



Article Double-Center-Based Iris Localization and Segmentation in Cooperative Environment with Visible Illumination

Jiangang Li and Xin Feng *

College of Computer Science and Engineering, Chongqing University of Technology, Chongqing 400054, China * Correspondence: xfeng@cqut.edu.cn

Abstract: Iris recognition has been considered as one of the most accurate and reliable biometric technologies, and it is widely used in security applications. Iris segmentation and iris localization, as important preprocessing tasks for iris biometrics, jointly determine the valid iris part of the input eye image; however, iris images that have been captured in user non-cooperative and visible illumination environments often suffer from adverse noise (e.g., light reflection, blurring, and glasses occlusion), which challenges many existing segmentation-based parameter-fitting localization methods. To address this problem, we propose a novel double-center-based end-to-end iris localization and segmentation network. Different from many previous iris localization methods, which use massive post-process methods (e.g., integro-differential operator-based or circular Hough transforms-based) on iris or contour mask to fit the inner and outer circles, our method directly predicts the inner and outer circles of the iris on the feature map. In our method, an anchor-free center-based double-circle iris-localization network and an iris mask segmentation module are designed to directly detect the circle boundary of the pupil and iris, and segment the iris region in an end-to-end framework. To facilitate efficient training, we propose a concentric sampling strategy according to the center distribution of the inner and outer iris circles. Extensive experiments on the four challenging iris data sets show that our method achieves excellent iris-localization performance; in particular, it achieves 84.02% box IoU and 89.15% mask IoU on NICE-II. On the three sub-datasets of MICHE, our method achieves 74.06% average box IoU, surpassing the existing methods by 4.64%.

Keywords: iris localization; iris segmentation; center-based object detection

1. Introduction

Iris recognition is one of the most reliable biometric technologies, and is widely applied in intelligence unlocking, border control, and forensics, among others [1–4]. A complete irisrecognition system usually consists of image acquisition, iris segmentation and localization, normalization, feature extraction, and matching. Figure 1 illustrates the key steps for iris-image preprocessing, which contains location, segmentation, and normalization. Iris localization aims to detect the inner and outer boundaries of an iris region; iris segmentation generates an iris mask to distinguish iris and non-iris pixels. The normalization process allows the alignment of any two iris images to be compared. After image preprocessing, we can achieve the normalized iris texture image and iris/noise mask image; therefore, an important part of the iris recognition pre-processing, iris segmentation and localization jointly define the region used for feature extraction and matching, directly affecting the overall iris-recognition performance.

Most previous studies have focused on iris images in cooperative environments (e.g., near-infrared illumination, high user collaboration, close acquisition distance, and stop-gaze verification), and two widely used localization methods are Daugman's integro-differential operator [5] and Wilde's circular Hough transforms [6]. Recently, it has been extended to non-cooperative environments (e.g., long distances, limited user cooperation, visible lighting, and mobile devices) because it requires minimally restrictive user cooperation



Citation: Li, J.; Feng, X. Double-Center-Based Iris Localization and Segmentation in Cooperative Environment with Visible Illumination. *Sensors* **2023**, *23*, 2238. https://doi.org/10.3390/ s23042238

Academic Editors: Youssef Chahir and Hassen Drira

Received: 13 December 2022 Revised: 7 February 2023 Accepted: 13 February 2023 Published: 16 February 2023



Copyright: © 2023 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/). and imaging conditions; however, iris images captured in non-cooperative environments often have various noise, such as gaze bias, iris rotation, aniridia reflection, specular reflection, motion/defocus blur, and eyelid/lash/hair/glasses occlusion (see Figure 2), thereby making iris localization challenging.





Figure 1. Iris-image preprocessing for the iris-recognition system, \otimes presents the pixel-wise multiplication.

Figure 2. Example of degraded iris images: (a) glasses occlusion, (b) blur, (c) absence, (d) gaze deviation, (e) eyelid occlusion, and (f) specular reflection.

Recent deep learning-based methods use convolution neural networks (CNNs) to separate the iris in the image and use circular Hough transform [6] thereafter to fit the innerand outer-circle parameters on the iris mask. These methods have achieved good results in iris mask segmentation, but localization is extremely rough and accuracy remains below the standard required for iris recognition. Other CNN-based methods, such as IrisParseNet [7] and NIR-Zhang [8], use the semantic segmentation method to predict inner-/outer-circle boundary masks, and utilize some post-processing thereafter to obtain parameterized inner and outer circles. Although localization is improved, it requires massive refined postprocessing, and the accuracy of the inner- and outer-circle localizations is easily influenced by a fuzzy contour boundary.

We use the preceding issues as bases in considering the following three aspects: (1) how toachieve a more accurate and robust localization method on iris images captured in a non-cooperative environment with visible illumination? (2) performing post-processing (e.g., integro-differential operator-based or circular Hough transforms-based method) on pupil or iris mask to fit the iris' inner and outer circle parameters (x, y, radius) is time-consuming and not robust, can the model directly output the localization results of the inner and outer circles without any post-processing? (3) On the basis of (2), can we obtain a simple but effective end-to-end model that integrates localization and segmentation in one network? For the above three aspects, and considering the natural biological characteristics of the iris region, we propose a double-center-based method to localize and segment the iris region. Our main contributions are summarized as follows:

(1) We propose to locate the iris's inner and outer circles as center points and regression radius on feature maps directly, thereby solving the problems of prediction inaccuracies and lacking robustness of iris localization on the visible iris image. Compared with existing methods that use massive post-processing on the predicted mask to obtain the circle parameters, our approach is post-processing-free and the localization module output is the final inner-/outer-circle localization result;

- (2) By explicitly analyzing the distribution of the center points of the pupil and iris, we propose a novel auxiliary sample strategy to accelerate model training on nonstandard iris images. These images have irrelevant face regions, such as chin, nose, and environments;
- (3) We design an end-to-end dedicated iris localization and segmentation framework, which is simple, effective, and achieves excellent performance in multiple benchmarks. The proposed method provides a good foundation for iris localization and segmentation in a non-cooperative environment.

The remainder of this paper is organized as follows: Section 2 introduces the related work on iris segmentation and localization. Section 3 presents details pertaining to the proposed iris localization and segmentation methods. Section 4 presents ablation, and compares and analyzes the experimental results. Section 5 summarizes the research and concludes the study.

2. Related Research

Most traditional iris-segmentation methods mainly use a circle or ellipse to locate the inner and outer boundaries of the iris, and then use the difference of the gray histogram to exclude any superimposed occlusions of eyelashes, shadows, glasses, or reflections, and infer the iris region [3].

For iris localization, two widely used baseline methods are Daugman's integrodifferential operator [5] and Wildes's Hough transforms [6]. The integro-differential operator searched for the largest difference of intensity over the circle parameter space, and it has achieved great precision; however, it may take a relatively longer time [9]. The circular Hough transforms found optimal circle parameters by a voting procedure in a binary edge image, and is often applied to detect boundaries of the circle or ellipse, but this transform is relatively non-sensitive to the broken contours of 2D objects in the binary image [10].

Based on the above two basic methods, many later proposed approaches made further improvements in accuracy and efficiency. For example, ref. [11] applied L_1 norm to suppress noisy texture before performing iris localization, ref. [12] applied region clustering before localization for narrowing the parameter search range, ref. [13] proposed integro-differential constellation to reduce the computation time, and ref. [14], applied the Viterbi algorithm on gradient maps of iris images to find coarse low-resolution contours. Those traditional methods rely considerably on prior assumptions and image low-level features, so they have low accuracy and poor robustness in dealing with noise.

Recently, deep learning-based the semantic segmentation method has achieved higher accuracy in iris segmentation. Different from the traditional pixel-based iris-segmentation method, the deep learning-based iris-segmentation methods utilize high-semantic features and estimate iris masks end to end. The first CNNs-based iris-segmentation method is HCNNs [15]. Other iris-segmentation methods based on CNNs include those in [16–21]. These methods use convolution to extract semantic features and segment the foreground of the iris, thereby achieving good iris mask segmentation.

Although CNNs-based iris-segmentation method has achieved good results, most of those methods only predict iris masks, and more important iris localization is not achieved. Ref. [15] first applied the circular Hough transforms method from original iris image to iris masks to generate candidate circular iris boundaries, and then use two quality measures to select the best inner and outer iris boundaries.

Current CNN-based methods first segment the image to achieve the iris mask and utilize a massive post-process, such as Daugman's integro-differential operator and Hough transform, to fit the inner and outer circles on the iris mask. As shown in Figure 3a, similar methods are also presented in [22,23]. These methods work well on iris images taken in a cooperative environment (e.g., near-infrared illumination, high user collaboration, close

acquisition distance, and stop-gaze verification), but in iris images taken in non-cooperative and visible light environments, the performance was mediocre. IrisParsetNet [7] first uses the segmentation method to predict the inner and outer boundaries of the iris. While the model predicts the iris mask, it also simultaneously predicts the pupil and iris outer boundary masks, as shown in Figure 3b. NIR-Zhang [8] uses two completely independent models responsible for iris segmentation and iris localization, as shown in Figure 3c. The localization model predicts pupil and iris masks and uses post-processing thereafter to fit the contour and circle parameters.; however, the massive post-process to fit a circle on the boundary mask is time-consuming and not robust on the iris images captured in non-cooperative environments with visible illumination. The reason is that the localization results are easily affected by irregular contour edges.



Figure 3. Illustrations of different methods for iris localization.

To obtain an efficient post-processing-free iris-localization solution and achieve a more accurate and robust iris localization on iris images captured in a non-cooperative environment, we propose a center-based iris localization and segmentation method and make localization rely on a feature map other than the iris mask (see Figure 3d). Inspired by the use of heat-map for face key point detection [24], we regard the inner/outer circles as two objects to locate. We first locate the double center and regress the radius thereafter based on the center. A segmentation branch is embedded after the location module, which is responsible for segmenting the iris and background pixels to obtain the iris mask. Our method is fast, accurate, and does not need any post-processing. The model's output is inner/outer circle (x_k, y_k, r_k), $k \in [inner, outer]$ and iris mask directly.

3. Methods

The architecture of the proposed ICSNet, as shown in Figure 4, mainly includes the localization and segmentation modules. The localization module consists of two subnets, namely, classification and radius regression heads, which are responsible for predicting the inner- and outer-circular bounding boxes. The segmentation module consists of RoI crop and the mask head, which are accountable for segmenting the iris mask over the detected region. As an auxiliary structure, the proposed concentric sample strategy only exists during the training stage and is cost-free in inference time.

First, this section introduces the localization module, which is our model's essential iris-localization component. Second, we present the mask-segmentation branch, which follows the localization module. Lastly, we introduce the sample strategy.



Figure 4. Architecture of ICSNet. Subnets in the three blue boxes are the center localization head, radius regression head, and mask head. Subnets in the orange box are the concentric sample operations, red arrow is the data flow, and our concentric sampling strategy is completely cost-free in inference time because auxiliary structures only exist during training. The localization module includes the classification and regression heads, which predict the circle bounding box. The mask module includes crop operation and mask head, which are responsible for cropping the iris region from the feature map and producing the segment iris mask.

3.1. Localization Module

3.1.1. Double-Center Localization

Current iris localization uses iris or contour mask to fit the inner/outer circles, and these fitting methods can be divided into integro-differential operator-based and circular Hough transforms-based methods; however, owing to the noise in a non-cooperative environment with visible light, these methods are susceptible to irregular mask or contour edges, resulting in poor robustness and inaccuracy. To address these issues, we propose a double-center-based iris-localization method, and the model relies on a feature map to directly predict the iris's inner and outer circles.

Given an input image $I \in R^{W*H*3}$, we adopt the Gaussian kernel as in CornerNet [25] to produce a ground truth heat map $Y_{x,y,c} \in [0,1]^{W*H*C}$, where (x, y) is the location and C is the number of center point categories. In our experiment, we set C = 2 (inner or outer circle). The heat map and input have the same size because our backbone is an encoder–decoder fully convolution network. A prediction $\hat{Y}_{x,y,c} = 1$ corresponds to a detected center point, while $\hat{Y}_{x,y,c} = 0$ is background.

For each ground truth center point (p_x, p_y) , we compute an equivalent $\overline{p} = \lfloor p \rfloor$; thereafter, we splat center points $(\overline{p}_x, \overline{p}_y)$ onto a heat map using a Gaussian kernel, which is the same as [26]. The peak of the Gaussian distribution is treated as a positive sample, while another pixel is treated as a negative sample. We adopt modified focal loss [27] to train.

The localization module follows behind the backbone. As shown in Figure 4, the module consists of three convolution layers and sigmoid function. The localization module predicts two heat maps, with all value range in [0,1], and the location of the maximum of one value in each heat map is the predicted center point $(x_k, y_k), k \in [inner, outer]$.

3.1.2. Radius Regression

Owing to the unique physiological characteristics of the iris region, the centers of the inner and outer circles appear in a pair; therefore, we set up two branches to be responsible for the inner- and outer-circle radius regressions.

Let $p_k = (x_k, y_k)$ be the circle box center points, where *k* represents the center point category (inner or outer circle). We utilize heat map \hat{Y} to predict center point p_k , and regress circle radius r_k thereafter for each center point p_k . We adopt an L_1 loss to train, which is defined as follows:

$$L_{size} = \frac{1}{2} \sum_{k=1}^{2} |\hat{r}_{p_k} - r_k|.$$
(1)

Thereafter, we directly use raw pixel coordinates for regression. To make the regression more accurate, we regard the samples in the Gaussian distribution region as positive samples, which are all responsible for predicting the object radius. Samples outside the Gaussian region are considered negative samples, and these samples do not need to predict the radius.

Additionally, we predict an offset for the center point to recover the discretization error caused by the integer operation and output stride. The offset is training with an L_1 loss. We take the pixel in the center of the object to predict the offset value directly, which is defined as follows:

$$L_{reg} = \lambda_{size} L_{size} + \lambda_{off} L_{off}.$$
 (2)

In our experiments, we set $\lambda_{size} = 1$ and $\lambda_{off} = 0.1$, following [28].

3.2. Segmentation Module

After the localization module process, we obtain the inner and outer center (\hat{x}_c, \hat{y}_c) and its radius \hat{r}_c . According to the detected center point of the outer circle, we crop the $\hat{r} \times \hat{r}$ RoI region on the feature map as the mask head input. We select only one iris RoI region according to the center point of the outer circle with confidence, thereby avoiding time-consuming post-processing.

The mask head comprises three convolution layers and sigmoid function. Each convolution layer consists of one 3×3 convolution layer, in which the stride is 2 and padding is 1, ReLU activation, and batch normalization (see Figure 4). After the mask branch process, we obtain the mask of the iris in the RoI region. The mask branch adopts the cross-entropy loss function to train.

3.3. Total Loss

Total loss *L* is composed of localization loss L_{cls} , regression loss L_{reg} , and mask loss L_{mask} , and weighted by three scalars, defined as follows:

$$L = w_{loc}L_{cls} + w_{reg}L_{reg} + w_{mask}L_{mask},$$
(3)

where L_{cls} is based on the modified focal loss [28], L_{reg} is calculated according to Equation (2), and L_{mask} adopts the cross-entropy loss. In our experiment, we set w_{loc} , w_{reg} , and w_{mask} as 1, 1, and 1, respectively.

3.4. Backbone

The backbone network of the model is an encoder–decoder full convolution network, which can accept any input size; moreover, the backbone is based on deep layer aggregation (DLA) [29]. In the encoder, we use a basic structure block in DLA. In the decoder, we use the up-sampling of bilinear interpolation and a single convolution layer, consisting of one 3×3 convolution layer with stride 2, ReLU activation, and batch normalization. The depth of the decoder and encoder is 5. We use a shortcut to connect the encoder block and decoder layer. The output feature map has 32 channels.

3.5. Concentric Sampling Strategy

Given the natural biological characteristics of the iris region, the centers of the inner and outer circles would overlap. Hindered by the camera device and non-cooperative environment, two centers in the captured image occasionally have a slight shift, as shown in Figure 5a,b, and annotated circles; moreover, the Gaussian heat map has overlapping regionsas shown in Figure 5c; therefore, we define a concentric Gaussian region (CGR) as the inner and outer circles sharing the same Gaussian value, which is represented as follows:

$$Y_{i,j} = \begin{cases} 1, & if \ Y_{i,j,inner} = 1 \ and \ Y_{i,j,outer} = 1 \\ 0, & if \ Y_{i,j,inner} = 0 \ or \ Y_{i,j,outer} = 0 \\ & min\left(0,5,\frac{1}{2}\sum\alpha_{c}Y_{i,j,c}\right), \ else \end{cases}$$
(4)

where $Y_{i,j,c}$ presents the original Gaussian distribution, in which the calculation method is the same as [30], *c* presents the category (inner or outer), and α is the hyper-parameter, presenting the weight of the inner and outer Gaussian distribution in the overlapping region; thereafter, the value is clipped to the range [0,0.5]. Locations of the inner/outer center points and their four directions are high activation points; hence, we set these values in the heat map as 1.0 and 0.8, respectively.



Figure 5. (a) Example of an iris, (b) inner-/outer-circle center, (c) and its Gaussian distribution. The centers of the inner and outer circles are highlighted in red and green, respectively. The Gaussian distributions of the inner and outer circles are shown in (c); the darker the color, the greater the value.

For the Gaussian distribution, we set the pixel at the circular box center region as the positive sample, while another pixel is the negative sample. We limit the inner Gaussian distribution and outer Gaussian distribution to the inner-circle region, avoiding the imbalance of loss contribution caused by the Gaussian distribution area and making the model focus immediately on important regions. The concentric sampling strategy is completely cost-free in inference time, as the auxiliary structures only exist during the training stage.

3.6. Double Center to Circular Boxes

During the inference, the model outputs 2 heat maps $\hat{Y}^{H \times W \times 1}$, where the values indicate the confidence scores of the inner and outer centers. The confidence score is sorted, and the coordinate position corresponding to the maximum value is the predicted target center point $(x_k, y_k), k \in [inner, outer]$. On the basis of the center point, the corresponding radius and offset can be directly derived.

For the center point position (x, y) and predicted radius \hat{r} and offset (δ_x, δ_y) , the corresponding circular bounding box coordinates are calculated as follows:

$$\hat{x} = x * s + \delta_x
\hat{y} = y * s + \delta_y'$$
(5)

where *s* is the scale ratio of the feature map and input size. Owing to the particularity of the backbone network, in the experiment, we set *s* = 1; therefore, the predicted radius is the size of the original pixel of the image, and $(\hat{x}, \hat{y}, \hat{r})$ is the final predicted result.

4. Experiments

4.1. Experimental Environment

We use Adam optimizer, and the model weight is initialized randomly. The initial learning rate is 0.001, weight decay is set as 0.0004, and mini-batch is 6. During training, the learning rate is automatically adjusted, and the total epoch is 120.

In the inference phase, we resize the input image to a specific resolution and forward it thereafter to directly obtain the predicted circle bounding boxes (inner and outer) and the iris mask.

4.2. Data Set and Evaluation Protocols

4.2.1. Data Set

NICE-II [31] comprises two non-overlapping subsets: (i) a training set with 1000 images from 171 subjects and (ii) a testing set with 1000 images from 150 subjects. We use all training sets to train and all testing sets to test.

MICHE [32] is composed of three sub-databases: GS4, IP5, and GT2. Images are captured in an uncontrolled environment using three mobile devices. MICHE-GS4 has 1297 images (663 and 634 indoor and outdoor images, respectively), IP5 has 1262 images (631 and 631 indoor and outdoor images, respectively), and MICHE-GT2 has 632 images (316 each for indoor and outdoor images). In the experiment, we use all indoor images for training and all outdoor images for testing for each sub-data set.

For the above four datasets, NICEII, MICHE-GT2, MICHE-GS4, and MICHE-IP5, the picture resolution (width × height) is 300×400 , 400×300 , 270×444 , and 270×444 , respectively. Since our network is a full convolutional encoder-decoder network, the image resolution that can be entered needs to be a multiple of 32; therefore, the input size of each dataset is 320×448 , 448×320 , 320×448 , and 320×448 , respectively. In addition to expanding the width and height by central cropping, there is no other data processing.

The images of these databases are captured from user non-cooperative and visible light environments, and taken by mobile devices rather than imaging sensors (such as infrared iris cameras); therefore, these images usually contain light reflection, gaze deviation, defocusing, mirror reflection, and other noises. Some examples of eye images and ground truths are shown in Figure 6. We aim to achieve more accurate localization of iris inner and outer boundaries, as well as eliminate these noises to obtain valid iris pixels.



Figure 6. Example images and corresponding ground truths (including iris inner boundary (green), iris outer boundary (red), and iris mask (white)) of four iris databases.

4.2.2. Evaluation Protocols

We adopt several evaluation protocols for the inner-/outer-circle localization and iris segmentation to evaluate the proposed method.

- (1). Localization. We compute the inner-/outer-circle box IoU mIoU^{box}, which ranges from [0,1]. The closer the value to 1, the better the localization. We also compute the Hausdorff distance, similar to [7]. We add points coordinate normalization, and the range of normalization Hausdorff distances is between [0,1]. The smaller the value, the higher the shape similarity;
- (2). Segmentation. We use E1_{mask} and E1_{norm}, which are the mask errors of the iris and normalized iris masks, respectively, which are the same as [7]. The value range is [0,1]. The smaller the value, the better the result. In addition, we use mIOU_{mask} to evaluate the segmentation performance, and the value range is [0,1]. The larger the value, the better the segmentation result.

4.3. Ablation Study

We conduct an ablation study to demonstrate the influence of the concentric sample. Results on the NICE-II [31] and MICHE [32] data sets are shown in Table 1. In the MICHE database, the images have irrelevant face regions, such as chin, nose, and environment, as shown in the middle column of Figure 6. Note that using this auxiliary training strategy brings a relatively significant performance gain of nearly 4.79%; moreover, the results show that making inner- and outer-circle Gaussian distributions share the same Gaussian value on the overlapping region can cause the model to focus on the critical region. Sampling over the central region is suitable for iris localization on non-standard iris images. In the NICE-II database, the images are in the eye region, as shown in the first column of Figure 6. Note that using CGR degrades the model performance. The low similarity in distribution between the NICE-II and MICHE data sets is the main reason. Using CGR on the NICE-II data set introduces many positive samples, which distract the model attention. After canceling CGR, the model only needs to focus on two positive samples (center point of the inner and outer circles), making the model more accurate.

Table 1. Ablation study of concentric sampling. "↑" denotes the improvement of localization performance after using concentric sampling strategy to auxiliary training.

Data Sets	None	DC-CGR	mIoU ^{box}	mIoU ^{box}	↑mIoU ^{box}
NICEII	\checkmark	\checkmark	73.99% 69.40%	94.06% 90.43%	-4.11%
MICHE-GS	\checkmark	\checkmark	59.01% 64.06%	80.66% 86.20%	5.29%
MICHE-IP	\checkmark	\checkmark	58.01% 63.07%	82.65% 87.34%	4.87%
MICHE-GT2		\checkmark	54.30% 57.47%	79.65% 86.06%	4.79%

4.4. Compared with Other Methods and Discussion

To verify the effectiveness of our model, we compare proposed method ICSNet with other state-of-the-art methods, FCEDN [18], IrisDenseNet [19], FRED-Net [20], IrisParseNet [7], and NIR-Zhang [8]. Tables 2 and 3 provide summaries of the performance comparison of the proposed approach with baseline method on iris localization and iris segmentation using the proposed evaluation. We also report the parameter amount and FLOPs in order to further evaluate the proposed approach.

For localization, FCEDN [18], IrisDenseNet [19], and FRED-Net [20] use an integrodifferential operator or circular Hough transform to fit the inner and outer circles on the predicted iris mask. IrisParseNet [7] predicts the iris mask and two iris edge masks, and then adopts the circle fitting method on the edge mask to obtain the circular inner and outer iris boundaries. NIR-Zhang [8] use an independent model to predict the inner contour and the outer contour, and then use a circular Hough transform to fit the inner circle and the

Table 2. Accuracy of iris segmentation and localization on NICE-II.								
Methods	mIoU ^{box}	mIoU ^{box}	mIoU ^{box}	mHdist ⁱⁿ	mHdist ^{out}	E1 ^{mask}	mIOU ^{mask}	E1 ^{norm}
FCEDN	0.5564	0.9083	0.7323	0.0298	0.0175	0.0108	0.7246	0.2323
IrisDenseNet	0.5761	0.9107	0.7434	0.0279	0.0174	0.0103	0.7320	0.2243
FRED-Net	0.6373	0.9034	0.7703	0.0238	0.0184	0.0122	0.7276	0.2112
IrisParseNet	0.7026	0.9019	0.8023	0.0151	0.0185	0.0096	0.8751	0.1284
NIR-Zhang	0.7308	0.9394	0.8351	0.0149	0.0134	0.0072	0.9013	0.0909
ICSNet (ours)	0.7399	0.9406	0.8402	0.0135	0.0102	0.0079	0.8915	0.0913

outer circle. As described in Section 3.1, our method relies on a feature map to predict the inner and outer circles directly.

Table 3. Accuracy of iris segmentation and localization on MICHE.

Methods	Data Sets	mIoU ^{box}	mIoU ^{box}	mHdist ⁱⁿ	mHdist ^{out}	E1 ^{mask}	mIOU ^{mask}	E1 ^{norm}
FCEDN	GS	0.2358	0.7743	0.0469	0.0268	0.0062	0.6067	0.3457
	IP	0.1735	0.7651	0.0518	0.0276	0.0046	0.6186	0.3152
	GT2	0.0849	0.6488	0.1053	0.0687	0.0100	0.5240	0.3827
IrisDenseNet	GS	0.3858	0.7733	0.0328	0.0346	0.0055	0.6956	0.3018
	IP	0.2317	0.7668	0.0409	0.0409	0.0070	0.6498	0.3013
	GT2	0.3998	0.7595	0.0434	0.0415	0.0043	0.7077	0.2882
FRED-Net	GS	0.5628	0.7710	0.0306	0.0314	0.0054	0.6388	0.2717
	IP	0.4014	0.7799	0.0360	0.0325	0.0052	0.6404	0.2440
	GT2	0.3767	0.7532	0.0474	0.0442	0.0048	0.7148	0.2418
	GS	0.5353	0.7640	0.0650	0.0743	0.0046	0.7362	0.2545
IrisParseNet	IP	0.5290	0.7071	0.0740	0.0875	0.0052	0.6823	0.2753
	GT2	0.5321	0.7771	0.0439	0.0519	0.0041	0.7504	0.1867
NIR-Zhang	GS	0.5327	0.8058	0.0408	0.0326	0.0051	0.6919	0.2605
	IP	0.5369	0.8436	0.0245	0.0224	0.0037	0.7429	0.2178
	GT2	0.5930	0.8516	0.0265	0.0260	0.0053	0.7039	0.2118
ICSNet (ours)	GS	0.6406	0.8620	0.0132	0.0162	0.0050	0.7001	0.2465
	IP	0.6307	0.8734	0.0165	0.0186	0.0036	0.7431	0.2099
	GT2	0.5747	0.8606	0.0381	0.0227	0.0061	0.6857	0.2319

It can be observed that the IrisParseNet [7] and NIR-Zhang [8] consistently outperforms the other three methods in four databases, in particular, three noisy visible MICHE databases, which show that fitting the circle on the contour mask is better than fitting it on the iris mask; moreover, our approach achieved better localization results on all databases than the above five methods. On the NICE-II, ICSNet achieves 84.02% box mIoU, outperforming other methods by nearly 1.02%; on the MICHE, ICSNet outperforms other methods by nearest 9.28% avg box mIoU, which all reflect that predicting the iris circle on the feature map is more accurate and robust.

The parameters and FLOPs for different methods (resolution: 320×320 pixels) are shown in Table 4. Note that our model only needs 47.01 GFLOPs. All experiment results show that our model is well balanced in terms of accuracy and speed.

Methods	Parameters	GFLOPs
FCEDN	115.18 M	62.65
IrisDenseNet	142.95 M	75.67
IrisParseNet	125.18 M	87.03
NIR-Zhang	95.61 M	50.12
ICSNet (ours)	69.53 M	47.01

Table 4. Comparison of parameters and GFLOPs.

To better compare the iris localization and segmentation results, we select certain challenging samples and present the normalization results. The visualization results of ICSNet and other methods on the different databases are shown in Figure 7. The innerand outer-circular boundaries are marked by green and red, respectively. If locating the inner and outer circles fails, then the normalized iris result will be empty. Evidently, for the irregular and noisy iris images formed under visible light illumination in a noncooperative environment, other mask-based fitting methods cannot achieve accurate and robust iris localization; however, our double-center-based localization method can handle this situation. This result indicates that the localization method of dependent feature maps is better than the fitting method of relying on iris/contour masks.



NICE-II

Figure 7. The results of the different methods on NICE-II [31] and MICHE [32]. For non-normalized and normalized masks, blue indicates true positive pixels, green represents false positive pixels, red represents false negative pixels. If locating the inner and outer circles fails, then the normalized iris result will be empty, which is indicated by a black bar image.

The existing methods use considerable post-processing to fit the inner- and outercircular parameters on the iris or contour mask. These methods rely substantially on the segmented iris mask instead of the image feature, have weak robustness, and low localization in a non-cooperative environment. Given the effects of illumination, reflection, and noise, the shape of the segmented iris mask is irregular. Our double-center-based localization method relies on feature map rather than the segmented iris or contour mask, and the model directly predicts the inner- and outer-circle positions; therefore, our model has good localization performance and strong robustness on iris images in a non-cooperative environment with visible illumination.

5. Conclusions

In this paper, we propose a novel double-center-based efficient iris localization and segmentation network. Unlike prior research utilizing post-process methods (e.g., circle Hough transform or integro-differential) on iris or contour mask to fit the inner and outer circles, ICSNet relies on feature map to directly predict the center and regress the radius; furthermore, we propose a novel auxiliary sample strategy to accelerate model training on non-standard iris images captured in a non-cooperative environment with visible illumination. The proposed approach is compared with other methods and evaluated using four representative iris image databases. Experiments show that our proposed method has excellent performance on iris localization.

Although the proposed method has many advantages above, it also has some defects. The serial network design of first localization and then segmentation increases the time for the model to process those two tasks; moreover, the method of optimizing radius regression with L_1 loss does not take into account the pairing of the inner and outer radius of the iris. Finally, although the model has achieved optimization on most indicators, the segmentation effect on some data sets still needs to be improved.

In the future, we will explore more efficient approaches to improving segmentation performance on irregular iris image captured under visible illumination, and consider optimizing the radius regression loss function; moreover, we will consider optimizing the network structure to deal with iris localization and iris segmentation tasks in parallel to reduce the total process time.

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

Funding: This work is partially supported by grants from the Natural Science Foundation of Chongqing, China (grant no. CSTB2022NSCQ-MSX0493), and is the key project of Chongqing Technology Innovation and Application Development (grant no. cstc2021jscx-dxwtBX0018).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

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

References

- Boles, W.W.; Boashash, B. A human identification technique using images of the iris and wavelet transform. *IEEE Trans. Signal Process.* 1998, 46, 1185–1188. [CrossRef]
- 2. Jain, A.K.; Chen, H. Matching of dental X-ray images for human identification. Pattern Recognit. 2004, 37, 1519–1532. [CrossRef]
- Ma, L.; Tan, T.; Wang, Y.; Zhang, D. Efficient Iris Recognition by Characterizing Key Local Variations. *IEEE Trans. Image Process.* 2004, 13, 739–750. [CrossRef] [PubMed]
- Daugman, J. New Methods in Iris Recognition. *IEEE Trans. Syst. Man Cybern. Part B (Cybern.)* 2007, 37, 1167–1175. [CrossRef] [PubMed]
- 5. Daugman, J. How iris recognition works. In *The Essential Guide to Image Processing*; Elsevier: Amsterdam, The Netherlands, 2009; pp. 715–739. [CrossRef]
- 6. Wildes, R.P. Iris recognition: An emerging biometric technology. Proc. IEEE 1997, 85, 1348–1363. [CrossRef]
- Wang, C.; Muhammad, J.; 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]
- Wang, C.; Wang, Y.; Zhang, K.; Muhammad, J.; Lu, T.; Zhang, Q.; Tian, Q.; He, Z.; Sun, Z.; Zhang, Y.; et al. NIR Iris Challenge Evaluation in Non-cooperative Environments: Segmentation and Localization. In Proceedings of the 2021 IEEE International Joint Conference on Biometrics (IJCB), Shenzhen, China, 4–7 August 2021. [CrossRef]
- Bowyer, K.W.; Hollingsworth, K.; Flynn, P.J. Image understanding for iris biometrics: A survey. Comput. Vis. Image Underst. 2008, 110, 281–307. [CrossRef]
- Jan, F.; Usman, I.; Agha, S. Iris localization in frontal eye images for less constrained iris recognition systems. *Digit. Signal Process.* 2012, 22, 971–986. [CrossRef]
- Zhao, Z.; Kumar, A. An Accurate Iris Segmentation Framework under Relaxed Imaging Constraints Using Total Variation Model. In Proceedings of the 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 7–13 December 2015; pp. 3828–3836. [CrossRef]
- Proença, H.; Alexandre, L.A. Iris segmentation methodology for non-cooperative recognition. *IEE Proc. Vis. Image Signal Process.* 2006, 153, 199. [CrossRef]

- 13. Tan, T.; He, Z.; Sun, Z. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *Image Vis. Comput.* **2010**, *28*, 223–230. [CrossRef]
- Sutra, G.; Garcia-Salicetti, S.; Dorizzi, B. The Viterbi algorithm at different resolutions for enhanced iris segmentation. In Proceedings of the 2012 5th IAPR International Conference on Biometrics (ICB), New Delhi, India, 29 March–1 April 2012; pp. 310–316. [CrossRef]
- Liu, N.; Li, H.; Zhang, M.; Liu, J.; Sun, Z.; Tan, T. Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In Proceedings of the 2016 International Conference on Biometrics (ICB), Halmstad, Sweden, 13–16 June 2016; pp. 1–8. [CrossRef]
- Othman, N.; Dorizzi, B.; Garcia-Salicetti, S. OSIRIS: An open source iris recognition software. *Pattern Recognit. Lett.* 2016, 82, 124–131. [CrossRef]
- 17. Jalilian, E.; Uhl, A. Iris segmentation using fully convolutional encoder–decoder networks. In *Deep Learning for Biometrics*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 133–155.
- Bazrafkan, S.; Thavalengal, S.; Corcoran, P. An end to end Deep Neural Network for iris segmentation in unconstrained scenarios. Neural Netw. 2018, 106, 79–95. [CrossRef] [PubMed]
- Arsalan, M.; Naqvi, R.A.; Kim, D.S.; Nguyen, P.H.; Owais, M.; Park, K.R. IrisDenseNet: Robust Iris Segmentation Using Densely Connected Fully Convolutional Networks in the Images by Visible Light and Near-Infrared Light Camera Sensors. *Sensors* 2018, 18, 1501. [CrossRef] [PubMed]
- 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]
- Hofbauer, H.; Jalilian, E.; Uhl, A. Exploiting superior CNN-based iris segmentation for better recognition accuracy. *Pattern Recognit. Lett.* 2019, 120, 17–23. [CrossRef]
- 22. Zhao, Z.; Kumar, A. A deep learning based unified framework to detect, segment and recognize irises using spatially corresponding features. *Pattern Recognit.* 2019, 93, 546–557. [CrossRef]
- Kerrigan, D.; Trokielewicz, M.; Czajka, A.; Bowyer, K.W. Iris recognition with image segmentation employing retrained off-theshelf deep neural networks. In Proceedings of the 2019 International Conference on Biometrics (ICB), Crete, Greece, 4–7 June 2019; pp. 1–7. [CrossRef]
- 24. Hassaballah, M.; Salem, E.; Ali AM, M.; Mahmoud, M.M. Deep Recurrent Regression with a Heatmap Coupling Module for Facial Land-marks Detection. *Cogn. Comput.* 2022, 1–15. [CrossRef]
- Law, H.; Deng, J. Cornernet: Detecting objects as paired keypoints. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 734–750.
- Duan, K.; Bai, S.; Xie, L.; Qi, H.; Huang, Q.; Tian, Q. Centernet: Keypoint triplets for object detection. In Proceedings of the 2019 IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 6568–6577.
- Lin, T.-Y.; Goyal, P.; Girshick, R.; He, K.; Dollar, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–27 October 2017; pp. 2980–2988.
- Liu, Z.; Zheng, T.; Xu, G.; Yang, Z.; Liu, H.; Cai, D. Training-time-friendly network for real-time object detection. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtually, 7 February 2020; pp. 11685–11692.
- Yu, F.; Wang, D.; Shelhamer, E.; Darrell, T. Deep layer aggregation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 2403–2412.
- 30. Ma, Y.; Liu, S.; Li, Z.; Sun, J. Iqdet: Instance-wise quality distribution sampling for object detection. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 19–25 June 2021; pp. 1717–1725.
- 31. Proenca, 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.* **2009**, *32*, 1529–1535. [CrossRef]
- De Marsico, M.; Galdi, C.; Nappi, M.; Riccio, D. FIRME: Face and Iris Recognition for Mobile Engagement. *Image Vis. Comput.* 2014, 32, 1161–1172. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.