PI-DECODER: Decoder Structure Designed for Precise Iris Segmentation and Location

Dongliang Wu, Yingfeng Liu, Ruiye Zhou, Huihai Wu
Shanghai University of Electric Power

1. Introduction

In our lab (Machine Vision and Intelligent Analysis Laboratory of Shanghai University of Electric Power), there is a group of people led by Professor Shen Wenzhong who focus on the research of iris research including iris location, segmentation, recognition and so on. In the past studies, we have tried to segment the iris accurately with machine learning. However, the purpose of this competition, IR-ISL 2021, is to perform iris segmentation and eye location simultaneously with the database collected under non-cooperative conditions. It is a challenging task, and we are very interested in this competition.

2. PI-decoder

We consider the two tasks of iris segmentation and location as a single task, a multi-class segmentation task of the human eye region, segmenting the iris, pupil edge mask, and iris edge mask. Finally, the corresponding inner boundary and outer boundary can be extracted by simple post-processing according to the pupil and iris edge masks' binary segmentation mask.

At the beginning of the competition, we tried to transform a regular segmentation network, directly outputting three binary segmentation masks at the end of the network. However, the results of the experiment were disappointing. The training results were unstable, and the effect was not good. This got us thinking.

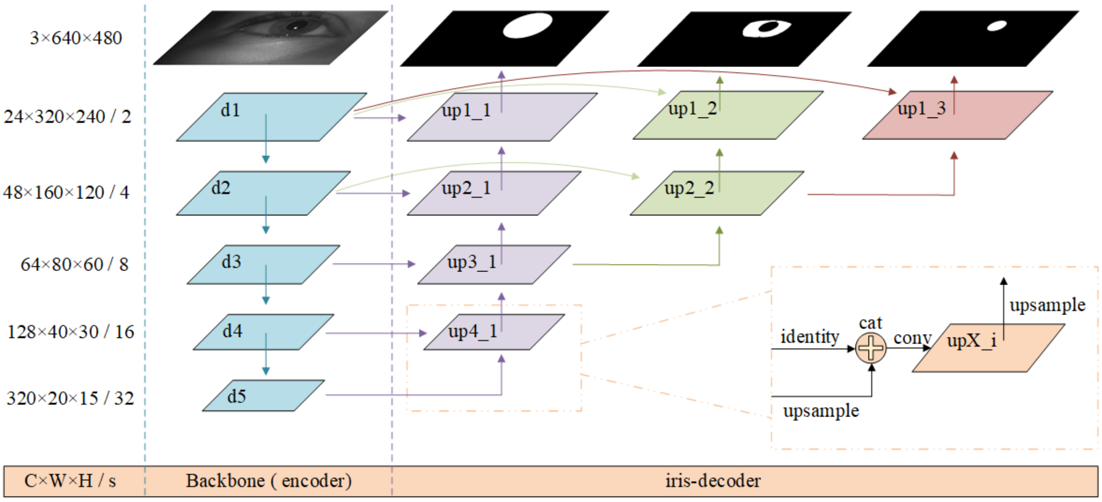

Figure 1 - The proposed PI-decoder structure, where di(i = 1,2,…,5) denote the feature maps of the backbone network and upX_i(X = 1,2,…,4; i = 1,2,3) are the feature maps in PI-decoder. C x W x H, at the bottom of the figure, denotes the channels, width and height of the feature maps, respectively. '/ s' is the down-sampling ratio of the feature maps to the input images. We choose the CASIA-Iris-Africa as example, the resolution of input image is 640 x 480.

Table 1 – The evaluation results of Iris Segmentation and Eye location, including the results of cross database evaluation. The metric mIOU on the right side of the table is the average of the three masks mIOU.

Method	Testing Dataset	Training Dataset	E1 (%)	E2 (%)	Hausdorff		mIOU (%)
					inner boundary	outer boundary	
PI-Decoder	Africa	Africa	**0.38**	**0.19**	**4.05**	**11.38**	**92.49**
		Asia	0.66	0.33	/	/	85.98
		Mobile	0.62	0.31	/	/	86.89
	Asia-Distance	Africa	0.6	0.3	/	/	85.87
		Asia	**0.3**	**0.15**	**3.12**	**6.01**	**93.82**
		Mobile	0.43	0.21	/	/	88.63
	Asia-Occlusion	Africa	0.75	0.38	/	/	80.92
		Asia	**0.42**	**0.21**	**2.95**	**6.54**	**90.77**
		Mobile	0.66	0.33	/	/	80.36
	Asia-Off-Angle	Africa	0.58	0.29	/	/	88.08
		Asia	**0.38**	**0.19**	**2.38**	**6.23**	**92.63**
		Mobile	0.54	0.27	/	/	85.18
	Mobile	Africa	0.56	0.28	/	/	87.62
		Asia	0.51	0.26	/	/	85.8
		Mobile	**0.4**	**0.2**	**1.69**	**6.91**	**91.87**

We always believe that in the human eye region, the relationship between the iris, pupil, and iris outer circle is family relationship. The multi segmentation task should not be a burden on the network, but mutually promote each other. For this, we redesigned the decoder in the encoder-decoder structure, to perform well for the tasks of iris segmentation and iris location. We called such a decoder structure PI-decoder, Precise-Iris-decoder.

As shown in figure 1, the feature map corresponding to the iris region was up-sampled from the feature map of the outer iris circle, and the feature map corresponding to the pupil region was up-sampled from the feature map of the iris region. As mentioned above, for the 'family' of iris, pupil, and iris outer circle, the iris is located in the outer iris circle, and the exposed boundary of the iris should completely coincide with the iris outer circle. The pupil is in the middle of the iris and should not cross the iris at all. Such a simple and effective idea helps us to design an PI-decoder. Our experiment proves that the PI-decoder is more stable and has better performance. Of course, this is the first attempt, and we will continue to improve it in the future.

3. Experiments

The proposed model is implemented in Pytorch, using Adam as the network optimizer and modified EfficientNetV2 as the backbone. The learning rate was 0.001, and used the cosine annealing scheduler with an adjustment period of 32 epochs. Furthermore, we manually split the dataset into two parts for training and validation according to the setting of 6 : 1. Experiments data are shown in Table 1.

4. Visualized Results

We put the visualized final results in the submitted folder:

NIR-ISL2021030902_1/visualized_result

Through these pictures, we can see our iris segmentation and location results intuitively.

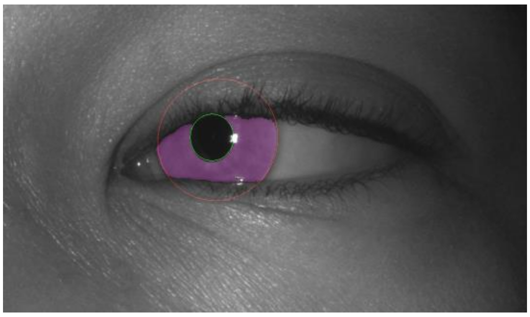

Figure 2 – visualized result in African-database.

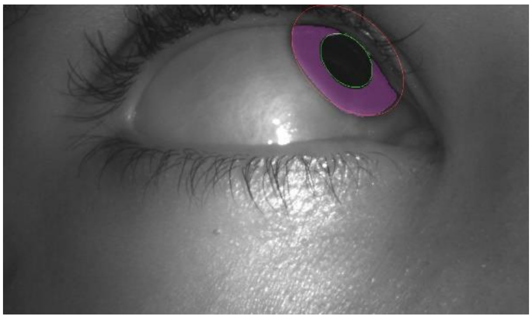

Figure 3 – off angle result.

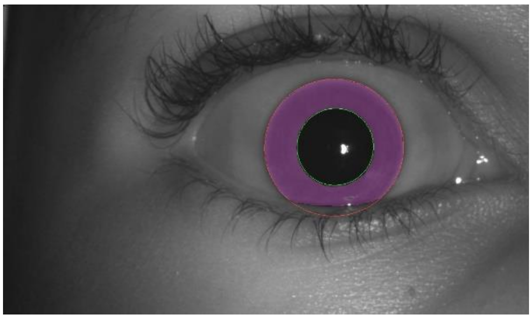

Figure 4 – Excellent iris segmentation and location result.