

The Impact of Preprocessing on Deep Representations for Iris Recognition on Unconstrained Environments

Luiz A. Zanlorensi*, Eduardo Luz†, Rayson Laroca*, Alceu S. Britto Jr.‡, Luiz S. Oliveira*, David Menotti*

*Department of Informatics, Federal University of Paraná (UFPR), Curitiba, PR, Brazil

†Computing Department, Federal University of Ouro Preto (UFOP), Ouro Preto, MG, Brazil

‡Postgraduate Program in Informatics, Pontifical Catholic University of Paraná (PUCPR), Curitiba, PR, Brazil

Abstract—The use of iris as a biometric trait is widely used because of its high level of distinction and uniqueness. Nowadays, one of the major research challenges relies on the recognition of iris images obtained in visible spectrum under unconstrained environments. In this scenario, the acquired iris are affected by capture distance, rotation, blur, motion blur, low contrast and specular reflection, creating noises that disturb the iris recognition systems. Besides delineating the iris region, usually preprocessing techniques such as normalization and segmentation of noisy iris images are employed to minimize these problems. But these techniques inevitably run into some errors. In this context, we propose the use of deep representations, more specifically, architectures based on VGG and ResNet-50 networks, for dealing with the images using (and not) iris segmentation and normalization. We use transfer learning from the face domain and also propose a specific data augmentation technique for iris images. Our results show that the approach using non-normalized and only circle-delimited iris images reaches a new state of the art in the official protocol of the NICE.II competition, a subset of the UBIRIS database, one of the most challenging databases on unconstrained environments, reporting an average Equal Error Rate (EER) of 13.98% which represents an absolute reduction of about 5%.

I. INTRODUCTION

Biometrics has many applications such as verification, identification, duplicity verification, which makes it an important research area. A biometric system basically consists of extracting and matching distinctive features from a person. These patterns are stored as a new sample which is subsequently used in the process of comparing and determining the identity of each person within a population. Considering that biometric systems require robustness combined with high accuracy, the methods applied to identify individuals are in constant development.

Biometric methods that identify people based on their physical or behavioral features are interesting due to the fact that a person cannot lose or forget its physical characteristics, as can occur with other means of identification such as passwords or identity cards [1]. The use of eye traits becomes interesting because it provides a framework for non-invasive screening technology. Another important factor is that biomedical literature suggests that irises are as distinct as other biometric sources such as fingerprints or patterns of retinal blood vessels [2].

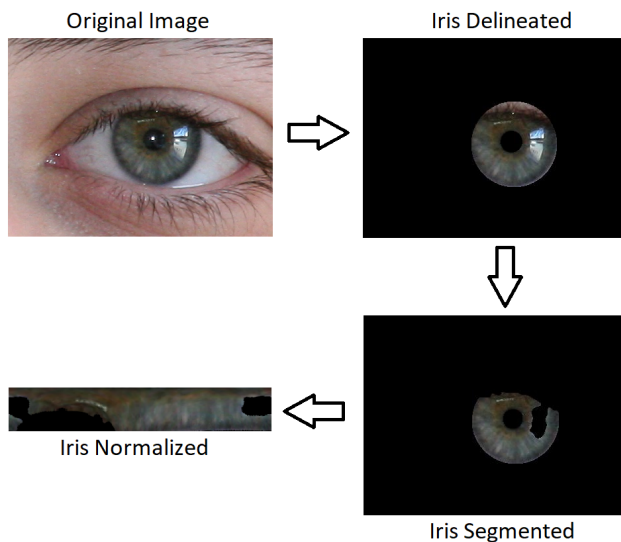


Fig. 1. Iris preprocessing stages.

Research using iris images obtained in near-infrared (NIR) showed very promising results and reported low rates of recognition error [1], [3]. Currently, one of the greatest challenges in iris recognition is the use of images obtained in visible spectrum under uncontrolled environments [4], [5]. The main difficulty in iris recognition using these images is that they may have problems such as noise, rotation, blur, motion blur, low contrast, specular reflection, among others. Generally, techniques such as normalization [6] and segmentation [4] are applied to correct or reduce these problems.

With the recent development of deep learning, some applications of this methodology in periocular [7], [8] and iris recognition [9]–[13] have been performed with interesting results being reported. The main problem with deep network architectures when trained in small databases is the lack of enough data for generalization usually producing over-fitted models. One solution to this problem is the use of transfer learning [14]. Another action is the use of data augmentation techniques [15], [16].

In this context, the goal of this work is to evaluate the impact of image preprocessing for deep iris representations acquired

on uncontrolled environments.

We evaluated the following preprocesses: iris delineation, iris segmentation for noise removal and iris normalization. The delineation process defines the outer and inner boundaries of the iris (i.e., the pupil). The segmentation process removes the regions where the iris is occluded. Finally, the normalization process consists of transforming the circular region of the iris from the Cartesian space into the Polar one resulting in a rectangular region.

To improve generalization and avoid overfitting, two Convolutional Neural Network (CNN) models trained for face recognition were fine-tuned and then used as iris representation (or features). From the comparison of the obtained results, we can observe that the approaches using only delineated but non-normalized and non-segmented iris image as input for the networks generated new state-of-the-art results for the official protocol of the NICE.II competition. We also evaluate the impact of the non-delineated iris images, the squared bounding box. We chose the NICE.II competition, which is composed as a subset of UBIRIS.v2 database, to evaluate our proposal because it is currently one of the most challenging databases for iris recognition on uncontrolled environments, presenting problems such as noise and different image resolutions.

Since iris recognition may refer to both identification and verification problems, it is important to point out that this paper addresses the verification problem, i.e., to verify if two images are from the same subject. Moreover, the experiments are performed on the NICE.II competition protocol which is a verification problem.

The remainder of this work is organized as follows. In Section II, we described the methodologies that achieved the best results in NICE.II so far and works that employed deep learning for iris recognition. Section III describes the protocol and database of NICE.II, and the metrics (i.e., Equal Error Rate (EER) and decidability) used in our experiments for comparison. Section IV presents and describes how the experiments were performed. The results are presented and discussed in Section V, while the conclusions are given in Section VI.

II. RELATED WORK

In this section, we briefly survey the best performing works in the NICE.II competition and then describe deep learning approaches for iris recognition.

A. The NICE.II competition

Among eight participants ranked in the NICE.II competition, six of them proposed methodologies considering only iris modality [17]–[22]. The other two fused iris and periocular region modalities [23], [24]. All these approaches were summarized in the paper “The results of the NICE.II Iris biometrics competition” [25] which appeared in 2012.

Taking advantage of periocular and iris information, the winner [23] of the NICE.II reported a decidability value of 2.57. Their methodology consists of image preprocessing, feature extraction from the iris and periocular region, matching

of iris and periocular region data, and fusion of the matching results. For iris feature extraction, ordinal measures and color histogram were used, while texton histogram and semantic information were employed for the periocular region. The matching scores were obtained using SOBoost learning [26] for ordinal measures, diffusion distance for color histogram, chi-square distance for texton histogram, and exclusive-or logical operator for semantic label. The approaches were combined through the sum rule at the matching level.

Using only features extracted from the iris, the best result in NICE.II achieved a decidability value of 1.82 and *EER* of 19%, ranking second in the competition [17]. In that work, Wang et al. first performed the segmentation and normalization of the iris images using the methodology proposed by Daugman [6]. Then, according to the normalization, irises were partitioned into different numbers of patches. With the partitions, features were extracted using Gabor filters. Finally, the AdaBoost algorithm was applied to select the best features and compute the similarity. To the best of our knowledge, this is the state of the art on the NICE.II database when only the iris is used as the biometric trait.

Another work using only iris features for recognition with images obtained in visible wavelength and in unconstrained environments is presented in [27]. Their methodology consists of four steps: segmentation, normalization, feature extraction and matching with weighted score-level fusion. Features such as wavelet transform, keypoint-based, generic texture descriptor and color information were extracted and combined. Results were presented in two subsets of UBIRIS.v2 and MobBIO databases, reporting an EER of 22.04% and 10.32%, respectively. Even though this result is worse than the best one obtained in the NICE.II, which is a subset of UBIRIS.v2 database, a direct comparison is not fair because the data used might be different.

B. Deep learning in iris recognition

Deep learning is one of the most recent and promising machine learning techniques [28]. Thus, it is natural that there are still a few works that use and apply this technique in iris images. We describe some of these works below.

The first work applying deep learning to iris recognition was the DeepIris framework, proposed by Liu et al. [9] in 2016. Their method was applied to the recognition of heterogeneous irises, where the images were obtained with different types of sensors, i.e., the cross-sensor scenario. The major challenge in this field is the existence of a great intra-class variation, caused by sensor-specific noises. Thus, handcrafted features generally do not perform well in this type of database. The study proposed a framework based on deep learning for verification of heterogeneous irises, which establishes the similarity between a pair of iris images using CNNs by learning a bank of pairwise filters. This methodology differs from the handcrafted feature extraction, since it allows direct learning of a non-linear mapping function between pairs of iris images and their identity with a pairwise filter bank (PFB) from different sources. Thereby, the learned PFBs can be used for

new and also different data/subjects, since the learned function is used to establish the similarity between a pair of images. The experiments were performed in two databases: Q-FIRE and CASIA cross sensor. Promising results were shown to be better than the baseline methodology, with EER of 0.15% in Q-FIRE database and EER of 0.31% in CASIA cross-sensor database.

Gangwar & Joshi [10] also developed a deep learning application for iris recognition in images obtained from different sensors, called DeepIrisNet. In their study, a CNN was used to extract features and representations of iris images. Two databases were used in the experiments: ND-iris-0405 and ND-CrossSensor-Iris-2013. In addition, two CNN architectures were presented, namely DeepIrisNet-A and DeepIrisNet-B. The former is based on standard convolutional layers, containing 8 convolutional layers, 8 normalization layers, and 2 dropout layers. DeepIrisNet-B uses inception layers [29] and its structure consists of 5 layers of convolution, 7 of normalization, 2 of inception and 2 of dropout. The results presented five comparisons: effect of segmentation, image rotation analysis, input size analysis, training size analysis and network size analysis. The proposed methodology demonstrated better robustness compared to the baseline.

The approach proposed by Al-Waisy et al. [11] consists of a multi-biometric iris identification system, using both left and right irises from a person. Experiments were performed in databases of NIR images obtained in controlled environments. The process has five steps: iris detection, iris normalization, feature extraction, matching with deep learning and, lastly, the fusion of matching scores of each iris. During the training phase, the authors applied different CNN configurations and architectures, and chose the best one based on validation set results. A 100% rank-1 recognition rate was obtained in SDUMLA-HMT, CASIA-Iris-V3, and IITD databases. However, it is important to note that this methodology only works in a closed-world problem, since the matching score is based on the probability that an image belongs to a sample of a class known in the training phase.

Also using NIR databases obtained in controlled environments, Nguyen et al. [12] demonstrated that generic descriptors using deep learning are able to represent iris features. The authors compared five CNN architectures trained in the ImageNet database [30]. The CNNs were used, without fine-tuning, for the feature extraction of normalized iris images. Afterward, a simple multi-class Support Vector Machine (SVM) was applied to perform the classification (identification). Promising results were presented in LG2200 (ND-CrossSensor-Iris-2013) and CASIA-Iris-Thousand databases, where all the architectures report better accuracy recognition than the baseline feature descriptor [31].

Other applications with iris images include spoofing detection [32], recognition of mislabeled left and right iris images [33], liveness detection [34], iris location [35], gender classification [36] and sensor model identification [37].

Also one can find the application of deep learning to the periocular and sclera regions, using images captured in

uncontrolled environments [7], [8], [38].

In all works found in the literature that apply deep learning for iris recognition, the input image used is the normalized one, where a previous iris location and segmentation is also required before the normalization process. In this work, we evaluate the use of different input images for learning deep representations for iris recognition (verification). The input images were created using three preprocesses: iris delineation, iris segmentation for noise removal and iris normalization.

III. PROTOCOL AND DATABASE

In this section, we describe the experimental protocol and database proposed in the NICE.II competition, which was used to evaluate the proposed methodology in this paper and in some of the related works.

The first iris recognition competition created specifically with images obtained in visible spectrum under uncontrolled environments is the Noisy Iris Challenge Evaluation (NICE). The NICE competition is separated into two phases. The first one, called NICE.I (2008), was carried out with the objective of evaluating techniques for noise detection and segmentation of iris images. In the second competition, NICE.II (2010), the encoding and matching strategies were evaluated. The NICE images were selected as a subset of images of a larger database, the UBIRIS.v2 [39], which in turn comprises of 11,102 images and 261 individuals.

The main goal of NICE.I [4] was to answer the following question: “Is it possible to automatically segment a small target as the iris in unconstrained data (obtained in a non-cooperative environment)?”. The competition was attended by 97 research laboratories from 22 countries, which received a database of 500 images to be used as training data in the construction of their methodologies. These 500 iris images were made available by the organizer committee, along with segmentation masks. The masks were used as ground truth to assess the performance of iris segmentation methodologies. For the evaluation of the submitted algorithms, a new database containing 500 iris images was used to measure the pixel-to-pixel agreement between the segmentation masks created by each participant and the masks manually created by the committee. The performance of each submitted methodology was evaluated with the error rate, which gives the average proportion of correctly classified pixels.

In order to guarantee impartiality and evaluate only the results of the feature extraction and matching, all the participants of the second phase of NICE (NICE.II) used the same segmented iris images, which were obtained with the technique proposed by Tan et al. [40], winner of NICE.I [4]. The objective of NICE.II was to evaluate how different sources of image noise obtained in an uncontrolled environment may interfere in iris recognition. The training database consisted of 1,000 images along with the corresponding segmented iris masks. The task was to build an executable that received as input a pair of iris images and their respective masks, generating a file with the corresponding scores with dissimilarity of irises (d) as an output. The d metric follows some conditions:

- 1) $d(I, I) = 0$
- 2) $d(I_1, I_2) = 0 \Rightarrow I_1 = I_2$
- 3) $d(I_1, I_2) + d(I_2, I_3) \geq d(I_1, I_3)$

For the evaluation of the methodologies proposed by the participants, another unknown database containing 1,000 images and masks was employed. Some samples randomly selected from these images are shown in Fig. 2. Consider $IM = \{I_1, \dots, I_n\}$ as a set of iris images, $MA = \{M_1, \dots, M_n\}$ as their targeting binary masks and $id(\cdot)$ as the identity function on an image. A comparison of one-against-all will return a match set $D^I = \{d_1^i, \dots, d_m^i\}$ and non match $D^E = \{d_1^e, \dots, d_k^e\}$ of dissimilarity scores, respectively, for the cases where $id(I_i) = id(I_j)$ and $id(I_i) \neq id(I_j)$. The evaluation of the algorithms was performed using the decidability scores d' [41].

The metric or index d' measures how well separated are two types of distributions, so the recognition error corresponds to the overlap area

$$d' = \frac{|\mu_E - \mu_I|}{\sqrt{\frac{1}{2}(\sigma_I^2 + \sigma_E^2)}} \quad (1)$$

where the means of the two distributions are given by μ_I and μ_E , and σ_I and σ_E represent the standard deviations.



Fig. 2. Samples from the NICE.II database.

Since the goal of the competition is the iris recognition in noisy images obtained in an uncontrolled environment, the methodology presented in this paper focuses only on the features extraction and matching using iris information. As shown in previous studies [23], [24], [42], [43], promising results were achieved using the fusion of iris and periocular modalities, and since this fusion depends on the quality of each modality, it is interesting to study each one of them independently, i.e., using only iris or periocular information.

Considering that the decidability metric measures how discriminating the extracted features are, we have used it to compare our method with the state of the art. We also report the EER, which is based on the False Acceptance Rate (FAR) and the False Rejection Rate (FRR). The EER is determined by the intersection point of FAR and FRR curves. All reported results were obtained in the official test set of the NICE.II competition.

IV. PROPOSED APPROACH

The proposed approach consists of three main stages: image preprocessing, feature extraction and matching, as shown in Fig. 4. In the first stage, segmentation and normalization techniques are applied. We perform data augmentation in the preprocessed images to increase the number of training samples. The feature extraction is performed with two CNN models, which were fine-tuned using the original images, and the images generated through data augmentation. Finally, the matching is performed using the cosine distance. These processes are best described throughout this section.

A. Image preprocessing

In order to analyze the impact of the preprocessing, six different input images were generated from the original iris image, as shown in Fig. 3. In the first image scheme, irises are normalized with the standard rubber sheet model [6] using an 8 : 1 aspect ratio (512×64 pixels). In the second image scheme, the iris images are also normalized, however, instead of the standard 8 : 1 ratio they were rearranged in a 4 : 2 ratio (256×128 pixels), so that less interpolation is employed in the resizing process. In the third and last schemes, no normalization is performed, applying only the delineated iris images as input to the models. Non-normalized images have different sizes, according to the image capture distance. The impact of the segmentation technique for noise removal was also evaluated in all representations. Note that all the iris images used as inputs for the feature representation models are resized using bi-cubic interpolation to 224×224 pixels.

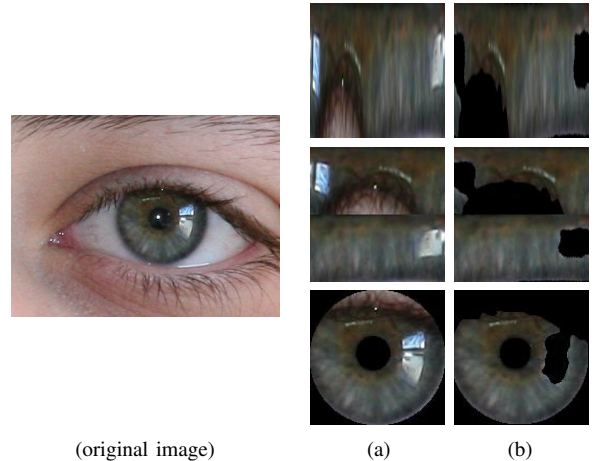


Fig. 3. Input Image. (a) and (b) show, respectively, non-segmented and segmented images for noise removal from the Nice.II database. From top to bottom, images with the aspect ratios of 8 : 1, 4 : 2 are shown, as well as non-normalized images.

The normalization through the rubber sheet model [6] aims to obtain invariance with respect to size and pupil dilatation. In the NICE.II database, the main problem is the difference of the iris size due to distances in the image capture. It is important to note that in non-normalized images, we use an arc delimitation preprocessing (i.e., two circles, an outer and an inner), based on the iris mask.

The segmentation process to noise removal was performed using the masks provided along with the database. These masks were obtained with the methodology proposed by Tan et al. [40], winner of NICE.I [4].

Considering these problems, the proposed approach aims to analyze the impact of non-normalization and non-segmentation of iris images to extract deep features.

B. Data Augmentation

Since the training subset has only 1,000 images belonging to 171 classes, it is important to apply data augmentation techniques to increase the number of training samples. The fine-tuning process can result in a better generalization of the models with more images. In this sense, we rotate the original images at specific angles.

The ranges of angles used were: -15° to 15° , -30° to 30° , -45° to 45° , -60° to 60° , -90° to 90° and -120° to 120° . For each range, the aperture is proportional, generating 4, 6 and 8 images for each original image, respectively. For example, considering the range -60° to 60° with 6 apertures, for each original image, another six were generated rotating -60° , -40° , -20° , 20° , 40° and 60° . Performing the validation of all these data augmentation methods on all input images, we determined (based on accuracy and loss) that the best range was -60° to 60° with 6 apertures. These parameters were applied to perform the data augmentation in the training set, totaling 7,000 images. Some samples generated by data augmentation can be seen in Fig. 5.

C. Convolutional Neural Network Model

For feature extraction of the iris images, the fine-tuning of two CNN models trained for face recognition were applied. The first model, called VGG, proposed in [44] and used in [7] for periocular recognition, has an architecture composed of convolution, activation (ReLU), pooling and fully connected layers. The second model (i.e., Resnet-50), proposed in [45] and trained for face recognition [46], has the same operations as VGG with the difference of being deeper and considering residual information between the layers.

For both models, we use the same architecture modifications and parameters described in [7]. In the training phase, the last layer (used to predict) was removed and two new layers were added. The new last layer, used for classification, is composed by 171 neurons, where each one corresponds to a class in the NICE.II training set and has a softmax-loss function. The layer before that is a fully-connected layer with 256 neurons used to reduce feature dimensionality.

The training set was divided into two new subsets with 80% of the data for training and 20% for validation. Two learning rates were used for 30 epochs, 0.001 for the first 10 epochs and 0.0005 for the remaining 20. Other parameters include momentum = 0.9 and batch size = 48. The number of epochs used for training was chosen based on the experiments carried out in the validation set (highest accuracy and lowest loss). For training (fine-tuning) the CNN models, the Stochastic Gradient Descent (SGD) optimizer was used. Similarly to [7], [47], we

do not freeze the weights of the pre-trained layers during training to perform the fine-tuning. The last layer of each model was removed and the features were extracted on the new last layer.

D. Matching

The matching was performed using the verification protocol proposed in the NICE.II competition. For this, the all against all approach was applied in the NICE.II test set, generating 4,634 intra-class pairs and 494,866 inter-class pairs.

The cosine metric, which measures the cosine of the angle between two vectors, was applied to compute the difference between feature vectors. This metric is used in information retrieval [48] due to its invariance to scalar transformation. The cosine distance metric is represented by

$$d_c(A, B) = 1 - \frac{\sum_{j=1}^N A_j B_j}{\sqrt{\sum_{j=1}^N A_j^2} \sqrt{\sum_{j=1}^N B_j^2}} \quad (2)$$

where A and B stand for the feature vectors. We also employed other distances such as Euclidean, Mahalanobis, Jaccard and Manhattan. However, due to its best performance, only the cosine distance is reported.

V. RESULTS

In this section, we present the results of the experiments for validating our proposal using the test set from the Nice.II competition database. Initially, the impact of the proposed data augmentation techniques is shown using non-segmented iris images. Then, we analyze the impact of both iris segmentation and iris delineation. Finally, we compare the best results obtained by our approaches with the state of the art. In all subsections, the impact of normalization is also analyzed. Note that in all experiments, the mean and standard deviation values from 30 runs are reported. For analyzing the different results, we perform statistical paired t-tests at a significance level $\alpha = 0.05$.

A. Data Augmentation

In the first analyses, we evaluate the impact of the data augmentation. For ease of analysis, all iris image employed in this initial experiment may contain noise in the iris region, i.e. no segmentation preprocessing is applied. As shown in Table I, in all cases where data augmentation is used, the decidability and EER values improved with statistical difference. Note that the models trained with data augmentation reported smaller standard deviation. In general, it is also observed that non-normalization yielded better results than 8 : 1 and 4 : 2 normalization schemes for both trained models, i.e. VGG and ResNet-50.

It is worth noting that the largest differences occurred in the non-normalized inputs, with greater impact specifically in the ResNet-50 model, where the mean EER dropped 7.53% and the decidability improved 0.6361 when applying data augmentation.

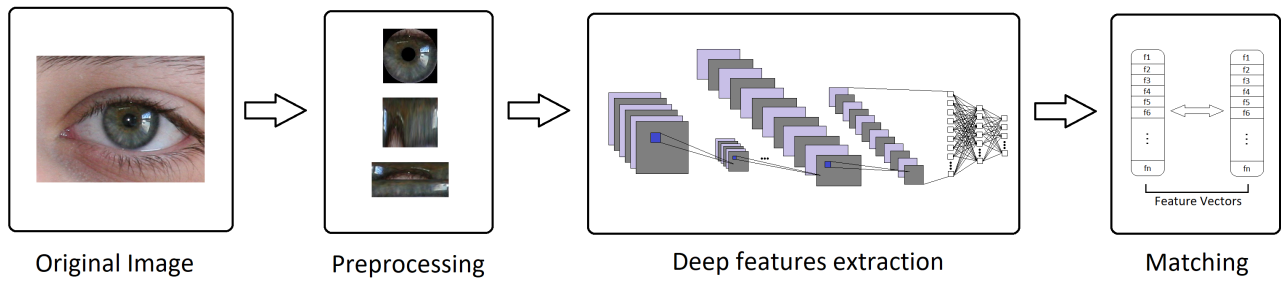


Fig. 4. The proposed approach.

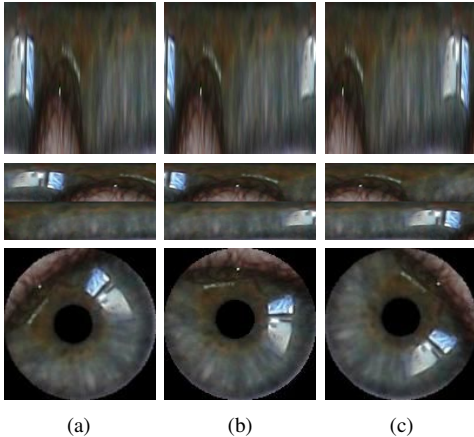


Fig. 5. Data augmentation samples: (a) -45° rotated images, (b) original images and (c) 45° rotated images.

B. Segmentation

In the second analysis, the impact of the segmentation for noise removal is evaluated. For such aim, two models are trained (fine-tuned): using segmented and non-segmented images, all with data augmentation.

As can be seen in Table I, for the VGG model segmentation has improved the results. On the other hand, for the ResNet-50 model, the non-segmented images have presented better results. For both models, statistical difference is achieved in two situations and in another one (light cyan color) there is no statistical difference.

Regarding the better results achieved by the ResNet-50 models when using non-segmented images, we hypothesized that this might be related to the fact that the ResNet-50 architecture uses residual information and it is deeper compared to VGG. Thus, some layers of ResNet-50 might be responsible for extracting discriminant patterns present in regions that were occluded in the segmented images, but not in non-segmented ones. Moreover, in segmented images, black regions (zero values) were employed for representing noise regions, and no special treatment was given for those regions.

It is noteworthy that segmentation is a complex process and might impact positively or negatively. However, as the best results here were achieved by the ResNet-50 models when using non-segmented images, we state that using the suitable

representation model the segmentation preprocessing can be disregarded.

Once again, non-normalization showed to provide better results in all scenarios, being more expressive here than in the data augmentation analysis.

TABLE I
IMPACT OF THE DATA AUGMENTATION (DA) AND SEGMENTATION (SEG.)
ON THE EFFECTIVENESS OF IRIS VERIFICATION FOR VGG AND
RESNET-50 NETWORKS.

Network	Norm.	DA Seg.	EER (%)	Decidability
VGG	8 : 1		26.19 ± 1.95	1.3140 ± 0.1246
VGG	8 : 1	✓	23.63 ± 1.33	1.4712 ± 0.0881
VGG	8 : 1	✓ ✓	22.58 ± 1.07	1.5437 ± 0.0697
ResNet-50	8 : 1		24.38 ± 1.41	1.4297 ± 0.0916
ResNet-50	8 : 1	✓	19.18 ± 0.75	1.7988 ± 0.0552
ResNet-50	8 : 1	✓ ✓	20.68 ± 1.39	1.6801 ± 0.1071
VGG	4 : 2		24.77 ± 1.42	1.4127 ± 0.1001
VGG	4 : 2	✓	18.74 ± 0.89	1.8527 ± 0.0712
VGG	4 : 2	✓ ✓	18.00 ± 0.93	1.9055 ± 0.0750
ResNet-50	4 : 2		22.78 ± 1.22	1.5307 ± 0.0853
ResNet-50	4 : 2	✓	17.11 ± 0.53	1.9822 ± 0.0482
ResNet-50	4 : 2	✓ ✓	17.44 ± 0.85	1.9450 ± 0.0803
VGG	Non-Norm		23.32 ± 1.10	1.4891 ± 0.0740
VGG	Non-Norm	✓	17.49 ± 0.90	1.9529 ± 0.0760
VGG	Non-Norm	✓ ✓	17.48 ± 0.68	1.9439 ± 0.0589
ResNet-50	Non-Norm		21.51 ± 0.97	1.6119 ± 0.0677
ResNet-50	Non-Norm	✓	13.98 ± 0.55	2.2480 ± 0.0528
ResNet-50	Non-Norm	✓ ✓	14.89 ± 0.78	2.1781 ± 0.0794

C. Delineation

Here we evaluated the impact on the recognition of using a usual delineated iris image and a non-delineated iris image, i.e., applying only the *squared* iris bounding box as input to the deep feature extractor. In both situations, non-normalized and non-segmented images are used. A delineated iris image and its corresponding bounding box (or non-delineated) are shown in Fig. 6.

The comparison of the results of this analysis is shown in Table II. Although the results reported by delineated iris images are better, there is no statistical difference. From this result, we state that the iris bounding box can be used as input for deep representation without the iris delineating (a.k.a. location) preprocessing.

Considering that the bounding box is not pure iris, comparison with other iris recognition methods may not be fair, since

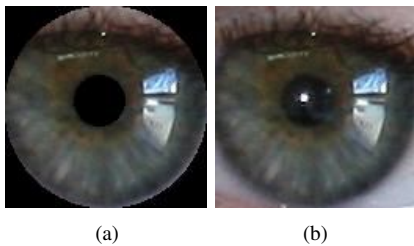


Fig. 6. Input images: (a) delineated iris and (b) non-delineated iris / bounding box version.

TABLE II
COMPARISON OF DELINEATED AND NON-DELINEATED IRIS IMAGES. BOTH WITH NO SEGMENTATION (FOR NOISE REMOVAL) AND NORMALIZATION.

Method	Delineated	EER (%)	Decidability
VGG	✓	17.49 ± 0.90	1.9529 ± 0.0760
VGG		17.52 ± 0.98	1.9652 ± 0.0790
Resnet-50	✓	13.98 ± 0.55	2.2480 ± 0.0528
Resnet-50		14.26 ± 0.47	2.2304 ± 0.0542

there may be discriminant patterns that have been extracted from regions outside the iris. Therefore, our methodology was compared with the state of the art using delineated iris images.

D. The state of the art

At last, the results attained with our models using non-normalized, non-segmented, and delineated iris images are compared with the state-of-the-art approaches and it is shown in Table III.

These experiments showed that the representations learned using deep models perform better the iris verification task on the NICE.II competition when the preprocessing steps of normalization and segmentation (for noise removal) are removed, outperforming the state-of-the-art method, which uses preprocessed images.

TABLE III
RESULTS ON THE NICE.II COMPETITION DATABASE. COMPARISON OF THE STATE OF THE ART WITH THE RESULTS ACHIEVED BY OUR PROPOSED APPROACHES USING NON-NORMALIZED, NON-SEGMENTED, AND DELINEATED IRIS IMAGES.

Method	EER (%)	Decidability
Wang et al [17]	19.00	1.8213
Proposed VGG	17.48	1.9439
Proposed ResNet-50	13.98	2.2480

VI. CONCLUSION

In this paper we evaluated the impact of iris image preprocessing for iris recognition on unconstrained environments using deep representations. Different combinations of (non-)normalized and (non-)segmented images as input for the system were evaluated.

Using these iris images, a fine-tuning process of two CNN architectures pre-trained for face recognition was performed. These models were applied to extract deep representations.

The matching, on a verification protocol, was performed with the cosine metric. A significant improvement in the results of both models was achieved using the proposed data augmentation approach. For both models, non-normalized iris achieved a better result. In addition, we verified that the use of non-delineated iris images is slightly worse than the ones attained when using delineated images, but no significant difference was reached. However, for a fair comparison with a state-of-the-art method, we used only delineated images because they represent the pure irises.

The experiments showed that the models learned on the ResNet-50 architecture using non-segmented and non-normalized images reported the best results achieving a new state of the art in the NICE.II official protocol – one of the most challenging databases on uncontrolled environments.

As future work, we intend to evaluate this approach in larger databases and the performance of other network architectures for feature extraction and also transfer learning from other domains than face recognition. Dealing with noisy iris regions and the analysis of its impact on iris recognition is the plan for future work as well.

ACKNOWLEDGMENT

The authors thank the *Conselho Nacional de Pesquisa e Desenvolvimento* (CNPq) (# 428333/2016-8, # 313423/2017-2 and # 307277/2014-3) for the financial support and also gratefully acknowledge the support of NVIDIA Corporation with the donation of the Titan Xp GPU used for this research.

REFERENCES

- [1] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.
- [2] R. P. Wildes, "Iris recognition: an emerging biometric technology," *Proceedings of the IEEE*, vol. 85, no. 9, pp. 1348–1363, Sep 1997.
- [3] P. J. Phillips, K. W. Bowyer, P. J. Flynn, X. Liu, and W. T. Scruggs, "The iris challenge evaluation 2005," in *IEEE International Conference on Biometrics: Theory, Applications and Systems*, Sept 2008, pp. 1–8.
- [4] H. Proença and L. A. Alexandre, "Toward covert iris biometric recognition: Experimental results from the NICE contests," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 798–808, 2012.
- [5] M. D. Marsico, M. Nappi, and H. Proença, "Results from MICHE II – Mobile Iris CHallenge Evaluation II," *Pattern Recognition Letters*, vol. 91, pp. 3–10, 2017.
- [6] J. Daugman, "High confidence visual recognition of persons by a test of statistical independence," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1148–1161, 1993.
- [7] E. Luz, G. Moreira, L. A. Z. Junior, and D. Menotti, "Deep periocular representation aiming video surveillance," *Pattern Recognition Letters*, 2017.
- [8] H. Proença and J. C. Neves, "Deep-PRWIS: Periocular Recognition Without the Iris and Sclera Using Deep Learning Frameworks," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 4, pp. 888–896, apr 2018.
- [9] N. Liu, M. Zhang, H. Li, Z. Sun, and T. Tan, "DeepIris: Learning pairwise filter bank for heterogeneous iris verification," *Pattern Recognition Letters*, vol. 82, pp. 154–161, 2016.
- [10] A. Gangwar and A. Joshi, "DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition," in *IEEE Intern. Conference on Image Processing*, 2016, pp. 2301–2305.
- [11] A. S. Al-Waisy, R. Qahwaji, S. Ipson, S. Al-Fahdawi, and T. A. M. Nagem, "A multi-biometric iris recognition system based on a deep learning approach," *Pattern Analysis and Applications*, Oct 2017.

- [12] K. Nguyen, C. Fookes, A. Ross, and S. Sridharan, "Iris recognition with off-the-shelf CNN features: A deep learning perspective," *IEEE Access*, vol. 6, pp. 18 848–18 855, 2018.
- [13] H. Proenca and J. C. Neves, "IRINA: Iris Recognition (Even) in Inaccurately Segmented Data," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 2017-Janua, jul 2017, pp. 6747–6756.
- [14] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Advances in neural information processing systems*, 2014, pp. 3320–3328.
- [15] P. Y. Simard, D. Steinkraus, and J. C. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Seventh International Conference on Document Analysis and Recognition, 2003. Proceedings.*, Aug 2003, pp. 958–963.
- [16] D. Ciregan, U. Meier, and J. Schmidhuber, "Multi-column deep neural networks for image classification," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, June 2012, pp. 3642–3649.
- [17] Q. Wang, X. Zhang, M. Li, X. Dong, Q. Zhou, and Y. Yin, "Adaboost and multi-orientation 2D Gabor-based noisy iris recognition," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 978–983, 2012.
- [18] K. Y. Shin, G. P. Nam, D. S. Jeong, D. H. Cho, B. J. Kang, K. R. Park, and J. Kim, "New iris recognition method for noisy iris images," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 991–999, 2012.
- [19] P. Li, X. Liu, and N. Zhao, "Weighted co-occurrence phase histogram for iris recognition," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 1000–1005, 2012.
- [20] M. D. Marsico, M. Nappi, and D. Riccio, "Noisy iris recognition integrated scheme," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 1006–1011, 2012.
- [21] P. Li and H. Ma, "Iris recognition in non-ideal imaging conditions," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 1012–1018, 2012.
- [22] R. Szewczyk, K. Grabowski, M. Napieralska, W. Sankowski, M. Zubert, and A. Napieralski, "A reliable iris recognition algorithm based on reverse biorthogonal wavelet transform," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 1019 – 1026, 2012.
- [23] T. Tan, X. Zhang, Z. Sun, and H. Zhang, "Noisy iris image matching by using multiple cues," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 970–977, 2012.
- [24] G. Santos and E. Hoyle, "A fusion approach to unconstrained iris recognition," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 984–990, 2012.
- [25] K. W. Bowyer, "The results of the NICE.II iris biometrics competition," *Pattern Recognition Letters*, vol. 33, no. 8, pp. 965–969, 2012.
- [26] Z. He, Z. Sun, T. Tan, X. Qiu, C. Zhong, and W. Dong, "Boosting ordinal features for accurate and fast iris recognition," in *2008 IEEE Conference on Computer Vision and Pattern Recognition*, June 2008, pp. 1–8.
- [27] E. Andersen-Hoppe, C. Rathgeb, and C. Busch, "Combining multiple iris texture features for unconstrained recognition in visible wavelengths," in *Intern. Workshop on Biometrics and Forensics (IWBF)*, 2017, pp. 1–6.
- [28] M. D. Marsico, A. Petrosino, and S. Ricciardi, "Iris recognition through machine learning techniques: A survey," *Pattern Recognition Letters*, vol. 82, pp. 106–115, 2016.
- [29] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 1–9.
- [30] J. Deng, W. Dong, R. Socher, L. J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, June 2009, pp. 248–255.
- [31] J. Daugman, "How iris recognition works," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 21–30, 2004.
- [32] D. Menotti, G. Chiachia, A. Pinto, W. R. Schwartz, H. Pedrini, A. X. Falcão, and A. Rocha, "Deep representations for iris, face, and fingerprint spoofing detection," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 4, pp. 864–879, April 2015.
- [33] Y. Du, T. Bourlai, and J. Dawson, "Automated classification of mislabeled near-infrared left and right iris images using convolutional neural networks," in *IEEE International Conference on Biometrics Theory, Applications and Systems (BTAS)*, Sept 2016, pp. 1–6.
- [34] L. He, H. Li, F. Liu, N. Liu, Z. Sun, and Z. He, "Multi-patch convolution neural network for iris liveness detection," in *2016 IEEE 8th International Conference on Biometrics Theory, Applications and Systems (BTAS)*. IEEE, sep 2016, pp. 1–7.
- [35] E. Severo, R. Laroca, C. S. Bezerra, L. A. Zanlorensi, D. Weingaertner, G. Moreira, and D. Menotti, "A benchmark for iris location and a deep learning detector evaluation," in *2018 International Joint Conference on Neural Networks (IJCNN)*, July 2018, pp. 1–7.
- [36] J. Tapia and C. Aravena, *Gender Classification from NIR Iris Images Using Deep Learning*, 2017, pp. 219–239.
- [37] F. Marra, G. Poggi, C. Sansone, and L. Verdoliva, "A deep learning approach for iris sensor model identification," *Pattern Recognition Letters*, 2017.
- [38] D. R. Lucio, R. Laroca, E. Severo, A. S. Britto Jr., and D. Menotti, "Fully convolutional networks and generative adversarial networks applied to sclera segmentation," *CoRR*, vol. abs/1806.08722, 2018. [Online]. Available: <http://arxiv.org/abs/1806.08722>
- [39] H. Proença, S. Filipe, R. Santos, J. Oliveira, and L. A. Alexandre, "The UBIRIS.v2: A database of visible wavelength iris images captured on-the-move and at-a-distance," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no. 8, pp. 1529–1535, Aug 2010.
- [40] T. Tan, Z. He, and Z. Sun, "Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition," *Image and Vision Computing*, vol. 28, no. 2, pp. 223–230, 2010.
- [41] J. Daugman, "The importance of being random: statistical principles of iris recognition," *Pattern Recognition*, vol. 36, no. 2, pp. 279–291, 2003.
- [42] N. U. Ahmed, S. Cvetkovic, E. H. Siddiqi, A. Nikiforov, and I. Nikiforov, "Using fusion of iris code and periocular biometric for matching visible spectrum iris images captured by smart phone cameras," in *2016 23rd International Conference on Pattern Recognition (ICPR)*. IEEE, dec 2016, pp. 176–180.
- [43] —, "Combining iris and periocular biometric for matching visible spectrum eye images," *Pattern Recognition Letters*, vol. 91, pp. 11–16, may 2017.
- [44] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *British Machine Vision Conference (BMVC)*, Sept 2015, pp. 1–12.
- [45] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, jun 2016, pp. 770–778.
- [46] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "VGGFace2: A dataset for recognising faces across pose and age," *CoRR*, 2017. [Online]. Available: <http://arxiv.org/abs/1710.08092>
- [47] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015, pp. 3431–3440.
- [48] S. Van Dongen and A. J. Enright, "Metric distances derived from cosine similarity and Pearson and Spearman correlations," *CoRR*, 2012. [Online]. Available: <http://arxiv.org/abs/1208.3145>