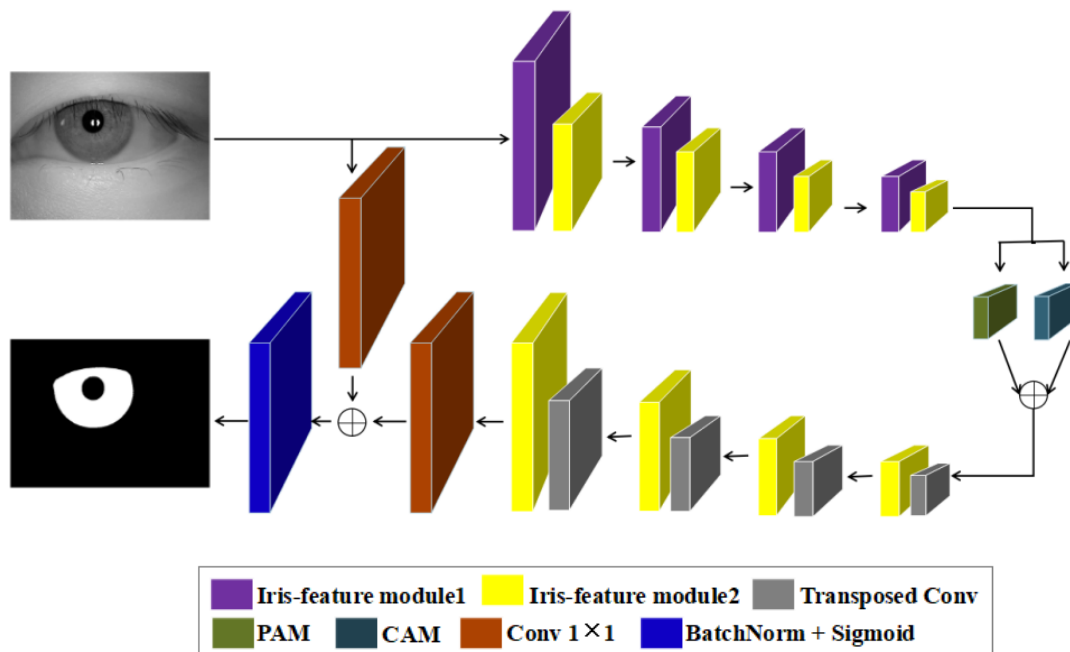


Figure 5. The network structure of PFSegIris.

Table 2. Network structure descriptions of PFSegIris. The input image dimension of $480 \times 640 \times 3$ are taken as an example.

Input Dimension	Operator Type	t	Stride	Output Dimension
$480 \times 640 \times 3$	Iris-feature module1	2	2	$240 \times 320 \times 64$
$240 \times 320 \times 64$	Iris-feature module2	6	1	$240 \times 320 \times 64$
$240 \times 320 \times 64$	Iris-feature module1	6	2	$120 \times 160 \times 128$
$120 \times 160 \times 128$	Iris-feature module2	6	1	$120 \times 160 \times 128$
$120 \times 160 \times 128$	Iris-feature module1	6	2	$60 \times 80 \times 256$
$60 \times 80 \times 256$	Iris-feature module2	6	1	$60 \times 80 \times 256$
$60 \times 80 \times 256$	Iris-feature module1	6	2	$30 \times 40 \times 512$
$30 \times 40 \times 512$	Iris-feature module2	6	1	$30 \times 40 \times 512$
$30 \times 40 \times 512$	PAM, CAM			$30 \times 40 \times 512$
$30 \times 40 \times 512$	Add			$30 \times 40 \times 512$
$30 \times 40 \times 512$	Transposed Conv		2	$60 \times 80 \times 256$
$60 \times 80 \times 256$	Iris-feature module2	6	1	$60 \times 80 \times 256$
$60 \times 80 \times 256$	Transposed Conv		2	$120 \times 160 \times 128$
$120 \times 160 \times 128$	Iris-feature module2	6	1	$120 \times 160 \times 128$
$120 \times 160 \times 128$	Transposed Conv		2	$240 \times 320 \times 64$
$240 \times 320 \times 64$	Iris-feature module2	6	1	$240 \times 320 \times 64$
$240 \times 320 \times 64$	Transposed Conv		2	$480 \times 640 \times 1$
$480 \times 640 \times 1$	Iris-feature module2	6	1	$480 \times 640 \times 1$
$480 \times 640 \times 1$	Conv 1×1		1	$480 \times 640 \times 1$
$480 \times 640 \times 1$	Add			$480 \times 640 \times 1$
$480 \times 640 \times 1$	BN + Sigmoid			$480 \times 640 \times 1$

2.4. Positioning of Iris Inner and Outer Circles

The positioning process for the inner and outer circles of the iris used the mask images output from the proposed PFSegIris, which accurately divided the effective region of iris and contained the position and radius information for the iris and pupil. In this paper, the contour detection method [30] was used to localize the inner and outer circles of the iris, and the coordinate and radius of the iris and pupil area were calculated. The difference between the localization algorithm in this paper and other iris localization algorithms is that the detection of the inner and outer circles was performed directly on the binary iris

mask images rather than on original iris images, avoiding the interference of eyelashes, eyelids, and lighting. The iris mask localization image is shown in Figure 6a, and its corresponding original iris localization image is shown in Figure 6b.

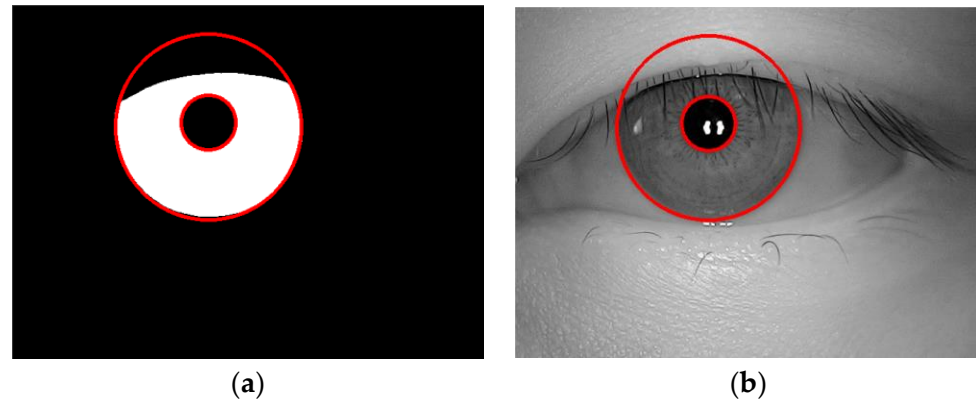


Figure 6. Iris mask localization image and its corresponding original iris localization image: (a) mask localization image; (b) original iris localization image.

3. Experiments and Analysis

3.1. Experimental Details

The proposed multi-source heterogeneous iris fast segmentation network, PFSegIris, was trained and tested using the TensorFlow 2.0 deep learning framework. We used an NVIDIA GTX 1080Ti GPU for training, and we trained the PFSegIris for 40 epochs. The mini-batch size was 4. Training was performed from scratch with original images without any pre-trained model. The optimizer was SGD with the initial learning rate of 0.01 and 0.9 momentum, and the standard weight decay was set to 0.0005.

3.2. Iris Datasets and Data Augmentation

A total of seven different iris datasets were used in the experiments, four of which were taken in near-infrared and three in visible light, containing images of iris regions of different sizes under various shooting conditions. In these seven datasets, images of the left and right irises from the same subject were classified into two classes, which meant that the left and right irises of the same subject could be randomly used with equal probability. The samples of seven iris datasets are shown in Figure 7. The use of seven iris datasets is described below.

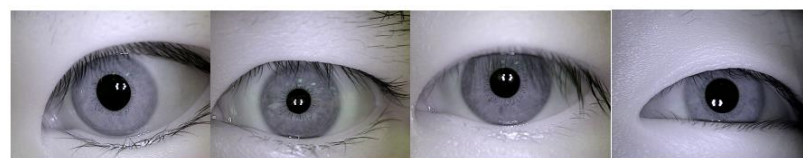
1. The JLU-6.0 iris dataset is characterized by large iris regions with various quality iris images, taken at close range and in near-infrared light. We selected 100 classes for the experiments. After data augmentation, the total number of images used was 10,006, of which the training set contained 7004 images, the validation set contained 2001 images, and the test set contained 1001 images.
2. The JLU-7.0 iris dataset has smaller iris regions compared with JLU-6.0 iris dataset and also has various quality iris images taken at close range and in near-infrared light. In all, 78 classes were selected for experiments, and after data augmentation, the total number of images was 986 for the test set entirely.
3. The iris images from the CASIA-Iris-Interval-v4 iris dataset are of high quality, with detailed features of the iris clearly visible, taken at close range and in near-infrared light. Overall, 120 classes were selected. A total number of 960 images were used after augmentation, all of which were used for testing.
4. The CASIA-Iris-Lamp-v4 iris dataset intentionally introduced ambient lighting variations, acquiring iris images with non-linear deformation. Since most of irises are from Orientals, the iris regions are heavily obscured by the upper and lower eyelids and eyelashes. Images were taken at close range and in visible light. We chose 120 classes for experiments. There were totally 9600 images after data augmentation, of which

the training set contained 6720 images, the validation set contained 1920 images, and the test set contained 960 images.

5. The UBIRIS.V2 iris dataset introduced more noises such as defocus blur, contact lens occlusion, and hair occlusion, taken at long distance and in visible light. From this dataset, 100 classes were selected. After data augmentation, 1000 images were used for the test set.
6. The main feature of the MICHE-I iris dataset is that it was acquired by mobile devices, containing more realistic noises. Most of images were obtained under unconstrained conditions, closer to realistic situations. We used 75 classes and 9360 images totally after augmentation. The training set contained 6552 images, the validation set contained 1872 images, and the test set contained 936 images.
7. The Mmu2 iris dataset has iris images from people of different races and ages in Asia, the Middle East, Africa, and Europe; with pupil spots, hair, eyelashes, eyebrows, and other noises; taken at close range and in near-infrared light. After augmentation, 974 iris images were all used for testing with 100 classes.

In order to avoid overfitting of the proposed PFSegIris and improve the generalization ability of the model, we performed data augmentation on the iris datasets used for experiments. Targeting the accurate segmentation of multi-source heterogeneous iris images and considering the actual application scenarios of iris recognition, the following data augmentation methods were designed and used.

1. Randomly flip horizontally with the probability of 0.5.
2. Randomly rotate within a certain angle with the probability of 0.8. The rotation angle we used was a maximum of 30 degrees to the left and right.
3. Randomly resize with the probability of 0.8. Four different scales of 0.5, 0.75, 1.25, and 1.5 times of the original images were used.
4. Randomly enhance the brightness and contrast with the probability of 0.5. The values that define the minimum and maximum adjustment of image brightness were 0.5 and 1.5. The values that define the minimum and maximum adjustment of image contrast were 0.5 and 1.



(a)



(b)



(c)

Figure 7. *Cont.*

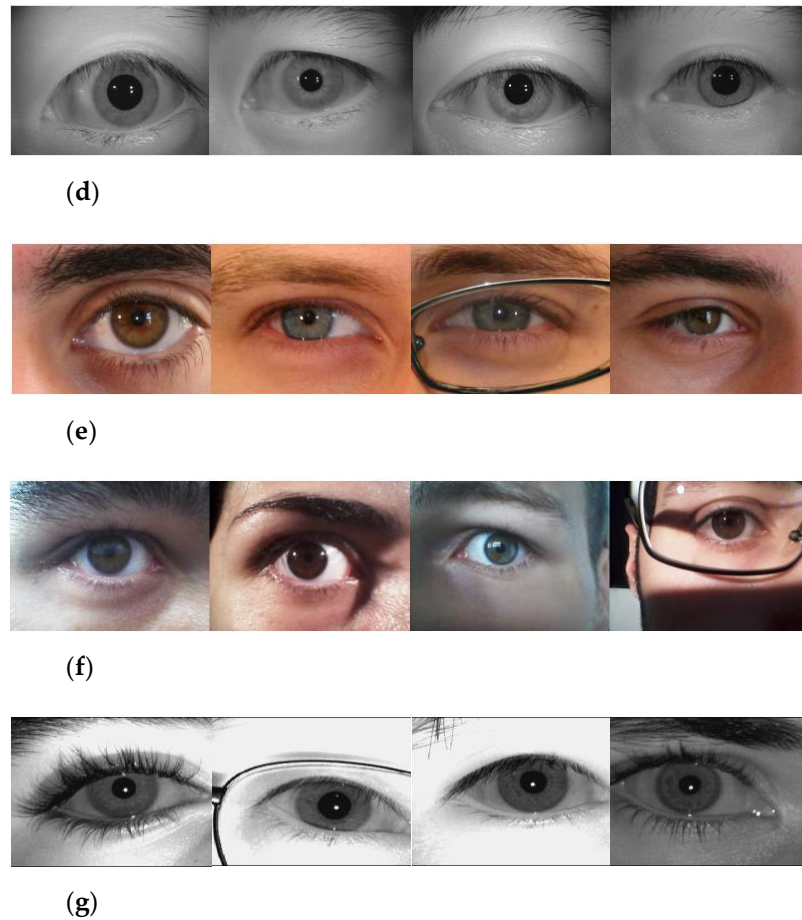


Figure 7. The samples of seven iris datasets: (a) JLU-6.0; (b) JLU-7.0; (c) CASIA-Iris-Interval-v4; (d) CASIA-Iris-Lamp-v4; (e) UBIRIS.V2; (f) MICHE-I; (g) Mmu2.

3.3. Evaluation Indexes

In order to evaluate the effectiveness of the proposed algorithm, the evaluation indexes [31] used in the experiments are described below.

1. mIoU (mean Intersection over Union) is a commonly used performance index in semantic segmentation, which is the average of the ratio of the intersection and union of the two sets of real and predicted values. The value of mIoU is in the range of [0, 1]. The closer the value is to 1, the higher the accuracy of segmentation is. mIoU was calculated using Equation (5).

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (5)$$

where p_{ii} denotes the pixel whose real pixel is i predicted to be i , and p_{ij} denotes the pixel whose real pixel is i predicted to be j . $k+1$ is the number of classes.

2. mPA (mean Pixel Accuracy) is a simple promotion of PA (Pixel Accuracy), which calculates the proportion of correctly classified pixels in each class and averages it over all classes. The value of mPA is in the range of [0, 1]. The closer the value is to 1, the higher the accuracy of segmentation is. mPA was calculated using Equation (6).

$$\text{mPA} = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (6)$$

where p_{ii} denotes the pixel whose real pixel is i predicted to be i , and p_{ij} denotes the pixel whose real pixel is i predicted to be j . $k + 1$ is the number of classes.

3. Precision and Recall are a pair of contradictory measures, and F1-score is the harmonic average of the two. The value of F1-score is in the range of $[0, 1]$. The closer the value is to 1, the higher the accuracy of segmentation is. F1-score was calculated using Equation (7).

$$F1 - score = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (7)$$

4. Time complexity:

- FLOPs is the abbreviation of floating point operations. A smaller value of FLOPs indicates that the network model is less computational intensive.
- Average Time is the average time to predict an iris mask image. The smaller the value of Average Time is, the better performance the practical application has.

5. Space complexity:

- Params is the total number of weight parameters of all parameter layers in the network model. The smaller the value of Params is, the smaller the number of parameters of the network is.
- Storage Space is the storage space of the network model. The smaller the Storage Space value is, the better the practical application of the model is.

3.4. Experimental Results and Analysis

3.4.1. Mixed Iris Dataset for Training and Testing

In this experiment, the training set and validation set iris images of the three datasets JLU-6.0, CASIA-Iris-Lamp-v4, and MICHE-I were selected and mixed to form a mixed iris training set and validation set. Then, training, validation, and testing were performed. Figure 8 shows the loss curves of the proposed PFSegIris on the mixed iris training set and validation set. Figure 9 shows the accuracy curves of PFSegIris on the mixed iris training set and validation set. Figure 10 shows the iris segmentation results of PFSegIris on the JLU-6.0, CASIA-Iris-Lamp-v4, and MICHE-I test sets, respectively. Table 3 gives the evaluation indexes of PFSegIris on the JLU-6.0, CASIA-Iris-Lamp-v4, and MICHE-I test sets, respectively.

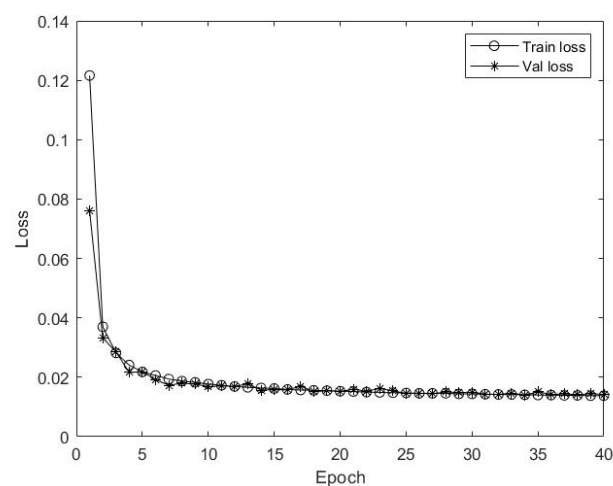


Figure 8. The loss curves of PFSegIris on the mixed iris training set and validation set.

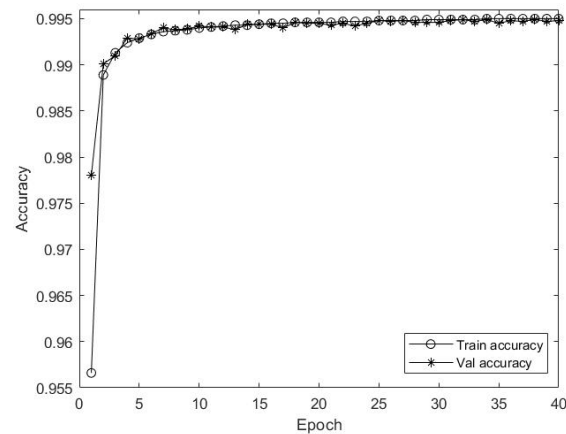


Figure 9. The accuracy curves of PFSegIris on the mixed iris training set and validation set.

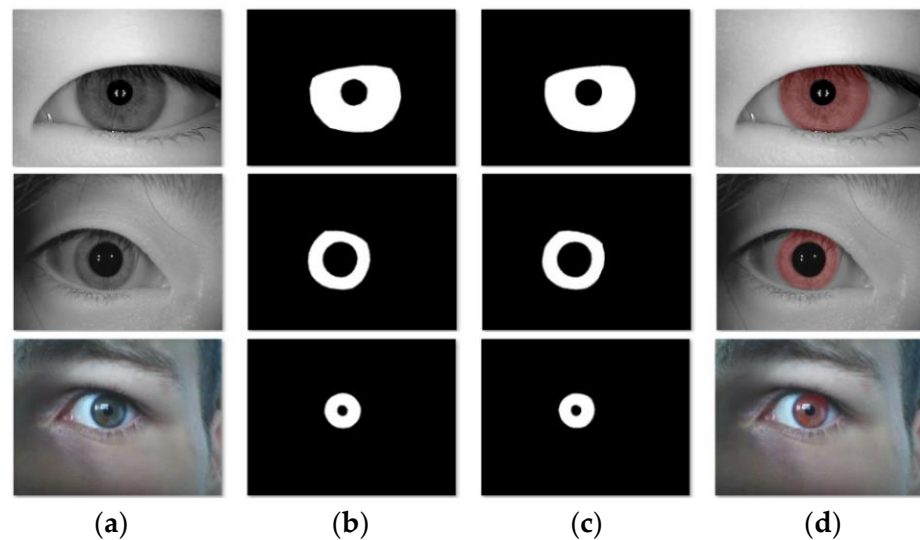


Figure 10. The segmentation results of PFSegIris on the JLU-6.0, CASIA-Iris-Lamp-v4, and MICHE-I test sets, respectively: (a) original images; (b) ground truth; (c) segmentation results; (d) merged images.

Table 3. The evaluation indexes of PFSegIris on the JLU-6.0, CASIA-Iris-Lamp-v4, and MICHE-I test sets, respectively.

Dataset	mIoU/%	mPA/%	F1-Score/%	Average Time/s
JLU-6.0	97.38	99.33	98.68	0.12
CASIA-Lamp	97.15	98.97	97.91	0.10
MICHE-I	96.24	97.11	97.13	0.06

Conclusions can be obtained as follows after the analysis of Figures 8–10 and Table 3:

1. The loss and accuracy curves converged fast with increasing epochs in the process of training and validation, and the curves were stable without oscillations, indicating that the model did not overfit, which avoids the emergence of complex models.
2. PFSegIris performed well on all three kinds of heterogeneous iris images and handled the details of small targets well. It performed best on the JLU-6.0 iris dataset. The main reason was that the JLU-6.0 dataset has high-quality near-infrared images with large iris area, while the other two are visible-light iris images with more noise.
3. The average prediction times of the three heterogeneous iris images were all short.

3.4.2. Cross-Dataset Evaluation of PFSegIris

To further verify that the iris segmentation algorithm designed in the paper has universality and fast segmentation performance for multi-source heterogeneous iris images, the model trained on a mixed iris dataset in Section 3.4.1 was tested directly on the four datasets JLU-7.0, CASIA-Iris-Interval-v4, UBIRIS.V2, and Mmu2, respectively. Figure 11 shows the cross-dataset segmentation results of PFSegIris on the four datasets JLU-7.0, CASIA-Iris-Interval-v4, UBIRIS.V2, and Mmu2, respectively. Table 4 shows the cross-dataset evaluation indexes on the four datasets.

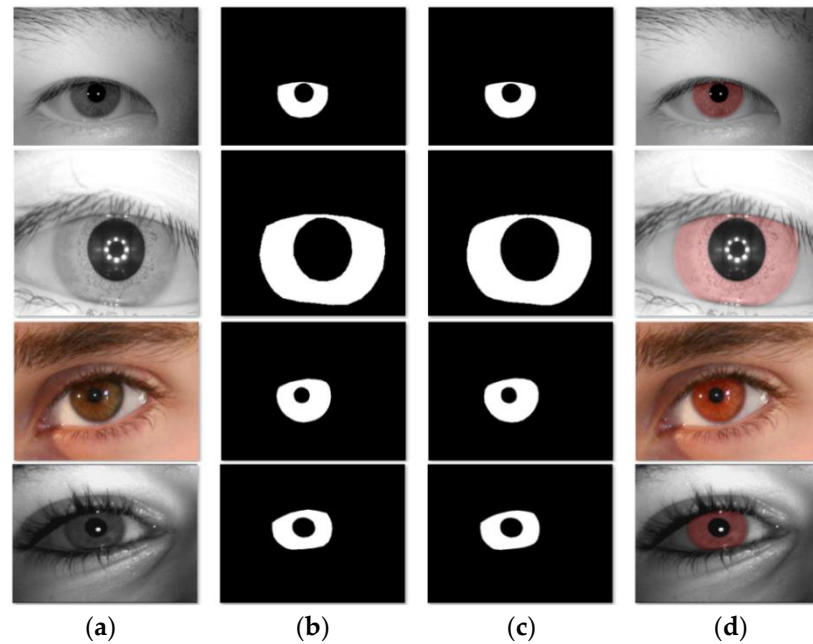


Figure 11. The cross-dataset segmentation results of PFSegIris on the JLU-7.0, CASIA-Iris-Interval-v4, UBIRIS.V2, and Mmu2 iris datasets, respectively: (a) original images; (b) ground truth; (c) segmentation results; (d) merged images.

Table 4. The cross-dataset evaluation indexes of PFSegIris on the JLU-7.0, CASIA-Iris-Interval-v4, UBIRIS.V2, and Mmu2 iris datasets, respectively.

Dataset	mIoU/%	mPA/%	F1-Score/%	Average Time/s
JLU-7.0	97.06	98.56	97.68	0.11
CASIA-Interval	98.27	99.40	99.13	0.10
UBIRIS.V2	96.69	97.30	97.29	0.09
Mmu2	96.82	97.62	97.47	0.09

Conclusions can be obtained as follows after the analysis of Figure 11 and Table 4:

1. The test results on four cross-datasets were still good, indicating that PFSegIris had learned the real iris features, had a certain migration ability and generalization ability, and had universality for multi-source heterogeneous iris image segmentation.
2. Among the four heterogeneous datasets, the effect of the CASIA-Iris-Interval-v4 was remarkably better than that of the other three datasets. The reason was that the images from CASIA-Iris-Interval-v4 were of high quality in near-infrared light, with large iris regions and almost no noise interference. Iris images from UBIRIS.V2 had more noise and therefore showed less generalization ability.
3. Fast segmentation was still realized on four cross-dataset iris images.

3.4.3. Comparison with Existing Segmentation Algorithms

In this experiment, deep-learning-based algorithms FCN-8s [11], U-Net [13], FRED-Net [14], PI-Unet [15], and MFFIris-Unet [18] were each selected to be trained and tested on multi-source heterogeneous iris datasets. The test results were compared with PFSegIris on several evaluation indexes. Four datasets were selected for the experiment, which were JLU-6.0, CASIA-Iris-Lamp-v4, UBIRIS.V2, and MICHE-I. Table 5 gives the evaluation indexes of different algorithms on four iris datasets. Table 6 gives the comparison of different algorithms in terms of Params, Storage Space, and FLOPs.

Table 5. The evaluation indexes of different algorithms on four iris datasets.

Method	Dataset	mIoU/%	F1-Score/%	Average Time/s
FCN-8s	JLU-6.0	85.35	92.16	0.29
	CASIA-Lamp	83.24	90.73	0.28
	UBIRIS.V2	77.72	86.85	0.15
	MICHE-I	77.29	87.67	0.19
U-Net	JLU-6.0	89.01	93.98	0.96
	CASIA-Lamp	87.45	92.04	0.92
	UBIRIS.V2	82.21	91.77	0.61
	MICHE-I	81.54	90.25	0.58
FRED-Net	JLU-6.0	94.08	95.78	0.34
	CASIA-Lamp	92.50	94.63	0.31
	UBIRIS.V2	91.89	93.84	0.28
	MICHE-I	91.03	93.15	0.25
PI-Unet	JLU-6.0	95.58	97.03	0.21
	CASIA-Lamp	94.35	96.68	0.18
	UBIRIS.V2	92.33	95.52	0.27
	MICHE-I	93.67	94.26	0.35
MFFIris-Unet	JLU-6.0	96.17	97.56	0.15
	CASIA-Lamp	95.76	97.28	0.12
	UBIRIS.V2	95.32	96.62	0.11
	MICHE-I	94.75	96.59	0.09
PFSegIris	JLU-6.0	97.38	98.68	0.12
	CASIA-Lamp	97.15	97.91	0.10
	UBIRIS.V2	96.69	97.29	0.09
	MICHE-I	96.24	97.13	0.06

Table 6. Comparison of different algorithms in terms of Params, Storage Space, and FLOPs.

Method	Params/M	Storage Space/GB	FLOPs/G
FCN-8s	134.27	0.500	83.42
U-Net	31.03	0.116	62.06
FRED-Net	9.7	0.036	19.5
PI-Unet	2.96	0.011	1.60
MFFIris-Unet	1.95	0.007	0.74
PFSegIris	1.86	0.007	0.65

Conclusions can be obtained as follows after the analysis of Tables 5 and 6:

1. Compared with other methods, PFSegIris achieved higher segmentation accuracy on multi-source heterogeneous iris images, with mIoU reaching 97.38%, 97.15%, 96.69%, and 96.24% and F1-score reaching 98.68%, 97.71%, 97.29%, and 97.13% on four iris datasets of JLU-6.0, CASIA-Iris-Lamp-v4, UBIRIS.V2, and MICHE-I, respectively, further verifying the universal applicability of the algorithm proposed for multi-source heterogeneous iris segmentation.

2. Compared with other network models, PFSegIris had a faster segmentation speed on multi-source heterogeneous iris images. It is an algorithm that can quickly segment multi-source heterogeneous iris.
3. Compared with the classic semantic segmentation methods and the latest lightweight iris segmentation methods, the proposed algorithm has fewer parameters, less storage space, and less computation with obvious application advantages.

3.4.4. Ablation Study

To verify the effectiveness of each component of the proposed algorithm in improving the performance of iris segmentation, four different networks were designed for the ablation study. The baseline network was the network structure without a parallel dual attention mechanism. Two different networks were designed by adding the position attention module and the channel attention module, respectively, to the baseline network. The evaluation indexes of the four networks were compared on four heterogeneous iris datasets, and the results are shown in Table 7. Figure 12 shows the test results of the baseline network and the PFSegIris on the four datasets.

Table 7. The evaluation indexes of different networks on four iris datasets.

Network	Dataset	mIoU/%	F1-Score/%	Average Time/s
Baseline	JLU-6.0	96.42	97.78	0.24
	CASIA-Lamp	96.27	96.83	0.21
	UBIRIS.V2	95.80	96.47	0.19
	MICHE-I	95.41	96.39	0.11
Baseline + PAM	JLU-6.0	96.85	98.03	0.16
	CASIA-Lamp	96.63	97.14	0.14
	UBIRIS.V2	96.29	96.90	0.12
	MICHE-I	95.85	96.72	0.08
Baseline + CAM	JLU-6.0	97.03	98.32	0.14
	CASIA-Lamp	96.85	97.55	0.13
	UBIRIS.V2	96.41	97.01	0.11
	MICHE-I	96.03	96.95	0.07
PFSegIris	JLU-6.0	97.38	98.68	0.12
	CASIA-Lamp	97.15	97.91	0.10
	UBIRIS.V2	96.69	97.29	0.09
	MICHE-I	96.24	97.13	0.06

Conclusions can be obtained as follows after the analysis of Table 7 and Figure 12:

1. After adding the attention mechanism to the baseline network, the average time to predict heterogeneous iris images was greatly reduced.
2. With the attention mechanism, the model had higher accuracy than baseline segmentation, and both mIoU and F1-score were improved. The channel attention module had a stronger boosting performance than the position attention module.
3. It can be seen from the subjective vision that the baseline did not work well for iris segmentation with more noises and small targets, while our algorithm performed well and accurately segmented small iris regions on heterogeneous iris images, avoiding the interference of noise such as lighting. Thus, this experiment illustrates the effectiveness of the parallel dual attention mechanism combination to improve the segmentation precision of multi-source heterogeneous iris images.

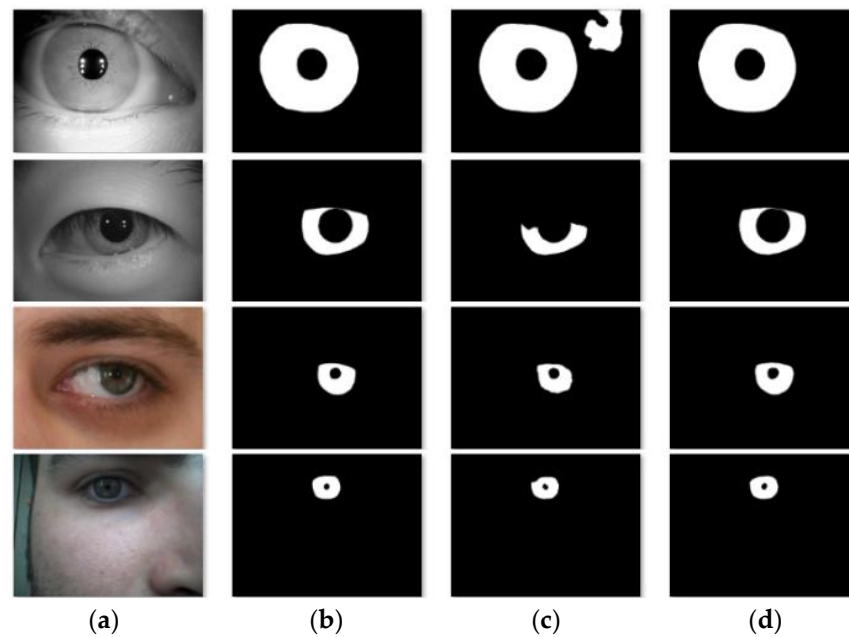


Figure 12. The segmentation results predicted by baseline and PFSegIris on four iris datasets: (a) original images; (b) ground truth; (c) segmentation results of baseline; (d) segmentation results of PFSegIris.

3.4.5. Comparison with Traditional Iris Positioning Algorithm

In order to further verify the accuracy of the algorithm proposed for multi-source heterogeneous iris segmentation, a comparative experiment was carried out with the widely used method OSIRIS V4.1 [32]. The contour detection method was performed on the mask images of the four heterogeneous iris test sets and positioning images of the iris area were obtained. The number of images with wrong position area was calculated. The original iris images of four heterogeneous iris test sets were input into OSIRIS V4.1, respectively, and the number of images with wrong position area was calculated. In this experiment, the three iris datasets, MICHE-I, UBIRIS.V2, and Mmu2, were not used because OSIRIS V4.1 is not applicable to iris images containing lots of realistic noises and taken from non-Asians. As the positioning effect is very poor and there is no comparability. The experimental results are shown in Table 8. The comparison of positioning results of PFSegIris and OSIRIS V4.1 for four heterogeneous iris test sets is shown in Figure 13.

Table 8. Comparison of experimental results.

Dataset	Method	Error Number	Error Rate
JLU-6.0	PFSegIris	23	0.023
	OSIRIS V4.1	285	0.285
JLU-7.0	PFSegIris	33	0.033
	OSIRIS V4.1	382	0.387
CASIA-Interval	PFSegIris	16	0.017
	OSIRIS V4.1	127	0.132
CASIA-Lamp	PFSegIris	29	0.030
	OSIRIS V4.1	368	0.383

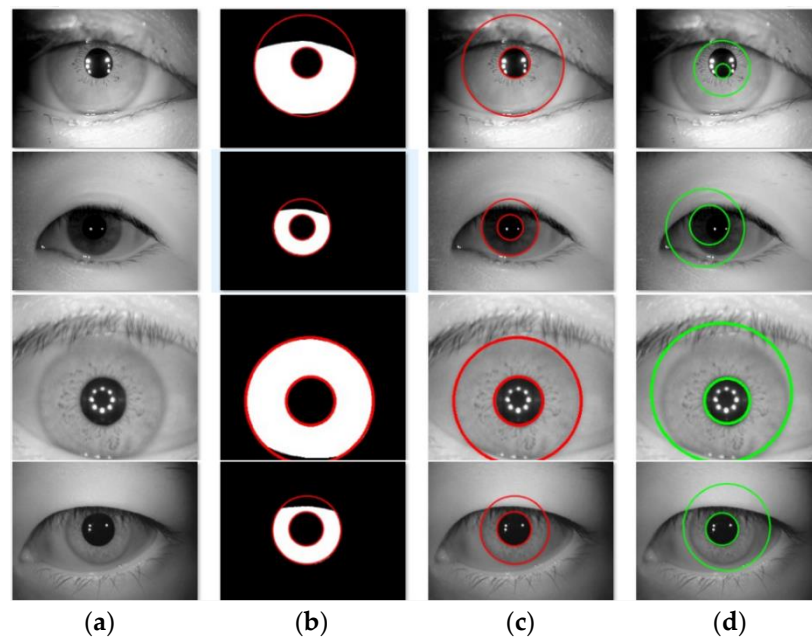


Figure 13. Positioning results comparison of PFSegIris and OSIRIS V4.1 on four iris test sets: (a) original images; (b) mask positioning images; (c) PFSegIris positioning images; (d) OSIRIS V4.1 positioning images.

We can see from Table 8 and Figure 13 that the proposed PFSegIris can divide the iris regions more accurately than the traditional positioning method. This is because the segmentation algorithm in this paper can avoid the problems of noise interference caused by eyelashes, eyelids, and lighting in the traditional method.

3.4.6. Qualitative Comparison of Extreme Images

Since most of the evaluation indexes above are of statistical significance, it is hard to evaluate the robustness of the algorithm proposed for extremely difficult cases. Therefore, we selected some representative hard samples from datasets and performed a qualitative comparison of the algorithms to obtain an objective and comprehensive evaluation.

In order to compare the performance of algorithms on hard samples, we randomly selected five hard samples with specular reflection, eyelid occlusion, eyelash occlusion, slanted eyes, and out-of-focus blur from the five different datasets JLU-7.0, CASIA-Iris-Lamp-v4, UBIRIS.V2, MICHE-I, and Mmu2. The three algorithms with the best results in Section 3.4.3 were used for comparison with the PFSegIris, which were MFFIris-Unet [18], PI-Unet [15], and FRED-Net [14]. Figure 14 shows the comparison of test results of PFSegIris and the other three algorithms on the five hard samples.

As can be seen from Figure 14, the proposed algorithm shows better segmentation results than the other three segmentation algorithms on five hard samples. Since the near-infrared images are less affected by ambient light, iris contours in these images are more obvious compared with those in the visible-light images. The segmentation performance of the proposed algorithm on the visible and near-infrared spectral iris images was more balanced than in other algorithms, and the segmentation results were closer to the ground truth with finer and smoother inner and outer iris edges. Therefore, it illustrates that PFSegIris performs more robustly in the face of challenging and hard samples and can identify the location of iris regions more accurately.

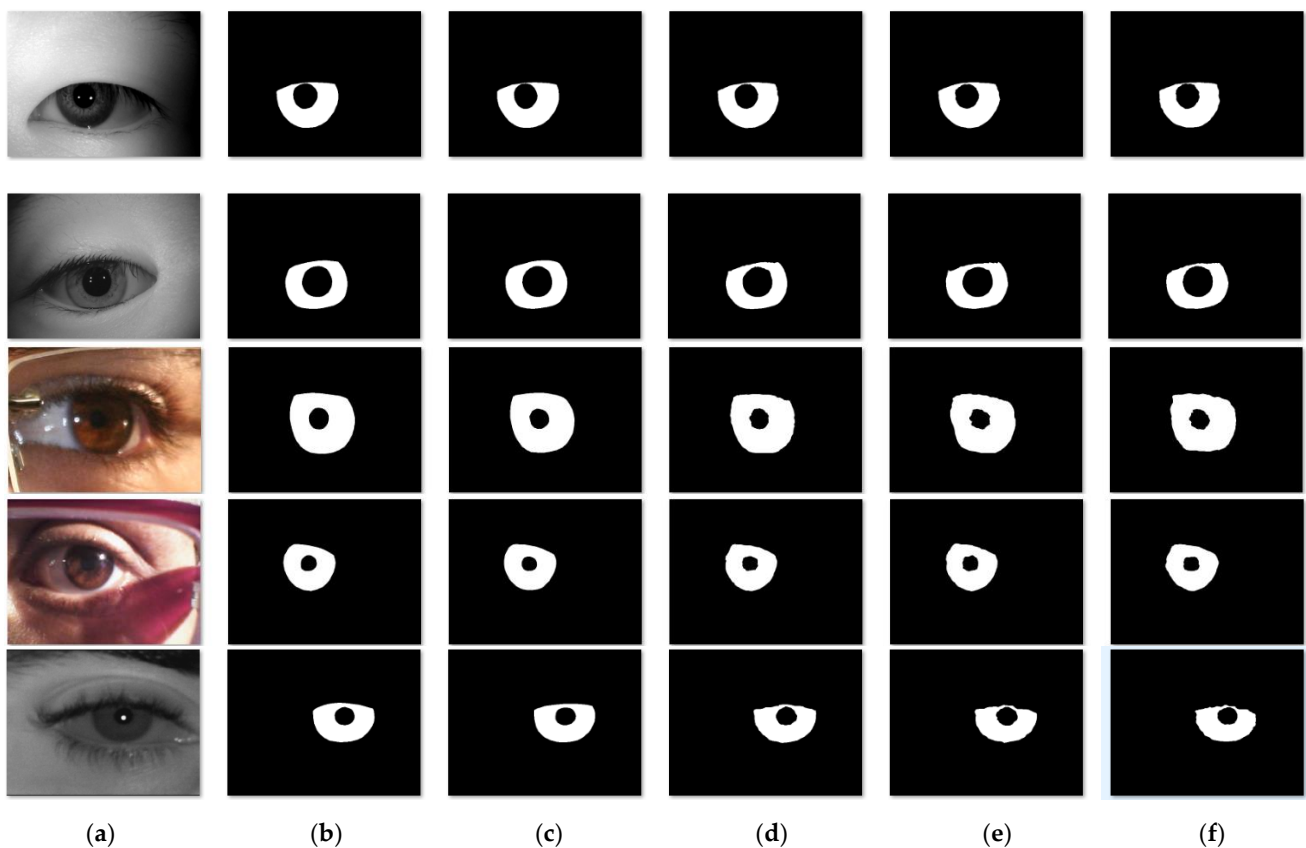


Figure 14. Test results comparison of PFSegIris and other three algorithms on five hard samples from JLU-7.0, CASIA-Iris-Lamp-v4, UBIRIS.V2, MICHE-I, and Mmu2, respectively: (a) original images; (b) ground truth; (c) segmentation results of PFSegIris; (d) segmentation results of MFFIris-Unet; (e) segmentation results of PI-Unet; (f) segmentation results of FRED-Net.

4. Conclusions

With respect to the problem that various iris segmentation methods work well for distinct iris datasets and poorly for multi-source heterogeneous iris images, a precise and fast segmentation algorithm, PFSegIris, was proposed for multi-source heterogeneous iris images. The algorithm fully extracted the information of iris images using the lightweight iris feature extraction module designed in this paper. The efficient parallel dual attention module suppressed the influence of noise and other interference factors, and further enhanced the discriminability of pixels in the iris region. A low-level skip connection supplemented the edge feature information. Experimental results on seven heterogeneous iris datasets from different sources showed that the proposed algorithm is superior in precision and speed to existing deep-learning-based iris segmentation algorithms and traditional positioning algorithms, with F1-score up to 99.13% and mIoU up to 98.27%. The shortest time was 0.06 s. The Params, Storage Space, and FLOPs of PFSegIris were only 1.86 M, 0.007 GB, and 0.65 G.

The algorithm proposed in the paper is able to segment multi-source heterogeneous iris quickly and precisely, providing sufficient and reliable feature information for the subsequent recognition process, which is of good practical application value on low-performance computing devices such as mobile terminals.

Author Contributions: Conceptualization, L.D. and X.Z.; methodology, L.D.; software, L.D.; validation, L.D.; formal analysis, L.D., Y.L. and X.Z.; investigation, L.D.; resources, L.D., Y.L. and X.Z.; data curation, L.D.; writing—original draft preparation, L.D.; writing—review and editing, L.D.; visualization, L.D.; supervision, L.D.; project administration, L.D.; funding acquisition, Y.L. and X.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Science Foundation of Jilin Province, grant number YDZJ202101ZYTS144; and the National Natural Science Foundation of China (NSFC), grant number 61471181.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: Thanks to the Jilin Provincial Key Laboratory of Biometrics New Technology for supporting this project.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wang, W.H.; Zhu, Y.; Tan, T.N. Identification based on iris recognition. *J. Autom.* **2002**, *28*, 1–10.
2. Jain, A.K. Technology: Biometric recognition. *Nature* **2007**, *449*, 38–40. [[CrossRef](#)] [[PubMed](#)]
3. Hollingsworth, K.P.; Bowyer, K.W.; Flynn, P.J. Improved Iris Recognition through Fusion of Hamming Distance and Fragile Bit Distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2465–2476. [[CrossRef](#)]
4. Kang, J.S. Mobile iris recognition systems: An emerging biometric technology. *Procedia Comput. Sci.* **2010**, *1*, 475–484. [[CrossRef](#)]
5. Jillela, R.; Ross, A.A. Methods for Iris Segmentation. In *Handbook of Iris Recognition*, 2nd ed.; Bowyer, K.W., Burge, M.J., Eds.; Springer: Cham, Switzerland, 2016; Volume 4, pp. 137–184.
6. Daugman, J.G. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 1148–1161. [[CrossRef](#)]
7. Wildes, R.P. Iris recognition: An emerging biometric technology. *Proc. IEEE* **1997**, *85*, 1348–1363. [[CrossRef](#)]
8. He, K.; Gkioxari, G.; Dollar, P.; Girshick, R. Mask R-CNN. In Proceedings of the IEEE ICCV, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
9. Girshick, R. Fast R-CNN. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 1440–1448.
10. Ren, S.; He, K.; Girshick, R.; Sun, J. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 1137–1149. [[CrossRef](#)]
11. Long, J.; Shelhamer, E.; Darrell, T. Fully Convolutional Networks for Semantic Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *39*, 640–651.
12. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [[CrossRef](#)] [[PubMed](#)]
13. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Cham, Switzerland, 2015; pp. 234–241.
14. Arsalan, M.; Kim, D.S.; Lee, M.B.; Owais, M.; Park, K.R. FRED-Net: Fully Residual Encoder-Decoder Network for Accurate Iris Segmentation. *Expert Syst. Appl.* **2019**, *122*, 217–241. [[CrossRef](#)]
15. Zhou, R.Y.; Shen, W.Z. PI-Unet: Research on precise iris segmentation neural network model for heterogeneous iris. *Comput. Eng. Appl.* **2021**, *57*, 223–229.
16. Wang, C.; Jawad, M.; Wang, Y.; He, Z.; Sun, Z. Towards Complete and Accurate Iris Segmentation Using Deep Multi-Task Attention Network for Non-Cooperative Iris Recognition. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 2944–2959. [[CrossRef](#)]
17. Li, Y.H.; Putri, W.R.; Aslam, M.S.; Chang, C.C. Robust Iris Segmentation Algorithm in Non-Cooperative Environments Using Interleaved Residual U-Net. *Sensors* **2021**, *21*, 1–21.
18. You, X.; Zhao, P.; Mu, X.; Bai, K.; Lian, S. Heterogeneous Noise Iris Segmentation Based on Attention Mechanism and Dense Multi-scale Features. *Laser Optoelectron. Prog.* **2021**, in press.
19. Fu, J.; Liu, J.; Tian, H.; Li, Y.; Bao, Y.; Fang, Z.; Lu, H. Dual Attention Network for Scene Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 3146–3154.
20. JLU Iris Image Database. Available online: <http://www.jlucomputer.com/index/irislibrary/irislibrary.html> (accessed on 7 September 2020).
21. CASIA Iris Image Database. Available online: <http://www.cbsr.ia.ac.cn/china/Iris%20Databases%20CH.asp> (accessed on 7 September 2020).
22. MMU2 Iris Image Database. Available online: <http://pesona.mmu.edu.my/~ccteo> (accessed on 7 September 2020).
23. Marsico, M.D.; Nappi, M.; Riccio, D.; Wechsler, H. Mobile Iris Challenge Evaluation (MICHE)-I, biometric iris dataset and protocols. *Pattern Recognit. Lett.* **2015**, *57*, 17–23. [[CrossRef](#)]
24. MICHE-I Iris Image Database. Available online: http://biplab.unisa.it/MICHE/index_miche.htm (accessed on 7 September 2020).
25. Proena, H.; Filipe, S.; Santos, R.; Oliveira, J.; Alexandre, L.A. The UBIRIS.v2: A Database of Visible Wavelength Iris Images Captured On-the-Move and At-a-Distance. *IEEE Trans. Pattern Anal. Mach. Intell.* **2010**, *32*, 1529–1535. [[CrossRef](#)]

26. UBIRIS.v2 Iris Image Database. Available online: <http://iris.di.ubi.pt/ubiris2.html> (accessed on 7 September 2020).
27. NICE.I Iris Image Database. Available online: <http://nice1.di.ubi.pt> (accessed on 7 September 2020).
28. Tan, M.X.; Le, Q.V. MixConv: Mixed depthwise convolutional kernels. *arXiv* **2019**, arXiv:1907.09595.
29. Zhang, H.; Goodfellow, I.; Metaxas, D.; Odena, A. Self-Attention Generative Adversarial Networks. *arXiv* **2018**, arXiv:1805.08318.
30. Suzuki, S.; Abe, K. Topological Structural Analysis of Digitized Binary Images by Border Following. *Comput. Vis. Graph. Image Process.* **1985**, *30*, 32–46. [[CrossRef](#)]
31. Wang, C.; Sun, Z. A Benchmark for Iris Segmentation. *J. Comput. Res. Dev.* **2020**, *57*, 395–412.
32. Othman, N.; Dorizzi, B.; Garcia-Salicetti, S. OSIRIS: An open source iris recognition software. *Pattern Recognit. Lett.* **2016**, *82*, 124–131. [[CrossRef](#)]