

NIR Iris Challenge Evaluation in Non-cooperative Environments: Segmentation and Localization

Caiyong Wang^{1,*}, Yunlong Wang², Kunbo Zhang², Jawad Muhammad², Tianhao Lu², Qi Zhang³, Qichuan Tian¹, Zhao Feng He⁴, Zhenan Sun^{2,†}, Yiwen Zhang⁵, Tianbao Liu⁵, Wei Yang⁵, Dongliang Wu⁶, Yingfeng Liu⁶, Ruiye Zhou⁶, Huihai Wu⁶, Hao Zhang⁷, Junbao Wang⁷, Jiayi Wang⁷, Wantong Xiong⁷, Xueyu Shi⁸, Shao Zeng⁸, Peihua Li⁸, Haodong Sun⁹, Jing Wang⁹, Jiale Zhang⁹, Qi Wang⁹, Huijie Wu¹⁰, Xinhui Zhang¹⁰, Haiqing Li¹⁰, Yu Chen¹¹, Liang Chen¹¹, Menghan Zhang¹¹, Ye Sun¹², Zhiyong Zhou¹², Fadi Boutros¹³, Naser Damei¹³, Arjan Kuijper¹³, Juan Tapia¹⁴, Andres Valenzuela¹⁴, Christoph Busch¹⁴, Gourav Gupta¹⁵, Kiran Raja¹⁵, Xi Wu¹⁶, Xiaojie Li¹⁶, Jingfu Yang¹⁶, Hongyan Jing¹⁶, Xin Wang¹⁷, Bin Kong¹⁷, Youbing Yin¹⁷, Qi Song¹⁷, Siwei Lyu¹⁸, Shu Hu¹⁸, Leon Premk¹⁹, Matej Vitek¹⁹, Vitomir Štruc¹⁹, Peter Peer¹⁹, Jalil Nourmohammadi Khiaraki²⁰, Farhang Jaryani²¹, Samaneh Salehi Nasab²², Seyed Naeim Moafinejad²³, Yasin Amini²⁴, Morteza Noshad²⁵

¹School of Electrical and Information Engineering, Beijing University of Civil Engineering and Architecture, China, ²CRIPAC, NLPR, CASIA, China, ³People's Public Security University of China, China, ⁴Beijing University of Posts and Telecommunications, China, ⁵School of Biomedical Engineering, Southern Medical University, China, ⁶Shanghai University of Electric Power, China, ⁷College of Science, Northeastern University, China, ⁸Dalian University of Technology, China, ⁹College of Sciences, Northeastern University, China, ¹⁰IriStar Technology Co., Ltd, China, ¹¹Xi'an Quanxiu Technology Co., Ltd, China, ¹²JiLin University, China, ¹³Fraunhofer Institute for Computer Graphics Research IGD, Germany; Mathematical and Applied Visual Computing, TU Darmstadt, Germany, ¹⁴Darmstadt University of Applied Sciences (Hochschule Darmstadt) and TOC Biometrics, ¹⁵Norwegian University of Science and Technology, Norway, ¹⁶Chengdu University of Information Technology, China, ¹⁷Keya Medical, Seattle, WA, USA, ¹⁸University at Buffalo, USA, ¹⁹University of Ljubljana, Slovenia, ²⁰Warsaw University of Technology, Poland, ²¹Arak University, Iran, ²²Lorestan University, Iran, ²³Shahid Beheshti University, Iran, ²⁴University of Kharazmi, Iran, ²⁵Stanford University, USA

*wangcayong@bucea.edu.cn, †znsun@nlpr.ia.ac.cn (corresponding author)

Abstract

For iris recognition in non-cooperative environments, iris segmentation has been regarded as the first most important challenge still open to the biometric community, affecting all downstream tasks from normalization to recognition. In recent years, deep learning technologies have gained significant popularity among various computer vision tasks and have also affected iris biometrics, especially iris segmentation. To investigate recent developments and attract more interest of researchers in the iris segmentation method, we organized the 2021 NIR Iris Challenge Evaluation in Non-cooperative Environments: Segmentation and Localization (NIR-ISL 2021) at the 2021 International Joint Conference on Biometrics (IJCB 2021). The challenge was used as a public platform to assess the performance of iris segmentation and localization methods on Asian and African NIR iris images captured in non-cooperative environments. The three best-performing entries achieved solid and satisfactory iris segmentation and localization results in most cases, and their code and models have been made publicly available for reproducibility research.

1. Introduction

Traditional iris recognition usually imposes many constraints to the user cooperation and imaging conditions, which seriously limits the application range of iris recognition. To solve this dilemma, there has been much recent work [5, 2, 12, 6] on the non-cooperative or less-constrained iris recognition (either at-a-distance, on-the-

move, with minor user cooperation, within dynamic imaging environments and using mobile devices). Under these circumstances, captured iris images inevitably suffer from all kinds of noise, such as occlusions due to eyelids or eyelashes, specular reflections, off-angle, or blur. To make full use of these noisy iris images, efficient and robust iris segmentation has been regarded as the first most important challenge still open to the biometric community, affecting all downstream tasks from normalization to recognition.

In 2007, Noisy Iris Challenge Evaluation - Part I (NICE.I) was held to benchmark the iris segmentation methods on the Noisy Visible Wavelength Iris Image Database (UBIRIS.v2) [5]. In 2013, Mobile Iris CHallenge Evaluation part I (MICHE I) was also held to evaluate the iris segmentation methods developed for visible iris images captured with mobile devices [2]. The two iris segmentation benchmarking competitions mainly focused on the VIS iris images from Caucasian people. In terms of evaluation metrics, only segmentation accuracy of noise-free iris masks was independently evaluated in the competition, but the localization accuracy of inner and outer boundaries of the iris was ignored or relied on an indirect evaluation based on the iris recognition performance, which was non-intuitive, time-consuming, and complex (relying on the downstream iris encoding/matching and larger iris datasets). Besides, most submitted methods in NICE.I and MICHE I were developed based on traditional image processing and pattern recognition methods, rather than deep learning technologies emerged in recent years.

To reflect latest developments of iris segmentation and offer new insights, the **2021 NIR Iris Challenge Evalu-**

Dataset	Resolution	Variability [†]	No. of training	No. of testing
CASIA-Iris-Asia	640 × 480	LR,OL,OA	496	899
CASIA-Iris-M1	400 × 400	MD,UI,VS,OL	1,500	1,500
CASIA-Iris-Africa	1088 × 640	UI,OL,OA,DK	370	370

[†] LR - long-range device, OL - occlusion, OA - off-angle, MD - mobile device, UI - uneven illumination, VS - various iris size, DK - dark coloured iris and skin.

Table 1. Overview of the datasets used for NIR-ISL 2021. Each of these is a subset of the corresponding original database.

tion in Non-cooperative Environments: Segmentation and Localization (**NIR-ISL 2021**) was organised as part of the 2021 International Joint Conference on Biometrics (IJCB 2021). The challenge aimed at benchmarking the methods of iris segmentation and localization for NIR iris images from Asian and African people captured in non-cooperative environments. Here we followed a recent line of our research on iris segmentation, *i.e.*, explicitly splitting the general iris segmentation task into the segmentation of noise-free iris mask and the localization of inner and outer boundaries of the iris, narrowly referred to as *iris segmentation* and *iris localization* [10]. The main reason for this is to consider that many recent deep learning based iris segmentation methods [4, 3, 9] were only designed to segment the noise-free mask but ignore the localization of iris boundaries. Such an incomplete solution makes it hard to effectively deploy in the conventional iris recognition pipeline. We employed diverse Asian and African iris datasets, which featured challenging iris images with manually labeled ground truths in non-cooperative environments to participants for developing and evaluating iris segmentation methods.

In the remainder of this paper, we first describe the training and testing datasets, the experimental protocol, and the evaluation metrics. Then, we summarize the participants and submitted algorithms of the challenge. Finally, we present, discuss the submitted prediction results quantitatively and qualitatively, and conclude the challenge.

2. Datasets

Three distinct NIR iris datasets built by CASIA* were used for NIR-ISL 2021: *i*) CASIA-Iris-Asia, *ii*) CASIA-Iris-M1 and *iii*) CASIA-Iris-Africa. All training iris images from three datasets were made available to the participants together with the ground truth segmentation masks and localization boundaries at the start of challenge. Besides, all testing iris images (without ground truths) were also released to the participants along with training datasets. Main characteristics of the competition datasets are summarized in Table 1. A more detailed description of the datasets is also provided as following.

CASIA-Iris-Asia contains various NIR iris images from

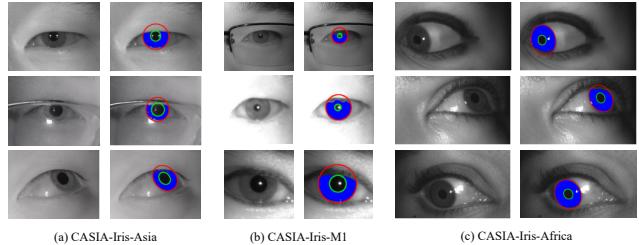


Figure 1. Sample images and corresponding ground truths (including inner iris boundary (green), outer iris boundary (red), and segmentation mask (blue)) for the datasets in the challenge.

Asian subjects in non-cooperative environments. Specifically, the dataset is a union of the previously collected long-range CASIA-Iris-Distance and the newly collected CASIA-Iris-Complex in complex environments, where CASIA-Iris-Distance was captured by a CASIA long-range iris camera, while CASIA-Iris-Complex was acquired in an indoor environment with a modified Canon EOS 1300D camera. Images from CASIA-Iris-Complex suffer from various noise, especially occlusions due to eyelids/eyelashes/glasses and off-angle. 500 occluded iris images were collected to form a CASIA-Iris-Complex-Occlusion subset, meanwhile 500 off-angle iris images were also collected to form a CASIA-Iris-Complex-Off-angle subset. Since the original images from CASIA-Iris-Complex captured the partial face, left and right eye regions of interest were manually extracted with resolution of 640 × 480 pixels for each subset image. The CASIA-Iris-Distance dataset used in the challenge consists of 400 iris images with resolution of 640 × 480 pixels from the first 40 subjects [4].

For training, 100 iris images were randomly selected from CASIA-Iris-Complex-Occlusion and CASIA-Iris-Complex-Off-angle subsets, respectively. In addition, in consistent with [4], the first 300 images from the first 30 subjects were served as the training data of CASIA-Iris-Distance. Hence, 500 iris images from three subsets were put together to develop the iris segmentation and localization model while the rest were employed for performance testing. Ground truth segmentation masks and inner and outer boundaries of the iris were annotated manually for all images in the CASIA-Iris-Asia dataset except for those extremely hard images. In summary, after removing those hard samples, the training set in the CASIA-Iris-Distance contains 296 iris images with ground truths and the testing set comprises 99 iris images with ground truths. For CASIA-Iris-Complex-Occlusion and CASIA-Iris-Complex-Off-angle, there were no hard samples to be removed, hence their training and testing sets are composed of 100 and 400 iris images with ground truths, respectively. Some sample images and corresponding ground truths are shown in Figure 1(a).

CASIA-Iris-M1 is a large-scale NIR mobile iris dataset

*<http://www.cripacsir.cn/dataset/>

comprising 11,000 images from 630 Asian subjects [12]. The dataset consists of three subsets: CASIA-Iris-M1-S1 (1,400 images, 70 subjects), CASIA-Iris-M1-S2 (6,000 images, 200 subjects), and CASIA-Iris-M1-S3 (3,600 images, 360 subjects). Images from the former two subsets were acquired with two versions of NIR iris imaging modules attached to a mobile phone through micro USB, while images from the last subset were captured using a mobile phone equipped with NIR iris-scanning technology. The variability of CASIA-Iris-M1 include uneven illumination, various iris size, occlusion of wearing glasses, *etc.* Similarly to CASIA-Iris-Complex, originally captured partial face images in the CASIA-Iris-M1 dataset were manually cropped to eye regions of interest and the cropped size was set to a resolution of 400×400 pixels. The challenge adopted cropped iris images and corresponding ground truths from [11] for training and evaluating the proposed model. Overall, the training set contains 1,500 images randomly and equally selected from three subsets while the testing set contains another 1,500 disjoint images selected in a similar manner. Some sample images and corresponding ground truths are shown in Figure 1(b).

CASIA-Iris-Africa was captured in Nigeria and the first large-scale black African iris dataset to the best of our knowledge. Iris images from the dataset have a resolution of 1088×640 pixels and were obtained under NIR illumination and in less-constrained environments using a hand held IKUSBE30. Various noise factors such as occlusions, off-angle, illumination variations, dark coloured iris and skin were introduced in the dataset and thus makes the accurate and robust iris segmentation and localization hard. In the challenge, 370 iris images of different types of noise were selected for training and another disjoint 370 iris images were selected for testing. Ground truth segmentation masks and inner and outer boundaries of the iris were manually annotated for all images in the dataset. Some sample images with different gaze angles and corresponding ground truths are shown in Figure 1(c).

3. Experimental protocol

In consideration of different characteristics and application scenarios of the datasets in the challenge, participants were asked to train each algorithm on 3 different versions of training data: *i*) CASIA-Iris-Asia, *ii*) CASIA-Iris-M1, and *iii*) CASIA-Iris-Africa. As a result of this setup, participants should obtain 3 distinct models after the training stage of each algorithm, with each model for the specific target. Participants could contribute more than a single submission, but each submitted algorithm should be trained according to the above requirements.

In the testing stage, participants were asked to perform a regular within-dataset evaluation for trained models. As a result, participants should have 5 distinct versions of testing

results for each submitted algorithm, where the results of CASIA-Iris-Asia were split into three parts corresponding to CASIA-Iris-Distance, CASIA-Iris-Complex-Occlusion, and CASIA-Iris-Complex-Off-angle, respectively. The testing results to be submitted had to contain binary iris segmentation masks and binary inner and outer boundaries of the iris corresponding to the test images, where non-zero valued pixels representing the foreground (valid iris texture or boundary) region and zero valued pixels representing everything else. Note that the binary inner and outer boundaries of the iris should be single-pixel wide.

In addition to the testing results, participants were also required to submit a technical document of 1 or 2 pages containing a general description of their proposed algorithms. To make the submitted testing results reproducible, participants were further encouraged to release their executable programs or models with relevant documentation.

4. Evaluation metrics

The performance of each participating algorithm was evaluated by the organisers of NIR-ISL 2021 from two aspects: 1) iris segmentation and 2) iris localization. Specifically, the following evaluation metrics were computed based on the submitted binary segmentation or localization predictions and corresponding ground truths for each algorithm:

Iris segmentation. The challenge used the two metrics[†] proposed by the NICE.I competition [5] to evaluate the performance of iris segmentation.

The first metric $E1$ denotes the average proportion of corresponding disagreeing pixels over all the images, which could be calculated through a pixel-wise XOR operator \otimes between the predicted iris mask and ground truth iris mask. Mathematically, it is formulated by the following equation:

$$E1 = \frac{1}{n \times h \times w} \sum_i \sum_j M(i, j) \otimes G(i, j) \quad (1)$$

where i and j are the coordinates of pixels in the predicted iris mask M and ground truth iris mask G , h and w represent the height and width of the testing image, respectively, and n is the amount of testing images.

The second metric $E2$ is used to compensate the disproportion between the priori probabilities of "iris" and "non-iris" pixels in the images. Specifically, it is defined as the average of the false positives (fp) and false negatives (fn) rates:

$$E2 = \frac{1}{2 \times n} \sum_i (fp + fn) \quad (2)$$

where fp denotes the proportion of background pixels incorrectly retrieved as iris pixels, fn denotes the proportion

[†]<http://nice1.di.ubi.pt/evaluation.htm>

of iris pixels incorrectly retrieved as background pixels, and n is the amount of testing images.

Values of $E1$ and $E2$ are bounded in $[0, 1]$, where the lower and higher values represent better and worse iris segmentation results, respectively.

Iris localization. Inspired by the GlaS Challenge Contest [7], the challenge adopted the Dice index and the Hausdorff distance[‡] to evaluate the performance of iris localization.

Given \hat{B} , a set of pixels within a predicted inner or outer boundary of the iris, and \hat{G} , a set of pixels within a corresponding ground truth iris boundary, the Dice index is defined as follows:

$$\text{Dice}(\hat{G}, \hat{B}) = \frac{2|\hat{G} \cap \hat{B}|}{|\hat{G}| + |\hat{B}|} \quad (3)$$

where $|\cdot|$ denotes set cardinality, and \hat{B} and \hat{G} essentially represent the results of the original predicted iris boundary B as well as its ground truth boundary G being filled internally, respectively. To borrow a phrase from the iris segmentation mask, the internally filled iris boundary is also simply referred to as the iris boundary mask.

The Dice index ranges over the interval $[0, 1]$, where the higher the value, the higher the overlapping ratio between the predicted iris boundary mask and its ground truth mask, suggesting the better iris localization accuracy. The overall Dice index for the inner or outer boundary of the iris was given by the average of corresponding Dice indexes on all testing images. The challenge finally took the mean value of overall Dice indexes of the inner and outer of the iris, termed as $mDice$, for ranking.

Given B , the predicted inner or outer boundary of the iris, and G , the corresponding ground truth boundary, the Hausdorff distance is employed to measure their shape similarity as:

$$H(G, B) = \max\{\sup_{x \in G} \inf_{y \in B} d(x, y), \sup_{y \in B} \inf_{x \in G} d(x, y)\} \quad (4)$$

where $d(x, y)$ denotes the distance between pixels $x \in G$ and $y \in B$. Since the Hausdorff distance is with regard to the image resolution, it is further normalized by image width to eliminate the impact of different image size.

It is clear that the smaller the normalized Hausdorff distance is, the better the predicted iris boundary approaches its ground truth boundary, suggesting the higher localization accuracy. Similarly to the Dice index, the challenge calculated the mean normalized Hausdorff distance for the inner or outer boundary of the iris. At last, the average value of two mean normalized Hausdorff distances was used for ranking, termed as $mHdis$.

[‡]https://warwick.ac.uk/fac/cross_fac/tia/data/glascontest/evaluation

No.	Team ID	Team Name	Affiliation
1	NIR-ISL2021022201	Fraunhofer IGD	Fraunhofer Institute for Computer Graphics Research IGD, Darmstadt, Germany; Mathematical and Applied Visual Computing, TU Darmstadt, Darmstadt, Germany
2	NIR-ISL2021023301	belaTOC	Darmstadt University of Applied Sciences (Hochschule Darmstadt) and TOC Biometrics
3	NIR-ISL2021022601	University of Ljubljana	University of Ljubljana
4	NIR-ISL2021022801	Cosmic Vision	Norwegian University of Science and Technology (NTNU), Norway
5	NIR-ISL2021030101	DreamMaker	Jilin University
6	NIR-ISL2021030401	Lan Yang Sprint Team	School of Biomedical Engineering, Southern Medical University
7	NIR-ISL2021030701	DLUT_VLG	Dalian University of Technology
8	NIR-ISL2021030801	instar	IRISTAR TECHNOLOGY LTD.
9	NIR-ISL2021030902	SUEP-Pixsur	Shanghai University of Electric Power
10	NIR-ISL2021031301	KartalOI	Warsaw University of Technology, Poland; Arak University, Iran; Lorestan University, Iran; Shahrood Beheshti University, Iran; University of Kerman, Iran; Shandev University, USA
11	NIR-ISL2021031401	insight	College of Sciences, Northeastern University, Shenyang, China
12	NIR-ISL2021031601	IRIS team	Xi'an Quanxin Technology Co. Ltd.
13	NIR-ISL2021031901	CKB	Chengdu University of Information Technology, Chengdu, China; Keya Medical, Seattle, WA, USA; University at Buffalo, Buffalo, NY, USA
14	NIR-ISL2021041402	EyeCool	College of Science, Northeastern University

Table 2. Summary of teams who submitted their competition results.

Ranking scheme. For the overall results, each submitted entry was assigned one ranking score per evaluation metric and set of testing data, using a standard competition ranking[§]. Since there were 4 evaluation metrics ($E1$ and $E2$ for iris segmentation, $mDice$ and $mHdis$ for iris localization) and 5 sets of testing data (CASIA-Iris-Distance, CASIA-Iris-Complex-Occlusion, CASIA-Iris-Complex-Off-angle, CASIA-Iris-M1, and CASIA-Iris-Africa), the total number of ranking scores was 20. The sum of these 20 ranking scores was then used for the final ranking. To be specific, a smaller sum stands for a better overall performance, therefore the entry with smallest sum came first in the final ranking.

5. Summary of submitted algorithms

We received 30 team registrations for NIR-ISL 2021, who were affiliated with different academic institutes or industrial organizations in different countries. Only 14 teams took part in the final round and submitted a total of 27 valid entries (models) for both iris segmentation and iris localization. Table 2 gives a summary of successfully participating teams, while a brief description of the submitted algorithms is presented below ¶.

Fraunhofer IGD. The proposed model is termed as *Eye-MS-3* (NIR-ISL2021022201), which extends over Multi-scale CNN-based segmentation solutions (Eye-MS) [1] originally developed in OpenEDS Semantic Segmentation Challenge. The detailed extension includes two parts: 1) having an additional input to the first module of Eye-MS from a separately trained encoder network, and 2) adaptation to produce three outputs, the iris segmentation mask, the inner iris map, and the outer iris map as defined in [11]. The model takes the single-channel gray iris image normalized to a range of 0-1 as input, and resizes the images to 256×256 . Heavy data augmentation and a Intersection over Union (IoU) loss are employed to train the model. In the testing stage, the predicted segmentation mask is round-

[§]<https://en.wikipedia.org/wiki/Ranking>

[¶]The KartalOI team is not reported because of issues with the submitted results.

ed to the nearest integer values to represent the discrete labels, and the produced inner iris map and outer iris map are post-processed to produce inner iris boundary and outer iris boundary. The post-processing pipeline includes finding the largest contour in the binary map, then fitting an ellipse on these contour points. The ellipse, with the line width of one pixel, is found for each of both maps and is the final produced iris inner boundary and iris outer boundary.

hda/TOC. The team first proposes two methods for semantic segmentation of the ocular regions. The first method *hda-toC1* (NIR-ISL2021022301_1) is based on their previously proposed DenseNet10 [8] with the output of three classes: background, pupil and iris. The DenseNet10 model is designed to be efficient, consisting of a feature extractor and two paths (Down-sampling and one Up-sampling), and the amount of parameters is only 674,082. The second method *hda-toC2* (NIR-ISL2021022301_2) is based on MobilenetV2 and UNet (MobileUNet) with the output of two classes: background and iris. It has 8,872,002 parameters. The two methods employ an adaptive histogram equalization to pre-process iris images to improve the quality of the images before feeding them to the segmentation model. During the training, a series of aggressive data augmentation operations based on 'imgaug library'^{||} are applied to the training datasets, generating a large number of challenging images to improve the quality of model training. After a good segmentation of the iris and pupil, a post-processing method based on 'Mass Center' is used to estimate the boundaries of the iris and the pupil. The method searches the vertical and horizontal boundaries (edges) of the pixels, then calculates the radii of the pupil and the iris using the searched information.

University of Ljubljana. The team submission (NIR-ISL2021022601) is designed around a U-Net segmentation model, where the output layer is replaced by three parallel layers that allow to train the model in a multi-task manner. Specifically, the model is designed to produce three distinct segmentation outputs: (i) the segmentation mask corresponding to the iris region, (ii) the map of the outer iris boundary, and (iii) the map of the inner iris boundary. Learning to generate all three outputs in multi-task fashion allows the training procedure to exploit correlations and similarities between the different tasks and produce more reliable segmentation results. Since the inner and outer boundaries of the iris predicted by the model are often not contiguous, a least-squares fitting procedure is used to fit ellipses to the generated boundary maps. This procedure then generates the final predictions for the iris localization solution that is submitted together with the binary mask of the iris region for scoring. The model is trained with a cross entropy loss applied to each of the three outputs. Data augmentation is also used to avoid over-fitting and ensure com-

petitive performance.

Cosmic Vision. A modified Mask R-CNN (NIR-ISL2021022801) with ResNet-101 model as a backbone is proposed for the challenge. Specifically, the iris images are firstly pre-processed to ameliorate the quality of images using de-noising filtering and sharpening techniques. Then the model employs Feature Pyramid Network (FPN) to generate bounding boxes and segmentation masks for each instance of an iris region. It results in the segmented masks with three regions like hollow portion of Iris (iris segmentation mask), pupil region (inner boundary mask of the iris), and solid region of Iris (outer boundary mask of the iris). At last, the boundary detection algorithms such as Hough-Circle and Blob analysis have been applied to these outputs for getting the inner and outer boundaries of the iris.

DreamMaker. The team adopts the architecture of U-net (NIR-ISL2021030101) with the VGG model pre-trained on the ImageNet dataset as the backbone for the challenge. A carefully designed intermediate transition module and a new attention mechanism are proposed to improve the accuracy of network outputs. The model finally generates the iris segmentation mask, the inner boundary mask of the iris, and the outer boundary mask of the iris, then followed by a post-processing method applied to the inner and outer boundary masks of the iris for obtaining the iris boundary predictions to be submitted. The model is trained using heavy data augmentation operations to enrich the training data.

Lao Yang Sprint Team.** The classic U-Net architecture with ResNet-34 as the backbone is proposed as the network framework (NIR-ISL2021030401). To solve the challenge, besides the iris segmentation, the team also regards the iris localization as a segmentation task to segment the mask inside the iris boundary. Since the multitasking model for iris segmentation and localization could not achieve the desired performance, the team finally resorts to two independent models based on U-Net to perform two tasks, respectively. The only difference between the segmentation network and the localization network is the last head layer, where the segmentation head contains a convolution layer for outputting iris mask, and the localization head contains two parallel convolution layers for outputting the inner boundary mask and the outer boundary mask, respectively. The model is trained using a combination of the dice loss and binary cross-entropy loss. In addition, a heavy data augmentation is also required in the training.

At the core of this submission is a novel transfer learning based training procedure to improve the generalization ability of the model. All the data from CASIA-Iris-Asia, CASIA-Iris-M1, and CASIA-Iris-Africa are first combined together to train an iris pre-trained model separately for iris segmentation and localization on the basis of ImageNet pre-

^{||}The code and models are available at <https://github.com/whisney/NIR-ISL2021-Transfer-learning>.

^{**}<https://github.com/aleju/imgaug>

trained model. Then the iris pre-trained model is fine-tuned on each subdataset with a small learning rate and a cross validation setting, generating multiple prediction models. In the testing, Test-Time Augmentation (or TTA for short) based on model ensemble is applied to testing images to make a better prediction. The inner and outer boundaries of the iris to be submitted are extracted from the predicted iris boundary masks by a mathematical morphology-based edge detection method and further refined by ellipse fitting.

DLUT_VLG. The Mask R-CNN is exploited as the base framework for the challenge. For the iris localization, the circle is used to fit the inner and outer boundaries of the iris in most cases, hence the detection part of the Mask R-CNN is adapted to locate the circumscribing square of the inner and outer circles of the iris, and thus the parameters of iris circles are obtained. In some non-ideal cases, especially in Off-angle iris images, ellipse fitting is applied to the masks predicted by the segmentation part of the Mask R-CNN to obtain the iris boundaries. For the iris segmentation, the segmentation part of the Mask R-CNN allows the model to simultaneously predict the iris, the pupil and the background pixels. The model is trained using the datasets provided by the challenge and additional (external) datasets with ground truths. The team finally trained 2 different models with 2 different backbones, *i.e.*, ResNet50 and ResNeSt, then merged the result of the 2 models to achieve better predictions. A threshold based post-processing operation is used to remove the reflections in the segmentation mask. The team finally submitted 7 entries (from NIR-ISL2021030701_1 to NIR-ISL2021030701_7) for evaluation, where the iris localization part kept the same and the iris segmentation part had some minor differences that were not explained by the team.

iristar. The team builds a high-accuracy network (NIR-ISL2021030801) based on the U-Net structure for the challenge. The output of the network is multi-branch, including iris mask, iris inner boundary and iris outer boundary. In order to improve the accuracy of the network, the team modifies the network structure and loss function. Some special modules such as ResNet and Squeeze-and-Excitation (SE) module based attention-net are embedded in the network. Besides, the network is trained with a weighted fusion of two loss functions, including the dice loss for the iris segmentation branch, and the focal loss for the inner and outer boundary segmentation branch. In order to increase the richness of the training data, different methods are used for data augmentation. After the inference of the network, some necessary post-processing operations are employed to refine the predictions, *e.g.*, noise removal, erosion, dilation.

SUEP-Pixsur^{††}. The team proposes a model called PI-decoder (NIR-ISL2021030902) to perform iris segmen-

tion and localization simultaneously. The team considers the two tasks of iris segmentation and localization as a single task, *i.e.*, a multi-class segmentation task of the human eye region, including the iris mask, the inner boundary mask of the iris, and the outer boundary mask of the iris. Considering that there is a close spatial connection and prior knowledge between the iris region, the pupil region (*i.e.*, the inner boundary mask of the iris), and the outer boundary of the iris, *i.e.*, the iris region is located in the outer boundary of the iris, while the pupil region is located in the middle of the iris region and should not cross the iris region, a novel encoder-decoder structure (*i.e.*, PI-decoder) is carefully designed to explicitly exploit the "family" relationship to mutually promote the iris segmentation and the iris localization.

Specifically, the model first uses a modified EfficientNetV2 as the encoder backbone, then followed by three parallel decoder branches of decreasing layers for predicting the outer boundary mask of the iris, the iris mask, and the inner boundary mask of the iris, respectively. Skip connection is used to connect the encoder and decoder layers of the same level for refining the middle and the final decoding predictions. Furthermore, the decoder of the iris mask also connects the middle feature maps of decoding the outer boundary mask of the iris, while the decoder of the inner boundary mask of the iris connects the middle feature maps of decoding the iris mask. Such a simple method is expected to exploit the "family" relationship in the human eye region to achieve a more stable and high-performance model. Finally, the corresponding inner and outer boundaries of the iris are extracted by simple post-processing operations on the predicted iris boundary masks.

insight. The iris segmentation task employs an improved U-Net model, where the Efficient-Net network acts as the encoder backbone. In order to reduce the model complexity, the decoder does not contain complex modules, only embedding some attention modules, *e.g.*, spatial attention module, channel attention module. Besides, the spatial attention module is also applied to skip connection layers of the encoder and the decoder. The iris localization task is regarded as a segmentation problem, where the inner and outer boundary masks of the iris are segmented. In this way, the iris segmentation and localization tasks are combined together and solved by a unified U-Net network. The final iris boundaries to be submitted are obtained through some post-processing operations on the predicted iris boundary masks. The team finally submitted 2 entries (*i.e.*, NIR-ISL2021031401_1, NIR-ISL2021031401_2) for evaluation, where their model structures are the same, but the training details have some minor differences.

IRIS team. The team proposes a lightweight iris segmentation and localization method. The training dataset is enhanced by data augmentation and used to train the mod-

^{††}The code and models are available at <https://github.com/swezin/PI-DECODER>.

el. The testing image is then sent to the trained model to obtain the iris mask, and the inner and outer boundaries of the iris. Some post-processing operations are further used to refine the predictions. The team finally submitted 4 entries (from NIR-ISL2021031601_1 to NIR-ISL2021031601_4) for evaluation, where the main difference between entries is in the backbone network. Due to the business interests, some important algorithm details were not introduced by the team.

CKB. The team employs three separate Mask R-CNN models with the resnet50 and FPN networks as the backbone for segmenting the iris mask, the inner and outer boundary masks of the iris, respectively. Compared with the single multi-task model, the multi-model setting clearly reduces the complexity of the task and the difficulty of convergence. The testing images are processed by the trained model to get the binary iris mask, and the inner and outer boundary masks of the iris, where the latter produces the inner and outer boundaries of the iris by edge extraction. The proposed model is encoded as NIR-ISL2021031901.

EyeCool^{††}. An improved U-Net model is first employed to segment the iris mask, the inner boundary mask of the iris and the outer boundary mask of the iris simultaneously. The improvements are two-fold, on the one hand, the EfficientNet-B5 is used as the encoder to improve the model’s feature extraction capability, on the other hand, a boundary attention module is proposed to improve the ability of the decoder to pay attention to iris boundaries. The core of the latter is inserting multiple side-out layers after the decoder convolutional layers to predict the iris boundary heatmaps, then these predicted iris boundary heatmaps act as the attention weights to make the decoder focus on the iris boundaries.

During the training, a joint supervision of Dice loss and MSE loss is used to update the model parameters, where the Dice loss is responsible for calculating the classification loss of the segmentation part, and the MSE loss is responsible for calculating the regression loss of the iris boundary heatmaps. To obtain the inner and outer boundaries of the iris to be submitted, the contours of the predicted inner and outer boundary masks of the iris are extracted firstly, and then the Least Square based ellipse fitting method is used to refine the iris boundaries. The team finally submitted 4 entries (from NIR-ISL2021041402_1 to NIR-ISL2021041402_4) that differed only in training techniques for evaluation.

6. Results and Discussion

In this section, the results of NIR-ISL 2021 are presented and discussed. A comprehensive analysis about the perfor-

^{††}The code and models are available at <https://github.com/neu-eyecool/NIR-ISL2021>.

Method	Iris Segmentation			Iris Localization			Rank Sum
	E1 (Rank)	E2 (Rank)	Rank Sum	mDice (Rank)	mHdis (Rank)	Rank Sum	
NIR-ISL2021030401	7	7	14	20	22	42	86
NIR-ISL2021030902	20	20	40	28	44	72	112
NIR-ISL2021041402_2	38	38	76	30	28	58	134
NIR-ISL2021041402_3	34	34	68	35	37	72	140
NIR-ISL2021041402_1	54	54	108	36	35	71	179
NIR-ISL2021030701_5	37	37	74	56	52	108	182
NIR-ISL2021031401_2	42	42	84	50	57	107	191
NIR-ISL2021030801	36	36	72	55	73	128	200
NIR-ISL2021030701_1	51	51	102	56	52	108	210
NIR-ISL2021030701_7	54	54	108	56	52	108	216
NIR-ISL2021030701_6	59	59	118	56	52	108	226
NIR-ISL2021031601_3	70	69	139	53	39	92	231
NIR-ISL2021031601_4	75	75	150	46	35	81	231
NIR-ISL2021031601_1	74	74	148	48	49	97	245
NIR-ISL2021031401_1	63	64	127	60	60	120	247
NIR-ISL2021030701_3	73	73	146	56	52	108	254
NIR-ISL2021031601_2	99	99	198	46	18	64	262
NIR-ISL2021030701_2	80	80	160	56	52	108	268
NIR-ISL2021030701_4	80	80	160	56	52	108	268
NIR-ISL2021030101	81	81	162	111	98	209	371
NIR-ISL2021022201	86	86	172	98	105	203	375
NIR-ISL2021022301_1	106	106	212	120	123	243	455
NIR-ISL2021022801	115	115	230	121	115	236	466
NIR-ISL2021031901	126	126	252	121	122	243	495
NIR-ISL2021022601	129	129	258	128	128	256	514
NIR-ISL2021022301_2	128	128	256	134	133	267	523

Table 3. Summary results of different entries on both the iris segmentation and localization tasks. The entries are listed in a descending order based on their rank sum on all evaluation measures. For each evaluation metric, the first-place entry is labeled in red, the second-place entry is labeled in green, while the third-place entry is labeled in cyan. The top-3 performing entries on any one metric are labeled in yellow.

mance of the submitted entries is conducted in the following subsections.

6.1. Quantitative Evaluation

According to the experimental settings, each submitted entry is evaluated on the multiple testing sets using multiple metrics. Multiple calculated metric scores and corresponding rankings are finally aggregated to compare the overall performance of different entries. We first present the prediction results for iris segmentation and localization separately, then we summarize the results to discuss the performances of the submitted entries. Table 3 gives a summary of the overall evaluation results achieved by each entry from all testing sets. More complete ranking results from the challenge can be found in the supplementary material.

Results of Iris Segmentation. In Table 3, to reflect the overall segmentation performance of each entry on different datasets, $E1$ (Rank) or $E2$ (Rank) is calculated as the sum of metric rankings on all testing set, which are obtained based on the corresponding metric score. For instance, NIR-ISL2021030401 achieved the $E1$ of 0.3584%, 0.3961%, 0.3014%, 0.5293%, and 0.3748% on the CASIA-Iris-Distance, CASIA-Iris-Complex-Occlusion, CASIA-Iris-Complex-Off-angle, CASIA-Iris-M1, and CASIA-Iris-Africa, respectively. By comparing the $E1$ score with that of other entries, the $E1$ ranking of NIR-ISL2021030401 on the former testing sets was 2, 1, 1, 2, and 1, respectively. Hence, the $E1$ (Rank) of NIR-ISL2021030401 was 7(2+1+1+2+1). Furthermore, the $E1$ (Rank) and $E2$ (Rank) are added together to form the Rank Sum metric of iris seg-

mentation.

As can be seen from Table 3, different entities demonstrate almost identical rankings on the $E1$ and $E2$, which reflects that the functionality of the two metrics are equivalent to some extent. This may be the reason why the NICE.I competition [5] only adopted the $E1$ as the measure for ranking the competition methods. Further observations can see that the diverse results on different entities. The best solution for iris segmentation is NIR-ISL2021030401 with the smallest $E1$ (Rank) of 7 and the smallest $E2$ (Rank) of 7, yielding the smallest Rank Sum of 14. The second best solution for iris segmentation is NIR-ISL2021030902 (20, 20, 40), while NIR-ISL2021041402.3 achieves the third best solution (34, 34, 68). By contrast, the last five solutions, *i.e.*, NIR-ISL2021022301_1, NIR-ISL2021022801, NIR-ISL2021031901, NIR-ISL2021022601 and NIR-ISL2021022301_2, produce the worst results with the $E1$ (Rank) and $E2$ (Rank) above 100, and are far behind the top-3 performing solutions. Furthermore, the ranking gap between the best solution and the second best solution is as high as 13 for both $E1$ and $E2$, indicating the superior iris segmentation performance of the best solution, *i.e.*, NIR-ISL2021030401.

Results of Iris Localization. In Table 3, for each entry, the localization metric $mDice$ (Rank) or $mHdis$ (Rank) as well as yielding Rank Sum is also calculated in a similar manner to iris segmentation. However, different from $E1$ and $E2$ in the iris segmentation, the $mDice$ and $mHdis$ do not always maintain consistency, resulting in entities with better $mDice$ (Rank) being worse at $mHdis$ (Rank), *e.g.*, NIR-ISL2021030902 ($mDice$: the second-place vs. $mHdis$: outside the top three). This may because that the Dice index calculates the overlapping ratio between the predicted iris boundary mask and its ground truth, which is not exactly equivalent to the shape similarity between the predicted iris boundary and its ground truth measured by the Hausdorff distance.

The top three entries w.r.t. $mDice$ are NIR-ISL2021030401, NIR-ISL2021030902 and NIR-ISL2021041402.2, whereas on the $mHdis$ metric, NIR-ISL2021031601_2, NIR-ISL2021030401 and NIR-ISL2021041402.2 come at the top. Further observation reveals that the ranking gap between the top three entries is relatively low, being below 10 on both metrics, which suggests the quite good and close iris localization performance achieved by them. Besides, NIR-ISL2021030401 and NIR-ISL2021041402.2 are consistently ranked in the top three on both metrics, indicating their superior iris localization performance. Overall, if both metrics of iris localization are considered simultaneously, we can rank the submitted entries according to the Rank Sum metric. As a result, the top three entries are NIR-ISL2021030401, NIR-ISL2021041402.2 and NIR-ISL2021031601_2. In

addition, the former five worst-performing entries in iris segmentation also perform worst in iris localization and are far behind the top three entries.

Summary of Evaluation Results. From the perspective of conventional iris recognition, iris segmentation and localization are necessary procedures in the iris preprocessing. Therefore, judging whether a submitted entity is suitable for deployment as a pre-processing plug-in in an actual iris recognition system requires to consider both the iris segmentation and localization performance. As shown in Table 3, the entries are listed according to the order of their rank sum on the iris segmentation and localization. The lower the rank sum, the higher the overall performance.

In summary, the top three entries of NIR-ISL 2021 according to the overall rank sum in descending order are NIR-ISL2021030401 with Rank Sum of 56, NIR-ISL2021030902 with Rank Sum of 112, and NIR-ISL2021041402.2 with Rank Sum of 134. Clearly, NIR-ISL2021030401 retains a considerable advantage over the NIR-ISL2021030902 and NIR-ISL2021041402.2 in terms of Rank Sum, suggesting its state-of-the-art iris segmentation and localization performance. Further analysis indicates that the high performance is attained by NIR-ISL2021030401 first because the two independent single-tasking model setting makes the model training easier than the multi-tasking model used in the NIR-ISL2021030902 and NIR-ISL2021041402.2. Secondly, a novel transfer learning training technique improves the generalization ability of the model and model ensemble based Test-Time Augmentation refines the prediction results, which are lacking in the NIR-ISL2021030902 and NIR-ISL2021041402.2. Thirdly, thanks to a stronger ResNet-34 as the backbone, NIR-ISL2021030401 achieves more powerful feature extraction capability than the lightweight EfficientNetV2 and EfficientNet-B5 used in the NIR-ISL2021030902 and NIR-ISL2021041402.2, respectively.

Then we further find that NIR-ISL2021030902 is superior to NIR-ISL2021041402.2 probably first because NIR-ISL2021030902 adopts three parallel and specific decoder branches for multi-task prediction, obviously having more powerful feature decoding capability than the single decoder module in the NIR-ISL2021041402.2. Also NIR-ISL2021030902 exploits additional prior knowledge (*or “family” relationship*) between iris mask and iris boundaries to mutually promote both challenge tasks, resulting in a more stable and high-performance model, which is ignored in NIR-ISL2021041402.2. Nevertheless, NIR-ISL2021041402.2 and its other variants from the team NIR-ISL2021041402 still produce competitive results, especially thanks to the proposed boundary attention module, which helps the decoder to pay more attention to iris boundaries in the multi-task prediction. Hence a more accurate iris localization performance is attained (the best ranks 2st).

To get better insight into the key of effective models, we also perform negative case analysis to seven worst models, *i.e.*, NIR-ISL2021030101, NIR-ISL2021022201, NIR-ISL2021022301_1, NIR-ISL2021022801, NIR-ISL2021031901, NIR-ISL2021022601, and NIR-ISL2021022301_2, whose overall Rank Sum reaches more than 300 and far exceeds the remaining models with reliable performance. From their proposed algorithms, we summarize the possible causes of poor performance: *i*) most of less competitive models adopt the off-the-shelf U-Net or Mask R-CNN, *etc.*, but lack significant improvements or customized model design for the challenge, *e.g.*, NIR-ISL2021030101, NIR-ISL2021022201, NIR-ISL2021022801, NIR-ISL2021031901, NIR-ISL2021022601, *ii*) iris localization relies heavily on post-processing on iris mask, ignoring explicit localization model learning and prediction, *e.g.*, NIR-ISL2021022301_1, NIR-ISL2021022301_2, and *iii*) regarding iris boundary map rather than iris boundary mask as the model target may result in discontinuous and incomplete iris boundaries as the imbalanced problem of positive (boundary) and negative (non-boundary) samples in the iris boundary map is easy to yield an inferior model [11], *e.g.*, NIR-ISL2021022601.

6.2. Qualitative Evaluation

In this section, we further carry out a qualitative comparison of some submitted entries in terms of predicted binary iris segmentation mask and inner and outer boundaries of the iris. Here, two representative and challenging samples in CASIA-Iris-Distance, CASIA-Iris-Complex-Occlusion, CASIA-Iris-Complex-Off-angle, CASIA-Iris-M1, and CASIA-Iris-Africa were separately selected to demonstrate the varying iris segmentation and localization performance of the submitted entries in different non-cooperative environments, as illustrated in Figure 2, Figure 3, Figure 4, Figure 5, and Figure 6.

More specifically, (a) and (e) show the original iris images, and corresponding ground truths of iris mask (blue) and inner (green) and outer (red) boundaries of the iris, respectively. Then (b-d) shows the iris segmentation (left) and localization (right) results from the best three entries of NIR-ISL 2021, *i.e.*, NIR-ISL2021030401, NIR-ISL2021030902, and NIR-ISL2021041402_2, respectively, where false positive, false negative, and true positive pixels are labeled in green, red, and blue in the segmentation results, respectively. Similarly, (f-h) shows the iris segmentation (left) and localization (right) results from the worst three entries of NIR-ISL 2021, *i.e.*, NIR-ISL2021031901, NIR-ISL2021022601, and NIR-ISL2021022301_2, respectively.

As can be visually observed, the best three entries are superior to the worst three entries in most images when

looking at the quality of iris segmentation and localization, which is consistent with quantitative conclusions in Section 6.1. It indicates that the best three entries are robust with different types of degraded iris images in non-cooperative environments, but further improvements are still necessary to better deal with some hard cases, such as Figure 2 (the top image), Figure 3 (the top image), and Figure 5 (the bottom image). Further observation finds that NIR-ISL2021041402_2 has a better localization accuracy than other two best entries in the Figure 2 (the top image) and Figure 5 (the bottom image), which is most likely due to the boundary attention module it proposed. In addition, if the boundary predicted by NIR-ISL2021031901 is fitted via ellipses, the iris localization performance may be further improved. NIR-ISL2021022301_2 used only circles to approximate iris boundaries, which may be unsuitable for off-angle iris images, *e.g.*, Figure 4, and Figure 6.

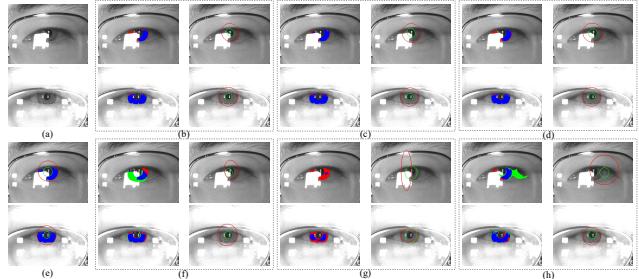


Figure 2. Samples of iris segmentation and localization results from best and worst three entries on the CASIA-Iris-Distance.

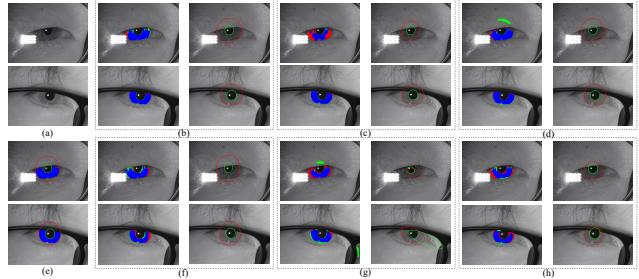


Figure 3. Samples of iris segmentation and localization results from best and worst three entries on the CASIA-Iris-Complex-Occlusion.

7. Conclusions

This paper presented a summary of the 2021 NIR Iris Challenge Evaluation in Non-cooperative Environments: Segmentation and Localization (NIR-ISL 2021), which was held in conjunction with the 2021 International Joint Conference on Biometrics (IJCB 2021). The goal of the challenge was to bring together researchers interested in the iris segmentation problem, to validate the performance of their existing or newly developed algorithms on the same non-cooperative Asian and African NIR iris datasets. In

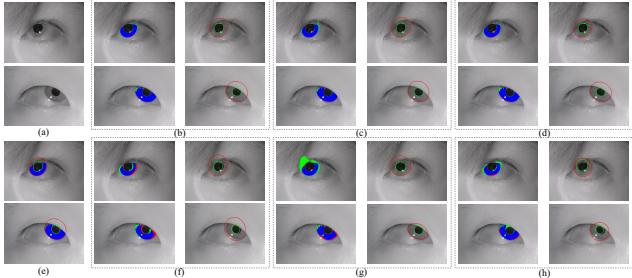


Figure 4. Samples of iris segmentation and localization results from best and worst three entries on the CASIA-Iris-Complex-Off-angle.

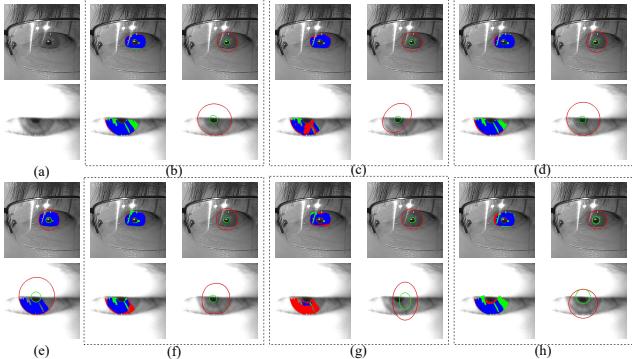


Figure 5. Samples of iris segmentation and localization results from best and worst three entries on the CASIA-Iris-M1.

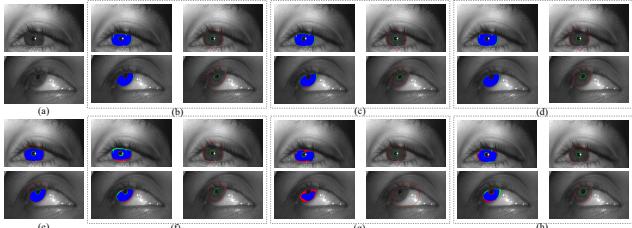


Figure 6. Samples of iris segmentation and localization results from best and worst three entries on the CASIA-Iris-Africa.

the final round, a total of 27 valid entries from 14 teams participated in the group evaluation. The top three entries achieved solid and robust iris segmentation and localization results in most noisy iris images. The datasets used in the challenge will be made publicly available at the challenge website (<https://sites.google.com/view/nir-is12021/home>).

In the future, we hope to continue the series of challenges to boost the research on more lightweight and weakly supervised even unsupervised iris segmentation and localization methods for practical deployment and reducing reliance on manual annotations.

Acknowledgment

This work is supported by the Beijing University of Civil Engineering and Architecture Research Capacity Promotion Program for Young Scholars (Grant No. X21079), the National Natural Science Foundation of China (Grant No. U1836217, 62006225, 62071468, 61906199) and the National Key Research and Development Program of China (Grant No. 2017YFC0821602).

tion Program for Young Scholars (Grant No. X21079), the National Natural Science Foundation of China (Grant No. U1836217, 62006225, 62071468, 61906199) and the National Key Research and Development Program of China (Grant No. 2017YFC0821602).

References

- [1] F. Boutros, N. Damer, F. Kirchbuchner, and A. Kuijper. Eye-mms: Miniature multi-scale segmentation network of key eye-regions in embedded applications. In *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, pages 3665–3670, 2019.
- [2] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler. Mobile iris challenge evaluation (miche)-i, biometric iris dataset and protocols. *Pattern Recognition Letters*, 57:17–23, 2015.
- [3] R. R. Jha, G. Jaswal, D. Gupta, S. Saini, and A. Nigam. Pixisegnet: pixel-level iris segmentation network using convolutional encoder-decoder with stacked hourglass bottleneck. *IET Biometrics*, 9(1):11–24, 2020.
- [4] N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan. Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In *IAPR International Conference on Biometrics (ICB)*, pages 1–8, 2016.
- [5] H. Proen  a and L. A. Alexandre. The nice. i: noisy iris challenge evaluation-part i. In *IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS)*, pages 1–4, 2007.
- [6] S. D. Shirke and C. Rajabshushnam. Local gradient pattern and deep learning-based approach for the iris recognition at-a-distance. *International Journal of Knowledge-based and Intelligent Engineering Systems*, 25(1):49–64, 2021.
- [7] K. Sirinukunwattana, J. P. W. Pluim, H. Chen, X. Qi, P. Heng, Y. B. Guo, L. Y. Wang, B. J. Matuszewski, E. Bruni, U. Sanchez, et al. Gland segmentation in colon histology images: The glas challenge contest. *Medical Image Analysis*, 35:489–502, 2017.
- [8] A. Valenzuela, C. Arellano, and J. E. Tapia. Towards an efficient segmentation algorithm for near-infrared eyes images. *IEEE Access*, 8:171598–171607, 2020.
- [9] V. Varkarakis, S. Bazrafkan, and P. Corcoran. Deep neural network and data augmentation methodology for off-axis iris segmentation in wearable headsets. *Neural Networks*, 121:101–121, 2020.
- [10] C. Wang, J. Muhammad, Y. Wang, Z. He, and Z. Sun. Towards complete and accurate iris segmentation using deep multi-task attention network for non-cooperative iris recognition. *IEEE Transactions on information forensics and security*, 15:2944–2959, 2020.
- [11] C. Wang, Y. Wang, B. Xu, Y. He, Z. Dong, and Z. Sun. A lightweight multi-label segmentation network for mobile iris biometrics. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 1006–1010, 2020.
- [12] Q. Zhang, H. Li, Z. Sun, and T. Tan. Deep feature fusion for iris and periocular biometrics on mobile devices. *IEEE Transactions on Information Forensics and Security*, 13(11):2897–2912, 2018.