Iris segmentation in an image using U-Net convolutional neural network architecture

Yulia Ganeeva
Geoinformatics and Information Security department
Samara National Research University
Samara, Russia
jganeeva99@mail.ru

Evgeny Myasnikov
Geoinformatics and Information Security department
Samara National Research University;
Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS
Samara, Russia
mevg@geosamara.ru

***Abstract*—The accuracy of segmentation plays an important role in the methods for personal identification by iris images. In this paper, we study the iris segmentation method based on the convolutional neural network of the U-Net architecture. As part of the research, manual segmentation of the images of the used dataset was performed, the optimal network training parameters were determined, and the segmentation quality was evaluated. All studies in the paper were conducted using the open MMU Iris Image Database. The results showed that the studied approach provides high precision segmentation of the iris images.**

Keywords—accuracy, convolutional neural network, iris, segmentation, U-Net.

I. Introduction

Recently, the most widespread methods of identifying a person use biometric data (face images [1, 2], hand geometry [3, 4], fingerprints [5, 6], iris, etc.). Iris identification is one of the most accurate and reliable methods of biometric identification since the texture of the iris is remarkably stable over time. The accuracy of segmentation plays an important role in achieving high-quality indicators of identification by the iris image. Since the iris is an annular area between the pupil and the sclera in the image, the solution to the segmentation problem is often reduced to approximating the inner and outer borders of the iris with circles. For this purpose, most works use the Daugman method [7], and the Hough transform for the detection of circles [8], as well as other methods, for example, based on the analysis of the distribution of boundary points [9].

Unfortunately, the segmentation of the iris performed by these methods is often not accurate enough due to the partial overlap of the iris with the eyelids and eyelashes, as well as the appearance of glare from the light source. Also, the shape of the iris itself is not always well approximated by circles. Thus, a relevant task is to improve the quality of iris segmentation.

With the success of deep neural network models, researchers are increasingly turning their attention to convolutional neural networks to further improve the accuracy of existing iris segmentation methods. To solve the problem of iris segmentation, we propose to use an approach based on training the convolutional neural network (CNN) of the U-Net architecture [11,12].

U-Net is considered one of the standard CNN architectures for image segmentation tasks when you need not only to define the entire image class but also to segment regions, i.e. create a mask that will divide the image into several classes. One of the main advantages of the network, which is essential in the context of this work, is the ability to train it on a small amount of data.

The paper has the following structure. Section 2 is devoted to the description of the studied approach to iris image segmentation, a description of the network architecture, the features of data preparation, and training. Section 3 describes the dataset used in the paper and provides the results of experimental studies. The paper ends up with the conclusion and reference list.

II. Method

A. *The U-Net Architecture*

U-Net is a famous CNN architecture for solving biomedical problems (segmentation of various types of cells, determining the boundaries between dense cell structures, etc.) and other image segmentation tasks. The main advantage of this model is the ability to learn on small datasets and, at the same time, show high accuracy, which is a common problem for computer vision problems.

The network architecture is shown in Fig. 1.

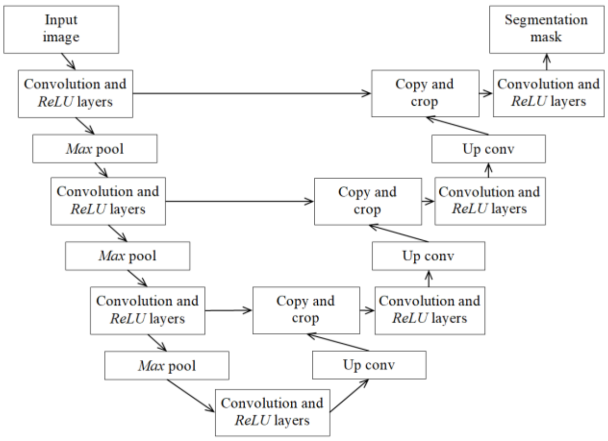

Fig. 1. The U-Net architecture [11].

As you can see in the figure, the architecture consists of a contracting path (left) to capture the context and an expanding path (right) to allow precise localization. There are two classes in the paper: iris and non-iris.

The contracting path corresponds to the typical structure of a convolutional neural network. It consists of re-applying two convolutions (kernel size 3x3), followed by a layer of ReLU and max-pooling (kernel size 2x2) with step 2.

Each step of the expanding path contains an expanding feature map layer (up-conv), the reverse of the max-pooling, followed by a convolution (kernel size 2x2) that reduces the number of feature channels. This is followed by concatenation with the appropriate cropped feature map (copy and crop) from the contracting path and two convolutional layers (kernel size 3x3), each of which is followed by a ReLU layer.

B. Model training and preparatory stages

At the stage of preparation for training, a pre-trained model was taken [13]. The dataset was divided into training and validation sets. The color source images were converted to grayscale, and reduced to a size of 320 x 320 pixels. Pixel values were scaled to the range [0;1]

Binary cross-entropy was used as a loss function, which determines how the network quality is estimated on training data. Adam was chosen as the optimizer determining the mechanism for updating the weights [14].

III. EXPERIMENTS AND RESULTS

For experimental research, implementation in Python was performed using the Keras, Numpy, OpenCV, and Matplotlib libraries.

The MMU Iris Database [15] was used as a dataset of iris images. The base contained 5 images for the left and right eyes of 45 people, which gave a total of 450 images. To train and evaluate the quality of segmentation, manual segmentation of all images from the set was performed.

Nine experiments were conducted to determine the best parameters. The neural network was trained using the found parameters for 50 epochs. At the stage of searching for the best parameters during training, the pre-trained neural network [13, 16] was configured on the above data with the following parameters: the learning rate for the first, second and third experiments was equal to 10^{-3}, for the fourth, fifth and sixth - 10^{-4}, for the seventh, eighth and ninth - 10^{-5}, the number of training epochs (epochs) was 10, the size of the random subsample (batch size) used to estimate the gradient for the first, fourth, and seventh experiments was 10, the second, fifth, and eighth - 20, and the third, sixth, and ninth - 40.

To monitor the training process, the value of cross-entropy and segmentation accuracy was evaluated using the validation set at the end of each epoch of training.

Segmentation accuracy was estimated as the proportion of correctly classified pixels in the images of the validation set.

The values of the loss function and the accuracy of the trained model on the training and validation data for various values of the batch size and learning rate hyperparameters are shown in Table 1.

The plot of changes in the accuracy values and loss function in the first 10 epochs of neural network training for the two best models can be seen below in figures 2 and 3.

As a result of experiments, the following values of training parameters were selected: learning rate was equal to 10^{-4}, the size of a random sub-sample (batch size) was 40. Training a neural network with the specified parameters for 50 epochs allowed us to achieve accuracy and loss function values equal to 99.53 and 0.0071, respectively, on the validation sample.

The results of the studied approach for the resulting model are shown below in figure 4.

TABLE I. TRAINING RESULTS FOR DETERMINING THE BEST PARAMETERS

	Learning rate	*Batch size*	*Accuracy*	*Value of the loss function*
Training data	10^{-3}	10	97.71	0.0262
		20	97.84	0.0293
		40	95.61	0.0771
	10^{-4}	10	98.04	0.0165
		20	97.98	0.0178
		40	98.04	0.0182
	10^{-5}	10	98.64	0.0135
		20	98.85	0.0142
		40	97.76	0.0384
Validation data	10^{-3}	10	98.55	0.0245
		20	96.23	0.0322
		40	95.61	0.0812
	10^{-4}	10	99.11	0.015
		20	99.05	0.016
		40	99.47	0.009
	10^{-5}	10	99.19	0.0127
		20	96.89	0.0153
		40	99.03	0.0255

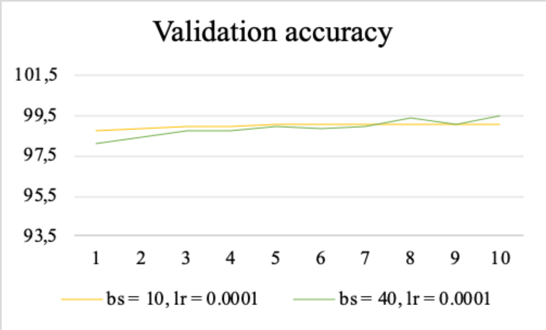

Fig. 2. Neural network training: plot of accuracy dependence on the epoch of neural network training.

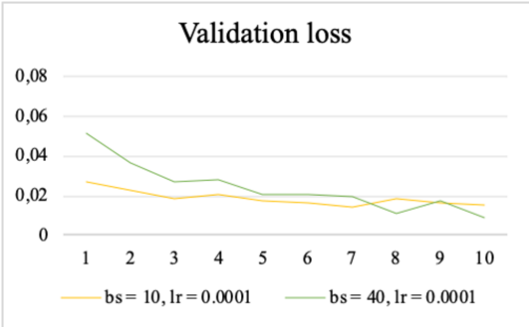

Fig. 3. Neural network training: plot of loss function dependence on the epoch of neural network training.

Besides, in this paper, we provide an example of iris segmentation using the previously developed technique [10].

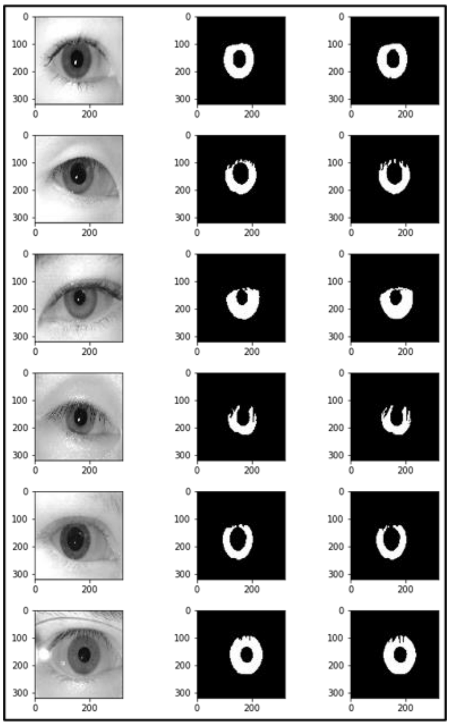

Fig. 4. Examples of image segmentation of the iris: the original image (left column), the segmentation mask generated using the studied approach (middle column), and the correct segmentation mask (right column).

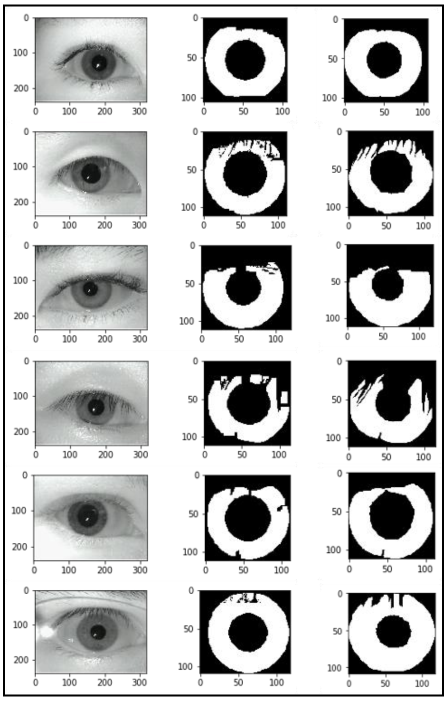

Fig. 5. Examples of image segmentation of the iris: the original image (left column), the segmentation mask generated using the previously developed approach (middle column), and the correct segmentation mask (right column).

A brief description of the previously proposed technique is given below:

1. Determination of the centers and radii of two circles approximated the pupil and the iris using the Hough transform.

2. Crop the image.

3. The resulting image is blurred to eliminate minor noise and smooth the iris.

4. Conversion to a monochrome image.

5. Contour detection.

6. Morphological transformation (closing gaps).

7. Construction of the mask of the open area of the iris (each pixel has a value of 1 if this pixel belongs to the iris, and 0 otherwise).

The use of the aforementioned technique allowed obtaining 90.67% of segment accuracy.

The accuracy of the method considered in this paper is 99.53%, which is 8.86% higher, the use of a CNN significantly compared to the previously proposed technique.

IV. Conclusion

In this paper, we study the iris segmentation method based on the convolutional neural network of the U-Net architecture. To configure and evaluate the method, manual segmentation of images from the open MMU Iris Image Database was performed.

Using the specified dataset, the best values of the training parameters were determined, and segmentation quality was estimated. Studies have shown that the approach under study allows us to achieve 99.53% segmentation on the validation dataset.

Acknowledgment

The work was partly funded by RFBR according to the research project 18-01-00748 in parts of «II. Method» - «III. Experiments and Results» and by the Russian Federation Ministry of Science and Higher Education within a state contract with the «Crystallography and Photonics» Research Center of the RAS in parts «I. Introduction» and «IV. Conclusion».

References

[1] V.B. Nemirovskiy, A.K. Stoyanov and D.S. Goremykina, “Face recognition based on the proximity measure clustering,” Computer Optics, vol. 40, no. 5, pp. 740-745, 2016. DOI: 10.18287/2412-6179-2016-40-5-740-745.

[2] Yu.V. Vizilter, V.S. Gorbatsevich, A.V. Vorotnikov and N.A. Kostromov, “Real-time face identification via CNN and boosted hashing forest,” Computer Optics, vol. 41, no. 2, pp. 254-265, 2017. DOI: 10.18287/2412-6179-2017-41-2-254-265.

[3] J. Hashemi and E. Fatemizadeh, “Biometric identification through hand geometry,” EUROCON – Int. Conf. ‘Computer as a Tool’, vol. 2, pp. 1011-1014, 2005.

[4] S.M. Prasad, V.K. Govindan and P.S. Sathidevi, “Bimodal personal recognition using hand images,” Proc. Int. Conf. Advances in Computing Communication and Control (ICAC3), pp. 403-409, 2009.

[5] W. Yuan, Y. Lixiu and Zh. Fuqiang, “A real time fingerprint recognition system based on novel fingerprint matching strategy,” 8th International Conference on Electronic Measurement and Instruments, pp. 1-81, 2007.

[6] M. Kaur, M. Singh, A. Girdhar and S. Parvinder, “Fingerprint verification system using minutiae extraction technique,” World Academy of Science Engineering and Technology, vol. 46, pp. 497-502, 2008.

[7] J.G. Daugman, “How iris recognition works,” Proc. IEEE Trans. Circ. Syst. Video Technol., vol. 14, no. 1, pp. 21-30, 2004.

[8] R. Gonzalez and R. Woods, “Digital Image Processing,” M: Technosphere, 2005, 1072 p.

[9] M.S. Semenov and E.V. Myasnikov, “A comparison of iris image segmentation techniques,” CEUR Workshop Proceedings, vol. 2210, pp. 163-169, 2018.

[10] Y.K. Ganeeva, V.A. Beketova and E.V. Myasnikov, “Development and software implementation of the method of segmentation of the iris in the image,” International Youth Scientific Conference "XV Korolev Readings", dedicated to the 100th anniversary of the birth of D.I. Kozlova, vol. 1, pp. 482-483, 2019.

[11] O. Ronneberger, P. Fischer and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” MICCAI, p. 234-241, 2015.

[12] A. V. Nikonorov, M. V. Petrov, S. A. Bibikov, V. V. Kutikova, A. A. Morozov and N. L. Kazanskiy, “Image restoration in diffractive optical systems using deep learning and deconvolution,” Computer Optics, vol. 41, no. 6, pp. 875-887, 2017. DOI: 10.18287/2412-6179-2017-41-6-875-887.

[13] Pre-trained U-Net model [Online]. URL: https://github.com/jus390/U-net-Iris-segmentation (12/23/2019).

[14] D. Kingma and J. Ba, “Adam: A Method for Stochastic Optimization,” arXiv: 1412.6980v8, 2017.

[15] MMU Iris Image Database: Multimedia University [Online]. URL: http://pesonna.mmu.edu.my/ccteo/.

[16] J. Lozej, B. Meden, V. Struc and P. Peer, “End-to-End Iris Segmentation Using U-Net,” IEEE International Work Conference on Bioinspired Intelligence (IWOBI), pp. 1-6, 2018