



EAI-NET: Effective and Accurate Iris Segmentation Network

Sanyam Rajpal¹, Debanjan Sadhya²(✉), Kanjar De³, Partha Pratim Roy⁴,
and Balasubramanian Raman⁴

¹ Indiana University Bloomington, Bloomington, IN, USA
srajpal@iu.edu

² ABV-Indian Institute of Information Technology and Management Gwalior,
Gwalior, India

debanjan@iiitm.ac.in

³ NTNU Norwegian University of Science and Technology, Gjøvik, Norway
kanjar.de@ntnu.no

⁴ Indian Institute of Technology Roorkee, Roorkee, India
{proy.fcs,balarfcs}@iitr.ac.in

Abstract. In iris-based biometric models, segmentation of the iris region from the rest of the eye is a crucial step. The quality of the segmented region directly affects the extracted iris features, which subsequently determines the overall recognition accuracy of the model. In this work, we propose *EAI-Net*, which is an effective and accurate iris segmentation network based on the U-Net architecture. In comparison to the previous works, we treat the segmentation process as a 3-class problem wherein the pupil, iris and the rest of the image are treated as separate classes. Furthermore, we have increased the complexity degree of our model by encoding the complex regions of the iris more efficiently. We have conducted both qualitative and quantitative assessments of our results over four benchmark iris databases - UBIRISv2, IITD, CASIAv4-Interval, and CASIAv4-Thousand. The obtained results demonstrate the superiority of our model over the other state-of-the-art deep-learning based approaches in solving the problem of iris segmentation in both the visible (VIS) and near-infrared (NIR) spectrum.

Keywords: Iris · Segmentation · U-Net

1 Introduction

Iris is the annular region in the eye which is present between the sclera and the pupil. It primarily consists of complex texture patterns which are unique to an individual. Biometric recognition systems which are operationally based on this particular trait are considered to be one of the most secure forms for entity authentication [3]. Furthermore, the advent of *mobile biometrics* has proliferated the use of these models in large-scale government and semi-government projects. Due to all these reasons, the development of accurate and robust iris-based

recognition systems which can work in unconstrained environments is an active area of research.

Segmentation of the iris region is arguably the most crucial stage in the entire recognition process. This important phase involves detecting and subsequently isolating the iris region from the corresponding input image. Importantly, the quality of the features extracted from the segmented area heavily relies on the accuracy of the associated segmentation procedure. As such, inaccurate iris segmentation results in the largest source of error for iris-based authentication models [5, 17]. The main factors which affect the segmentation process are: (i) occlusions caused due to eyelids and eyelashes, (ii) specular reflections and non-uniform illumination, (iii) imaging distance and, (iv) noise from the acquisition device (sensor) [17].

Our work in this paper proposes *EAI-Net*, which is an end-to-end deep-learning based segmentation model for non-ideal iris images that are characterized with real-world covariates such as variable imaging distances, subject perspectives and non-uniform lighting conditions. Our proposed model utilizes the U-Net architecture [18] for segmenting the iris region from their corresponding images. Importantly, this architecture can work with relatively few training images while yielding precise segmented regions. We have tested our model on four benchmark iris databases, for which our model comprehensively outperforms other deep-learning based studies.

2 Related Work

With the advent of deep neural networks, highly challenging problems in computer vision like object detection and object classification have shown excellent results. Some of the earliest works involved the use of Fully convolutional networks (FCN) [13] and Densely connected convolutional networks (DenseNet) [8] for performing the task of semantic segmentation. The use of deep-learning based models for iris segmentation was initially studied by Liu et al. [12] wherein Hierarchical convolutional neural network (HCNN) and Multi-scale fully convolutional network (MFCN) were introduced. Other deep models such as fully convolutional encoder-decoder networks [7] and a domain adaption technique for CNN based iris segmentation [6] were subsequently used in later works. Most recent works have utilized the design of Fully convolutional deep neural network (FCDNN) [1] and Generative adversarial networks (GAN) [2] for segmenting lower quality iris images which are obtained in the visible spectrum. The U-Net architecture has also been used in some previous works [11, 14]. However in our work, we have demonstrated that this architecture can give more accurate results when the iris and pupil sections of the eye are segregated. In such a scenario, the pupil is treated as a separate class and is not included with the background class. This feature facilitates the EAI-Net model in encoding the complex boundary of the iris region more accurately.

3 The EAI-Net Model

In this section, we describe in details the proposed EAI-Net model along-with the underlying U-Net architecture.

3.1 U-Net Architecture

In this paper, we have used U-Net to effectively learn the features from different regions of the eye. U-Net is one of the most popular architectures of convolutional neural networks which deals with the problem of end-to-end image segmentation. The initial U-Net model was successfully used for the segmentation of bio-medical images [18]. This architecture is basically an encoder-decoder model which consists of a contracting path (which works as an encoder) and an expanding path (which works as a decoder). Most of the operations in U-Net include convolution, which is followed by a non-linear activation function. In the contracting path, max-pooling operations are present for reducing the size of the feature maps. The expansion path consists of a sequence of up-convolutions in combination with the concatenation of high-resolution features from the contracting path. Each level in the U-Net architecture has four layer depth for extracting higher-level features from the iris. In each level of the U-Net, there is a convolution operation from a 3×3 kernel, which is followed by the ReLU activation function and batch normalization. Each max-pooling operation is performed by a factor of 2 for finding out the features at different scales. To avoid any information and content loss due to convolutions, skip connections are added. Similar to the contracting path, up-sampling with a factor of 2 is done in the expanding path for generating the upscaled maps. In each level of the expanding path, 3×3 kernel convolution operations are performed. This process is followed by the non-linear ReLU activation function and batch normalization (similar to the contracting path). We use a soft-max layer after the last convolutional operation in the expanding path for generating the final output segmentation mask. The implemented U-Net based architecture is illustrated in Fig. 1.

3.2 Pre-processing of Ground-Truth

The iris segmentation problem is generally treated as a 2-class problem where the iris is considered as the foreground and the rest of the image is considered as the background. The main issue in adopting such an approach is that the iris and pupil have similar visual appearances, for which their exact discrimination becomes very difficult. To address this problem, we modify the problem into a 3-class problem where the pupil and iris are treated as separate classes. This process enables the deep-neural network to learn distinguishing features between the iris and the pupil, which subsequently results in a more accurate segmentation of the iris region. We achieve this particular objective in our work by using elements from computational geometry. Specifically speaking, we convert the binary problem into a 3-class problem using a combination of convex hulls, fitting contours and morphological operations. Furthermore, we had to use a combination of the

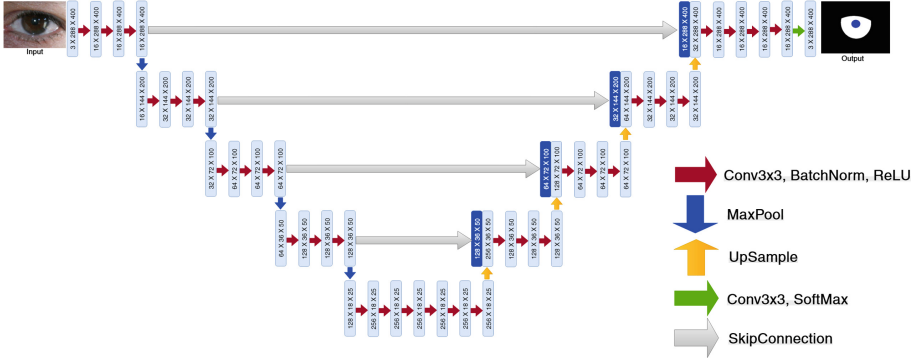


Fig. 1. The proposed *EAI-Net* model for iris segmentation.

convex hull with concave hull [15] and the morphological closing operation for generating the augmented ground-truth for the CASIAv4-T database. The reason for using these additional pre-processing operations was due to the presence of some poorly labeled noisy samples in this particular database. The process of generating the 3-class ground-truth where the classes are labeled as 0 (for background), 1 (for iris), and 2 (for pupil) is presented in Algorithm 1.

Algorithm 1. Generation of augmented Ground-truth

Input: $M \in \{0, 1\}^{M \times N}$

Output: $GT \in \{0, 1, 2\}^{M \times N}$

- 1: $GT \leftarrow 0$
 - 2: $contour \leftarrow findContour(M)$ using [19]
 - 3: $Area_List \leftarrow []$
 - 4: **for** $i \leftarrow 1$ to $|contour|$ **do**
 - 5: $Area_List.append(contour[i], i)$
 - 6: $sorted_List, index \leftarrow SORT(Area_List)$
 - 7: **for** $j \leftarrow 1$ to 2 **do**
 - 8: $Hull_Points \leftarrow CONVEX_HULL(contour(index[j]))$ using [4]
 - 9: $GT \leftarrow Fill_Contour(Hull_Points, GT, j)$
 - 10: **return** GT
-

4 Experimental Setup

In this section, we describe the experimental datasets and associated quantitative measures. We also elaborate on the network training process.

4.1 Database Description

We have performed extensive experiments on the following four publicly available benchmark iris databases: IITD-1 [10], UBIRISv2 [16], CASIAv4-Interval (further referred to as CASIAv4-I) and CASIAv4-Thousand (further referred to as CASIAv4-T)¹. We have specifically selected these four databases for validating our work due to the variability of both image quality and quantity in them. The ground-truth masks of the IITD, CASIAv4-I and UBIRISv2 database are provided by the University of Salzburg via their *IRISSEG-EP* package [5]². Alternatively, the ground-truth masks for the CASIAv4-T database are distributed by Bezerra et al. [2]. However, it should be noted that the ground-truths corresponding to all the images of the respective databases are not provided. For instance, the total number of available annotations for UBIRISv2 and CASIAv4-T are 2250 and 1000 respectively.

4.2 Evaluation Protocol and Metrics

To evaluate the performance of EAI-Net, we use the following statistical quantities: **NICE-I**, **NICE-II** [7], and **F1-Score**. The NICE-I and NICE-II scores represent the overall segmentation errors between the segmentation mask (obtained from the network) and the corresponding ground-truth mask. The NICE-I score estimates the segmentation error by computing the proportion of the disagreeing pixels between the two masks, whereas the NICE-II score is intended to balance the disproportion between the prior probabilities of iris and non-iris pixels in the images. The F1-Score is a standard measure of the segmentation accuracy. It represents the harmonic mean of the corresponding precision and recall values. All these three metrics are bounded in the range [0, 1].

4.3 Model Training Details

The entire framework for supervised iris segmentation has been implemented in Pytorch. Information like the number of channels, the number of filters, the type of connection and activation functions are visually depicted in Fig. 1. The receptive field has been kept identical for implementation in the different datasets. The batch size for training was kept at 4. All the experiments were conducted on a computer having Intel Xeon E5 processor with NVIDIA Quadro K620 2GB RAM graphics card. The model takes around 25 epochs to converge. We have used the *Adam Optimizer* [9] for conducting all the experiments. The hyperparameters associated with this optimizer include *learning rate* = 0.0001, $\beta_1 = 0.9$, and $\beta_2 = 0.999$. The learning rate was multiplied with 0.5 every time the validation loss did not decrease (validation was done after every 150 iterations). For training the U-Net, we have chosen *Categorical cross entropy* as the loss function.

¹ <http://biometrics.idealtest.org/>.

² <http://www.wavelab.at/sources/Hofbauer14b/>.

5 Results and Discussions

Now we present and analyze all of our obtained results. In accordance with the previous works, we perform both quantitative and qualitative assessment of our results.

5.1 Ablation Study

We initially perform an ablation study by comparing the traditional 2-class segmentation problem with the 3-class problem. As presented in Table 1, some improvements in performance can be immediately noticed when the iris, the pupil and the background were considered as separate classes. Specifically speaking, both the NICE-I and NICE-II error scores were relatively lower and the F1 score was comparatively higher for the 3-class problem. This trend was consistently noted for all the four iris databases. Hence these results vindicate the importance of segmenting the entire eye image into three distinct classes (instead of two).

Table 1. Average values of the evaluation metrics while considering 2-class and 3-class segmentation problems.

Database	NICE-I		NICE-II		F1-Score	
	2-class	3class	2-class	3-class	2-class	3-class
CASIAv4-I	0.0156	0.0152	0.0193	0.0193	0.9813	0.9842
IITD	0.0192	0.0185	0.0242	0.0231	0.9684	0.9764
UBIRISv2	0.0082	0.0073	0.0331	0.0316	0.9692	0.9699
CASIAv4-T	0.007	0.0054	0.0227	0.0204	0.9586	0.9785

5.2 Quantitative Evaluation

We quantitatively compare the performance of EAI-Net with the other state-of-the-art deep-learning based iris segmentation techniques. For evaluation purpose, we have used the performance measures explained previously in Sect. 4.2. The mean (μ) and standard deviation (σ) of these measures are presented in Table 2.

As observable, the best F1 Score of 0.9842 was obtained for CASIAv4-I, which indicates the presence of high precision and recall values. Alternatively, the least F1 Score of 0.9699 was noticed for the UBIRISv2 databases, which denotes relatively poor segmentation of the iris regions. This result can be aptly justified due to the presence of off-angle noisy iris samples in this database. Interestingly, low NICE-I scores of 0.0054 and 0.0073 were noticed for the CASIAv4-T and UBIRISv2 databases respectively. This particular outcome can be attributed to the fact that the area of the iris region is comparatively much smaller in the samples of these datasets. This resulted in a lesser number of disagreeing pixels

Table 2. Mean (μ) and standard deviation (σ) values of the evaluation metrics.

Database	NICE-I		NICE-II		F1-Score	
	μ	σ	μ	σ	μ	σ
CASIAv4-I	0.0152	0.0059	0.0193	0.0088	0.9842	0.0098
IITD	0.0185	0.0074	0.0231	0.0116	0.9764	0.0129
UBIRISv2	0.0073	0.0082	0.0316	0.0341	0.9699	0.0572
CASIAv4-T	0.0054	0.0036	0.0204	0.0195	0.9785	0.0337

between the ground-truth and the corresponding predicted mask, which consequently produced low NICE-I scores. Another noticeable observation pertains to the CASIAv4-T database. Although this database is characterized by covariates such as specular reflection and non-uniform illumination (much like UBIRISv2), the corresponding F1 score of 0.9785 is relatively high. One possible reason for this result might relate to its associated spectral band. Since all of the images for this database were captured in NIR, the iris regions had more richly structured textural information which the EAI-Net exploited.

The superiority of our framework over the other deep-learning based techniques is demonstrated in Table 3. For all the iris databases, our model results in comparatively better values of NICE-I, NICE-II and F1 Score. The best improvement in the segmentation error corresponded to the UBIRISv2 database, wherein a decrease of approximately 18.88% over the next best (lowest) reported result [12] was noted. Considering the quality of the samples in this database, this is a considerable improvement over the previous results. The only anomaly was noticed for the IITD database, for which a smaller error score of 0.0133 was observed in the GAN model [2]. However, it should be noticed that our U-Net based model is relatively more efficient than GAN in terms of the required memory resources.

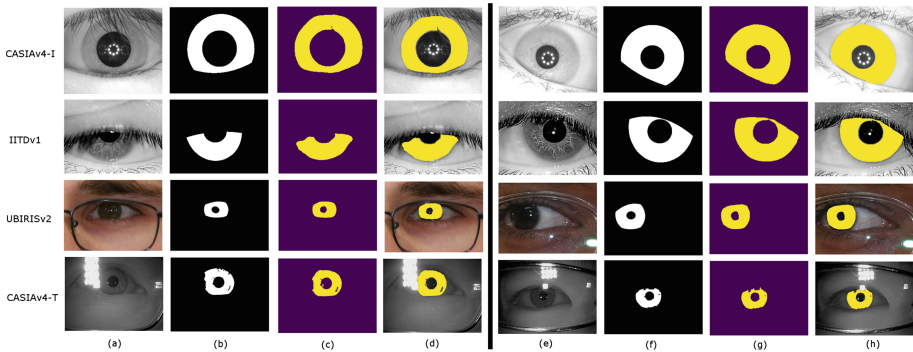
5.3 Qualitative Evaluation

Now we visually analyze a few instances of the iris segmentation results given by our model. Figure 2 illustrates sample results from the four databases used for our evaluation. As expected, the EAI-Net model gives excellent results for the CASIAv4-I and IITD datasets. Although both the UBIRISv2 and CASIAv4-T are very challenging iris dataset, EAI-Net works well on them too. As understandable from Fig. 2, our model effectively handles samples from both the VIS and NIR spectrum. Important covariates such as imaging-distance and camera angle are also efficiently supervised by our model.

The segmentation errors for some noisy samples are illustrated in Fig. 3. The EAI-Net model is unable to accurately segment the iris regions when it is affected by strong reflections and drooping eyelashes. Due to this reason, pre-processing these iris samples for eliminating the effects of these covariates would potentially improve the segmentation accuracy of our network. Noticeably, the sample from

Table 3. Comparative analysis of the average segmentation scores for the four iris databases.

Database	Technique	NICE-I	NICE-II	F1-Score
CASIAv4-I	Proposed EAI-Net	0.0152	0.0193	0.9842
	FCEDN (Basic variant) [7]	0.033	0.0382	0.9375
	FCEDN (Bayesian variant) [7]	0.0412	0.0362	0.925
	Domain adaptation [6]	0.033	0.038	0.937
IITD	Proposed EAI-Net	0.0185	0.0231	0.9764
	FCEDN (Basic variant) [7]	0.0277	0.0322	0.951
	FCEDN (Bayesian variant) [7]	0.0292	0.0337	0.95
	FCN [2]	0.0148	-	0.9744
	GAN [2]	0.0133	-	0.9584
	Domain adaptation [6]	0.027	0.032	0.951
UBIRISv2	Proposed EAI-Net	0.0073	0.0316	0.9699
	FCEDN (Basic variant) [7]	0.0262	0.0687	0.79
	FCEDN (Bayesian variant) [7]	0.0187	0.0675	0.8625
	FCN [2]	0.01	-	0.882
	GAN [2]	0.03	-	0.9142
	MFCN [12]	0.009	-	-
	HCNN [12]	0.011	-	-
CASIAv4-T	Proposed EAI-Net	0.0054	0.0204	0.9785
	FCN [2]	0.0061	-	0.9442
	GAN [2]	0.014	-	0.9538

**Fig. 2.** Qualitative analysis of the segmentation results for some selected iris samples. Columns (a), (e) represent the original images, columns (b), (f) represent the available ground-truths, columns (c), (g) represent the predicted iris masks, and columns (d), (h) represent the segmented iris region (yellow section). All the images are scaled uniformly for representational purpose (Color figure online).

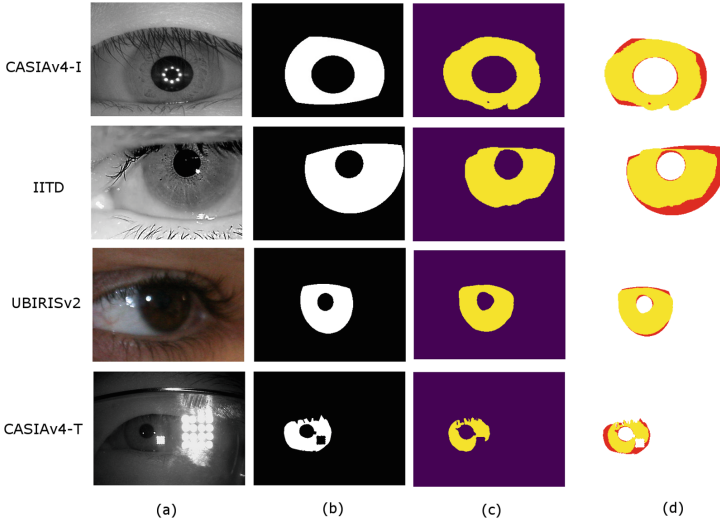


Fig. 3. Segmentation results for some selected noisy iris samples. Column (a) represents the original images, column (b) represents the corresponding ground-truths, column (c) represents the predicted iris masks, and column (d) shows the error in the predicted masks (red regions) (Color figure online).

the UBIRISv2 database is additionally characterized with low contrast since the entire UBIRISv2 database was collected in the VIS spectrum.

6 Conclusion

Our work in this paper introduces the *EAI-Net* model for accurately segmenting the iris region from eye images. While using conventional deep architectures, this problem is generally treated as a 2-class problem where the iris is considered as the foreground and rest of the eye is considered as the background. However, our proposed technique uses a combination of computational geometry techniques and morphological operations for pre-processing the ground-truth of the data while separating the pupil from iris. This 3-class ground-truth is subsequently used for training the U-Net architecture whose receptive fields have been calculated for accurately recognizing the structure of the iris. We have performed extensive empirical tests on four benchmark iris databases for demonstrating the efficacy of our model in both the visible and NIR spectrum. Importantly, *EAI-Net* is able to accurately segment the iris region for two of the most challenging iris databases, namely UBIRISv2 and CASIAv4-T. In the future extension of our work, we would investigate this model in combination with region proposal networks for extracting the iris region after initially localizing the eyes. Furthermore, we would like to focus on developing strategies that seek to optimize performance and computational aspects of the used architecture.

References

1. Bazrafkan, S., Thavalengal, S., Corcoran, P.: An end to end deep neural network for iris segmentation in unconstrained scenarios. *Neural Netw.* **106**, 79–95 (2018)
2. Bezerra, C.S., et al.: Robust iris segmentation based on fully convolutional networks and generative adversarial networks. CoRR. <http://arxiv.org/abs/1809.00769> (2018)
3. Daugman, J.: Information theory and the iriscodes. *IEEE Trans. Inf. Forensics Secur.* **11**(2), 400–409 (2016)
4. Graham, R.L., Yao, F.F.: Finding the convex hull of a simple polygon. *J. Algorithms* **4**(4), 324–331 (1983)
5. Hofbauer, H., Alonso-Fernandez, F., Wild, P., Bigun, J., Uhl, A.: A ground truth for iris segmentation. In: 2014 22nd International Conference on Pattern Recognition, pp. 527–532 (August 2014)
6. Jalilian, E., Uhl, A., Kwitt, R.: Domain adaptation for CNN based iris segmentation. In: 2017 International Conference of the Biometrics Special Interest Group (BIOSIG), pp. 1–6 (September 2017)
7. Jalilian, E., Uhl, A.: Iris segmentation using fully convolutional encoder–decoder networks. In: Bhanu, B., Kumar, A. (eds.) *Deep Learning for Biometrics*. ACVPR, pp. 133–155. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-61657-5_6
8. Jégou, S., Drozdal, M., Vazquez, D., Romero, A., Bengio, Y.: The one hundred layers tiramisu: fully convolutional densenets for semantic segmentation. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1175–1183. IEEE (2017)
9. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint. [arXiv:1412.6980](https://arxiv.org/abs/1412.6980) (2014)
10. Kumar, A., Passi, A.: Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recogn.* **43**(3), 1016–1026 (2010)
11. Lian, S., Luo, Z., Zhong, Z., Lin, X., Su, S., Li, S.: Attention guided U-Net for accurate iris segmentation. *J. Vis. Commun. Image Represent.* **56**, 296–304 (2018)
12. Liu, N., Li, H., Zhang, M., Liu, J., Sun, Z., Tan, T.: Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In: 2016 International Conference on Biometrics (ICB), pp. 1–8 (June 2016)
13. Long, J., Shelhamer, E., Darrell, T.: Fully convolutional networks for semantic segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3431–3440 (2015)
14. Lozej, J., Meden, B., Struc, V., Peer, P.: End-to-end iris segmentation using U-Net. In: 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOB), pp. 1–6 (July 2018)
15. Moreira, A., Santos, M.Y.: Concave hull: a k-nearest neighbours approach for the computation of the region occupied by a set of points (2007)
16. 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.* **32**(8), 1529–1535 (2010)
17. Proenca, H., Alexandre, L.A.: Iris recognition: analysis of the error rates regarding the accuracy of the segmentation stage. *Image Vis. Comput.* **28**(1), 202–206 (2010)
18. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015*. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
19. Suzuki, S., et al.: Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.* **30**(1), 32–46 (1985)