Study on Iris segmentation Algorithm based on Dense U-Net

Xiaoqiang Wu¹, Long Zhao²

¹College of Mechanical Engineering, Inner Mongolia University for the Nationalities, Tongliao 028000, China
²School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences), 250353, Jinan, China

Corresponding author: Long Zhao (e-mail: zhao_long@163.com)

This work was supported by Joint special of Shandong Natural Science Foundation (ZR2017LF020), in part by the National Natural Science Foundation of China (61806105), Supported by Program for Young Talents of Science and Technology in Universities of Inner Mongolia Autonomous Region' NJYT-19-B15.

ABSTRACT Iris segmentation is an essential process of iris recognition. Iris segmentation plays an important role in maintaining the accuracy of iris based on recognition system by limiting the errors in the current stage. However, its performance is affected by non-ideal conditions caused by ambient light noise and user non-cooperation. The existing segmentation methods based on local features cannot find the real iris boundary under these non-ideal conditions, and the errors generated in the segmentation stage will traverse to all subsequent stages, resulting in decreased accuracy and reliability. In addition, real iris boundaries need to be divided without additional denoising costs. Aiming at the problems of existing algorithms in complex scenes and cross-device applications, an Iris segmentation algorithm based on Dense U-Net is presented in this paper. Combining Dense network with U-Net network, Iris is segmented by taking advantage of dense U-Net network, which is narrower and has fewer parameters, and taking advantage of U-Net in semantic segmentation. Dense connected path is derived from dense connected network (Dense U-Net), in which improved information and parameters are helpful to reduce the training difficulty of deep network. The final segmentation accuracy was 98.36%. F1 is 97.07%. The experimental results prove the presented model can improve the accuracy, reduce the error rate, and assist doctors in the diagnosis of Iris Diseases effectively.

INDEX TERMS Iris Segmentation; Dense network; U-Net; accuracy; Iris Diseases

I. INTRODUCTION

Image segmentation is an important part of image processing. Based on texture, Gray scale, color, and shape, it can be divided into a number of areas with unique properties, and then the region of interest for further image processing. Image segmentation is a key step from image processing to image analysis. With the rapid development of computer and network communication technology and the popularization of big data, more and more activities related to identity authentication are involved in people's daily life and work, such as website login, e-commerce, banking, attendance system and so on. Identity authentication is to confirm the uniqueness and legitimacy of an individual's identity through a series of methods. Traditional authentication methods include token-based authentication system, such as identity card, passport, driver's license, and knowledge-based authentication system, such as various account passwords or personal identification numbers. Nowadays, with the rapid development of network communication technology, the traditional authentication method has many shortcomings, such as easy to lose, easy to forget, easy to be stolen and forged, which has been difficult to meet the growing demand for information security. Therefore, biometric authentication technology emerges as the times require. Because biometric features are unique and immutable for individuals, and there are no shortcomings such as loss or theft, biometric authentication system is more reliable and convenient than traditional authentication methods, and has great potential for development. Iris has the advantages of uniqueness, non-variability and easy collection [1], Iris identity recognition plays a more and more important role in the field of biometrics. Iris recognition is regarded as one of the human body identifiers because of its uniqueness and stability, which attracts more and more researchers' attention. Iris feature extraction mainly depends on whether the iris can be accurately located and segmented, and the quality of Iris segmentation is related to the research and analysis of the subsequent results. Generated by the complex collection environment, easy to cause the light intensity is not consistent, the upper and lower eyelid occlusion, shooting
angle caused by the mold, squint and other effects of complex segmentation. The process of Iris Outer Edge Search is complicated, time-consuming and biased. And Iris Inside Edge, suffer PALPEBRAL, eyelash to wait to affect lesser commonly but its occurrence contracts or dilate easily. In order to make the Iris Recognition Technology Work Normally, Iris segmentation is necessary for the iris image which usually contains other parts besides iris.

Due to the influence of eyelash occlusion, light reflection, shooting Angle and other noises, the acquired iris image is incomplete. Therefore, we divide the iris image into ideal iris and non-ideal iris, as shown in Figure 1. The ideal iris image is the complete iris image without occlusion. The non-ideal iris image is the iris image with occlusion.

![Figure 1. Schematic of iris type.](image)

(a) Ideal iris image; (b) The upper eyelid obscures the image (c) Lower eyelids block the image (d) Upper and lower eyelids block images

II. RELATED WORK

At present, there are many iris segmentation Algorithms, such as ring-shaped integro-differential Algorithm proposed by Daugman et al. [2], and algorithm that the edge detection is combined with Hough transform proposed by Wildes et al. [3,4] is the most representative method to locate the Iris region of human eye. Hilal et al. [5,8] proposed an improved Hough transform for Iris Segmentation, Zeng et al. [9] proposed an Iris segmentation method based on Image Alignment Huang et al. [10] proposed a self-affine fitting texture segmentation algorithm for rough classification of Iris images to detect the outer edge of the IRIS loop. Xu et al. [11] used geometric gray-scale projection to locate the inner edge of the IRIS. Although there are many iris segmentation Algorithms, domestic and foreign scholars have also proposed many improved algorithms, but the mainstream algorithm is generally based on fixed threshold image binarization, edge detection, Hough Transform, etc. And the fact that the center of the pupil does not coincide with the center of the Iris is not taken into account. This method needs to set the parameters artificially, the final accuracy of Iris segmentation is usually affected to some extent, and the robustness is not strong.

With the emergence of deep neural networks, a number of scholars have applied convolutional neural network (CNN) to Iris segmentation [12-14]. As we all know, CNN is a typical deep learning model. The segmentation of Iris Image using CNN can reduce the process of feature extraction and selection, and further improve the final accuracy. Liu et al. [15] has proposed a hierarchical convolutional neural network and multi-scale full convolution network (MFCN) to process noisy iris images. Jalilian et al. [16] has proposed three full convolution networks (FCEDN) for Iris segmentation. Yang et al. [17] has presented a model for Iris segmentation based on FCN and dilatation convolution, and trained and tested it on CASIA-IRIS-INTERVALO-V4.0 and Ubiris. Shabab et al. [18] proposed an end-to-end CNN for Iris Image Segmentation with low-quality, and achieved a good result. Lozej et al. [19] implemented the segmentation using U-Net model. Hofbauer et al. [20] used marked Iris images for Iris segmentation. However, although these methods proposed may improve one of the performances, such as accuracy and error rate, while improving one of the indicators, it may reduce other indicators. The generalization ability of the network is not strong.

In order to solve these problems and improve the accuracy of Iris segmentation, we combine Densenet with U-Net, and propose a new network structure, which not only reduces the network parameters, but also takes advantage of U-Net in semantic segmentation. Dense U-Net integrates dense connectivity into U-Net's contraction and expansion paths. Compared with traditional CNNs, dense connectivity reduces learning redundancy, enhances information flow, and reduces the number of parameters needed to achieve similar or better performance. Having a more compact CNN and fewer parameters requires less computation, resulting in faster image segmentation.

III. NETWORK ARCHITECTURE

We combined the Densenet and U-Net networks to add the dense blocks in the contraction and expansion paths.

A. Densenet

Convolutional neural network is the "brain" of most current artificial intelligence systems and is widely used in computer vision, speech recognition, natural language processing and
other fields. The rapid development of artificial intelligence in recent years is closely related to the innovation and breakthrough of deep convolutional network. Alpha Go, based on convolutional networks, beat out top human players in board game, the highest human intelligence game, and the automated medical imaging system, also based on convolutional networks has outperformed imaging specialists in lung disease detection, and has shown promise and value in areas such as autopilot, security, and finance.

In recent years, with the development of research, the research of dense connection convolutional neural network has made great achievements, which innovatively expands the one-dimensional layer-by-layer connection mode of traditional neural network to the three-dimensional cross-layer dense connection mode (Figure. 2). In the traditional neural network model, each layer of the network only gets the input signal from the upper layer, and then transmits the extracted features to the next layer. Dense U-Net is different: first, there is a direct connection between any two layers of the network, that is, the input of each layer of the network is a union of the outputs of all the previous layers. Each layer of Dense U-Net learns only a very limited number of features, which are passed directly to all the layers behind it as input. In order to reduce redundancy. In fact, the first is the premise of the second, there is no dense connection, each layer of the network must learn a lot of features, otherwise the training will appear under-fitting phenomenon. According to this new network architecture, Dense U-Net also proposed the concept of bottleneck layer and transition layer, which can compress the computation of each layer locally and improve the efficiency of feature Reuse Globally Thus further excavates and expands this model the superiority.

B. U-Net

U-Net includes the contraction path on the left and the expansion path on the right. The contraction path is a structure (stride: 2) which consists of several 33 convolution-added RELU active layers and 22 maxpooling structure (stride: 2). Each step of the expansion path consists of an upper sample, a 22 convolution (reducing the number of channels by half), a series of trimmed feature layers in the corresponding contraction path, and two 3 convolution relus. The last layer uses 11 convolutions to map 64 channels to the desired number of category categories. It has the following advantages; the network structure is shown in Figure 3.

1) U-Net is built on the network architecture of FCN. The authors modify and expand this network framework so that it can get very accurate segmentation results with very few training images.

2) Add the upper sampling phase and lot of feature channels, allowing more of the original image texture information in the high-resolution layers for dissemination

3) U-Net has no FC layer and uses valid to convolution all the way. This ensures that the result of segmentation is based on no missing context feature, so the size of the input and output images is not quite the same for very large input images, you can use overlap-strategy for seamless image output.

4) In order to predict the edge of the input image, it is possible to extrapolate the missing context information by mirroring the input image. In fact, it is also possible to input a large image, but this strategy is based on the case of insufficient GPU memory.
C. Dense U-Net

Our network diagram is shown in Figures 4 and 5.

FIGURE 3. The structure of Unet

FIGURE 4. The Dense U-Net frame diagram presented in this paper, which add the Dense block to the U-Net shrink and box path. The network input is x and the size is 320 × 320.

FIGURE 5. Four layered dense block with a growth rate of $g = 8$ and $g = 32$. Feature-maps learned in previous layers are concatenated with subsequent layers to learn the desired $g$ = 64 feature-maps.
Dense U-Net network is a common structure u-shaped network system based on semantic segmentation, which consists of two parts: encoding and decoding [21]. As shown in Figure 3, in the coding part, the convolution kernel pooling is used to extract features continuously, and the image is decomposed at multiple levels. Finally, through decoding process, up-sampling and convolution operation, the semantic segmentation at pixel level is carried out. Such a process enables the network to learn different global and local features of images at different scales [22]. Dense connections can increase the training speed. K represents the number of network feature graphs. Each layer of training is constantly updated, so the fusion layer in the dense fast network is the same. A link is made between the encoding and decoding so that the valid features of the learning can be up-sampled back to the same size as the input. Before each dense block of the up-sampling process, 1x1 convolution is used to simplify the connected 2f eigengraph to f/2.

In a dense block, the preceding convolutional layer is connected to all subsequent convolutional layers through a channel cascade [23]. The output of the $l^{th}$ layer has feature-map, and the input of the $F + k \cdot (l - 1)$ layer has feature-map, where is the number of feature-maps in the input, as shown in Figure 4. In order to improve computational efficiency, 1x1 convolution is applied before each 3x3 convolution to reduce the number of input feature graphs to 4. The final output of a dense block is the connection between the inputs and outputs of the $l$ layers.

The proposed CNN structure uses dense connections to decompose the size of convolution kernel between input and output [24]. In dense connection, the input image is convoluted by convolution kernel without size, and finally the output is merged. In this way, the features learned by the network are more abundant, and both local and global information can be taken into account. For CNN, features that can be learned through dense connection are more abundant [25].

In ReLU, small batch normalization is used. During training, it uses its mean and standard deviation in each batch and updates global statistical data with these values. Next is a layer that defines learning scale and bias.

**D. Network Parameters**

In the process of experimental training, we take binary cross entropy as a loss function. The output of the last layer is $O_i \in [0,1], y_i \in [0,1]$ is the real sample tag, and the loss function is defined as follows:

$$L_{bce} = \sum y_i \log o_i + (1 - y_i) \log (1 - o_i) \quad (1)$$

In the training process, Nesterov Adaptive Moment Estimation (Nadam) is used to train our proposed network, and other methods are available.

**IV. EXPERIMENT**

**A. Experimental Data**

CASIA-iris-interval-v4. 0 [26]: Dataset (casia-4i) dataset is a subset of the CASIA database. The data is collected in the indoor environment by the circular near-infrared LED (Light Emitting Diode) array near-infrared iris camera. Therefore, the image quality is good. The pixel size of the original data is 320×280. Because in the process of network design, the length and height of images must be the same, otherwise, network training will fail. For the need of network data, we change it to 320×320.

**B. Experimental Index**

The simulation platform of the experiment is PyCharm, using keras and TensorFlow port, the computer is configured as Inter (R) Core (TM) i7-8750H @2. 0GHz, 16GB memory, Nvidia GeForce GTX 1070 GPU (Graphics Processing Unit), using 64-bit Win10.

We train and test our network on casia-4i, and evaluate the network and segmentation using both objective and subjective methods. Our main goal is to extract the iris region. We plan to divide our segmentation results into four kinds of types, as shown in Figure 6. The correctly divided iris pixel was True Positive (TP), the incorrectly divided iris pixel was False Positive (FP), the non-iris pixel was wrongly divided into False Negative (FN), and the non-iris pixel was correctly divided into True Negative (TN).

![Schematic diagram of iris segmentation results](image)

**FIGURE 6. Schematic diagram of iris segmentation results**

We use nice1 and nice2 to evaluate the performance of our proposed network. In addition, $f_1$ is used to objectively assess the accuracy of our computed results. Many scholars in the field use nice1 to assess the error rate. The method is defined as follows:

$$nice1 = \frac{1}{N \times m \times n} \sum_{i=1}^{N} \sum_{j=\text{m,n}} G(i, j) \oplus O(i, j) \quad (2)$$

In the above equation, $N$ is the number of images, while $(m, n)$ is the spatial resolution of the image. $G(i, j)$ and $O(i, j)$ are the pixels of the ground truth image and the output image, respectively.
Nice2 is another index to assess computed results, which is mainly calculated by the average of false positive rate (FPR) and false negative rate (FNR). FPR and FNR are defined as follows:

\[
\begin{align*}
FPR &= \frac{FP}{FP + TN} \\
FNR &= \frac{FN}{FN + TP} \\
nice2 &= \frac{1}{2} (FPR + FNR)
\end{align*}
\] (3)

Accuracy is an important index to measure segmentation and classification, and its formula is as follows:

\[
\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}
\] (4)

C. Experimental Results

In order to evaluate the effectiveness of the proposed algorithm, we divide the experimental results into subjective and objective parts.

1) SUBJECTIVE EVALUATION

We randomly selected four groups of data from the test set for objection evaluation, as shown in Figures 7-10. The data of the four groups were ideal iris, upper eyelid occlusion, lower eyelid occlusion and upper and lower eyelid occlusion respectively.

As shown in Figure 7, it is the ideal iris segmentation result. It can be seen from the figure that the segmentation results in this paper are almost identical with the ground truth, and the contour segmentation of iris is complete. It can be seen from the segmentation gray scale that there is no over- and under-segmentation phenomenon in iris segmentation. The segmentation results can be applied to real security and other situations.

As shown in Figure 8, the upper eyelid covers the iris image. The position of the red box is the occlusion part of iris. The segmentation results were compared with the ground truth, and the proposed method showed that the binary image segmentation was of high accuracy and basically consistent with the ground truth. After observing the segmented gray image, the iris occlusion part is better segmented.

As shown in Figure 9, the lower eyelid covers the iris image. The position of the red frame is the lower eyelid occlusion. By comparing the results, the segmentation results were more accurate and not affected by the white area in the
pupil. After the observation of the gray scale after segmentation, iris features remain relatively complete.

As shown in Figure 10, the upper and lower eyelids block the iris image. The position of the red frame is upper and lower eyelid occlusion. By comparing the results, upper and lower occlusion can be distinguished clearly. It's not affected by the shaded area. Observation segmentation gray image, segmentation results are ideal, the basic iris region segmentation. It can be seen that our proposed algorithm is robust.

In order to observe the segmentation effect of the algorithm more intuitively, we randomly selected two groups of data for each type of non-ideal iris to conduct local amplification of upper and lower eyelids, so as to observe the segmentation effect more clearly. Figures 11 and 12 show the enlarged image of upper eyelid occlusion. Figures 13 and 14 show the enlarged image of lower eyelid occlusion, and Figures 15 and 16 show the enlarged image of upper and lower eyelid occlusion.

As shown in Figure 11, the first group of upper eyelid occlusion images. (a) Original image (b) The upper eyelid occlusion enlarges the image (c) Segmentation result (d) Segmentation results upper eyelid occlusion enlarges the image.

As shown in Figure 11, is the segmentation image of upper eyelid occlusion. Figure b is the enlarged image of the upper eyelid, Figure c is the segmented grayscale image, and Figure d is the enlarged image of the upper eyelid segmentation. As can be seen from Figure d, the upper eyelid was completely removed, and the iris part was retained completely. The segmentation marks were obvious, and the boundary between the upper eyelid and the iris was obvious.

As shown in Figure 12, similar to Figure 11, upper eyelid occlusion is used to segment the image. It can be observed from Figure d that our segmentation effect is obvious, the upper eyelid and iris are clearly distinguished.

As shown in Figure 13, the first group of lower eyelid occlusion images. (a) Original image (b) The lower eyelid occlusion enlarges the image (c) Segmentation result (d) Segmentation results lower eyelid occlusion enlarges the image.
As shown in Figure 13, it is the segmentation image of lower eyelid occlusion. Figure a is the original grayscale image, Figure b is the lower eyelid magnification image, Figure c is the segmentation grayscale image, Figure d is the lower eyelid segmentation image. From the magnification effect of Figure d, there are some burrs at the edges, which need to be smoothing in the later stage, but the iris area remains intact.

As shown in Figure 14, lower eyelid occlusion is also the segmentation image. Figure a is the original grayscale image, Figure b is the lower eyelid magnification image, Figure c is the segmentation grayscale image, Figure d is the lower eyelid segmentation method image. From the magnification effect of Figure d, the edges are smooth and the segmentation effect is significant.

FIGURE 15. The first group of upper and lower eyelid occlusion images. (a) Original image (b) The upper and lower eyelid occlusion enlarges the image (c) Segmentation result (d) Segmentation results upper and lower eyelid occlusion enlarges the image.

As shown in Figure 15, it is the segmentation effect of upper and lower eyelid occlusion at the same time. Figure a is the original image, Figure b is the enlarged image of upper and lower eyelid, Figure c is the grayscale segmentation image, and Figure d is the enlarged image of upper and lower eyelid segmentation. As can be seen from the segmentation results of upper and lower eyelids in Figure d, the segmentation of upper eyelid edge is relatively smooth and the effect is obvious, while the lower eyelid is a little over-divided, but the overall iris area remains intact.

FIGURE 16. The second group of upper and lower eyelid occlusion images. (a) Original image (b) The upper and lower eyelid occlusion enlarges the image (c) Segmentation result (d) Segmentation results upper and lower eyelid occlusion enlarges the image.

As shown in Figure 16, it is also the segmentation effect of upper and lower eyelid occlusion at the same time. Figure a is the original image, Figure b is the enlarged image of upper and lower eyelid, Figure c is the grayscale segmentation image, and Figure d is the enlarged image of upper and lower eyelid segmentation. As can be seen from the segmentation results of upper and lower eyelids in Figure d, there are some over-segmentation of upper and lower eyelids, but iris features can be completely separated and can be fully applied to real life. The reliability of the proposed algorithm is illustrated.

2) OBJECTIVE EVALUATION
In order to verify the effectiveness of the proposed algorithm, the Algorithm is compared with other algorithms in the literature, and the results are shown in Table 1.
<table>
<thead>
<tr>
<th>Method</th>
<th>nice1 (%)</th>
<th>nice2 (%)</th>
<th>f1 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional algorithm for iris segmentation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wahed[26]</td>
<td>6.08</td>
<td>8.42</td>
<td>89.49</td>
</tr>
<tr>
<td>Calt[27]</td>
<td>11.61</td>
<td>14.70</td>
<td>76.51</td>
</tr>
<tr>
<td>CNN for iris segmentation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FCEDNs-origional[28]</td>
<td>5.61</td>
<td>5.88</td>
<td>88.26</td>
</tr>
<tr>
<td>FCEDNs-basic[16]</td>
<td>4.48</td>
<td>4.38</td>
<td>90.72</td>
</tr>
<tr>
<td>FCEDNs-bayesian-basic[16]</td>
<td>3.91</td>
<td>4.07</td>
<td>91.92</td>
</tr>
<tr>
<td>Yang[17]</td>
<td>0.79</td>
<td>0.88</td>
<td>98.6</td>
</tr>
<tr>
<td>U-Net</td>
<td>0.61</td>
<td>1.27</td>
<td>97.23</td>
</tr>
<tr>
<td>proposed</td>
<td>0.31</td>
<td>1.22</td>
<td>97.07</td>
</tr>
</tbody>
</table>

As can be seen from Table 1, the proposed network segmentation error rate is low. First of all, compared with the traditional methods, this paper is much higher than the traditional methods, regardless of nice1, nice2 and f1 scores. Compared with Yang et al., our method scores 0.34% lower in nice2, and the f1 score is also 1.53% lower, but the nice1 score is increased by 0.48%.

In order to more intuitively see the advantages and disadvantages of our proposed algorithm and other advanced algorithms, as shown in Figures 17-19, it can be more clearly seen that our algorithm is superior to other algorithms by using the bar chart.

As can be intuitively seen from Figure 17, our nice1 index is significantly smaller than other literatures, indicating that our error segmentation rate is very low, and we can basically completely segment the iris.

As can be intuitively seen from Figure 18, our nice2 index is obviously smaller than that of most literatures, which is not much different from that of reference 17. However, combined with nice1, it can be seen that our comprehensive nice index is better than that of this literature.

As can be seen from Figure 19, our f1 index is slightly higher than that of some literatures, but combined with the other two indexes of nice1 and nice2, we can find that our algorithm is very effective.

V. CONCLUSION

With the development of mobile devices and social networks, activities related to identity authentication and information security are becoming more and more popular. Due to the inherent disadvantages such as being easy to be forgotten, lost, forged and stolen, the traditional identity authentication methods have been unable to meet people’s increasing demand for information security. Therefore, the biometric authentication technology based on fingerprint, human face and other biometric features has come into being. Compared with other biological features, Iris contains rich and unique texture, and has many advantages, such as uniqueness, stability, easy to collect, which makes iris play an important role in biometric authentication system. However, there are still a series of problems in current Iris segmentation and recognition Algorithms, for example, many existing iris recognition systems are designed on the basis of good iris images in the same database. The difficulty of segmentation and recognition of Heterogeneous Iris images is neglected. Based on this, this article from the Iris Own Physiological Structure Characteristics, has the important research significance and the application value. We propose to combine Dense and U-Net network, the final segmentation
accuracy rate is 98.36%, and the error segmentation rate is far lower than those advanced algorithms. Therefore, this algorithm can be applied to real life and can be applied to the diagnosis of Clinical Iris Diseases. In future, we will consider combining this algorithm with internet of things to realize the full automation from the data collection to the processing.

REFERENCE


Xiaoqiang Wu is an associate professor at the University of Inner Mongolia Nation. He is currently doing a doctorate at the Key Laboratory of Mechanism and Equipment Design of Ministry of Education of Tianjin University. His research interests are related to advanced manufacturing technology, computer algorithms and internet of things. He has published research papers at national and international journals, conference proceeding as well as chapters of books.

Long Zhao received M.S degree in Computer Science and Technology from Shandong Polytechnic University in 2009. He received his Ph.D. degree from Wuhan University in 2016. Recently, he is currently a lecturer in the School of Computer Science and Technology, Qilu University of Technology (Shandong Academy of Sciences). His research interests include image processing, machine learning, and knowledge discovery.