Deep Learning Utilized for Person Recognition Based on the Biometric Features of the Periocular Region

Leonardo Ferreira Nascimento
Dept. of computation
Instituto Federal de São Paulo
Barretos - SP, Brazil
f.leonardo@aluno.ifsp.edu.br

Jones Mendonça de Souza
Dept. of computation
Instituto Federal de São Paulo
Barretos - SP, Brazil
jonessouza@ifsp.edu.br

***Abstract*—This article proposes the use of deep learning technologies to perform visual biometric recognition. The results obtained by convolutional neural networks trained to perform multi-class classification based on the visual features of the human periocular region are presented and discussed, in addition to being compared with results obtained using pattern recognition for biometric recognition from human iris textures.**

***Index Terms*—Biometry, Artificial Intelligence, Computer Vision, Deep Learning, Periocular Region**

I. Introduction

Biometry has been used within the technology field to perform authentication as an alternative to possession or knowledge-based algorithms [1]. Between the methods of biometric recognition we can cite digital impression, facial, iris, voice, and hand-writing. Besides them, the periocular region of a human's face has shown evidence of possessing features capable of performing such authentication as well [2].

As being intrinsically non-transferable, the biometric approach of authentication provides an advantage over the possession or knowledge-based methods of authentication [1], thus becoming a valuable asset in the field of information security and others.

Even thought that the human iris can be deemed a better suited biometric characteristic for possessing epigenetic patterns [3], an efficient system performing recognition based on the features of the periocular region could be better suited for real-world applications, considering that, for being a larger region of interest, it would be less affected by the challenges described at [4] as acquisition of images of non-cooperative subjects and accurate segmentation of the specific area.

In terms of data, the information in the periocular region of a human's face can be obtained over the process of extraction of patterns between the eyes, the eyelids, the lashes, the skin, and even the iris, considering the quality of the image. Taking in consideration the amount of visual information, the use of convolutional neural networks to process the data and retrieve the information becomes a valid possibility, as it possess the ability to map given images inputs to corresponding categories, or in this case entities, by detecting discriminative abstract feature representations [5].

II. Purpose

The purpose of this research is to, through different databases of images, algorithms and pre-trained artificial intelligence models, be capable to develop, and then analyze,an artificial intelligence model with the ability to recognize multiple subjects based on the features of their periocular region.

III. Materials and Methods

In this research, we utilized two different datasets of images, *CASIA-Iris-V4-Distance* collected by the *Chinese Academy of Sciences' Institute of Automation (CASIA)* [6] and *UBIRIS.V2* [7]. Both datasets were utilized in the same process of training to generate a model of multi-class classification, which consists of utilizing a trained convolutional neural network as a base, and an algorithm of classification, this method is known as Transfer Learning. We utilized two different trained convolutional neural networks: *MobilenetV2.7* [8], and *InceptionV1* [9]. And three different algorithms of multi-class classification: *Naive Bayes, L-BFGS Maximum Entropy,* and *OneVersusAll* utilizing a binary classification *L-BFGS Logistic Regression* algorithm.

A. Datasets

CASIA-Iris-V4-Distance: From this dataset was utilized a subconjunt of 710 images, separated in 142 different subjects, being 5 images from which subject. This dataset contains facial images collected under near infrared illumination or synthesized, taken from a high resolution device in a 3 meters away range. This dataset was segmented through the use of the *Viola-Jones*' object detection algorithm [8] in three divisions, (Fig. 1) bi-periocular region, (Fig. 2) left periocular region, and (Fig. 3) right periocular region, resulting then in the creation of three different datasets, totalizing 2130 images.

UBIRIS.V2: From this dataset was utilized a subconjunt of 930 images, divided equally between (Fig. 4) left and (Fig. 5) right periocular regions. This dataset contains images captured on non-constrained conditions (at-a-distance, on-the-move and on the visible wavelength) of the periocular region of its subjects.

Fig. 1. Image of the bi-periocular region using the CASIA-Iris-V4.

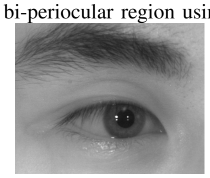

Fig. 2. Image of the left periocular region using the CASIA-Iris-V4.

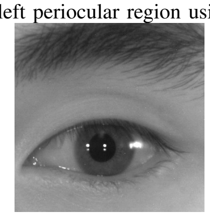

Fig. 3. Image of the right periocular region using the CASIA-Iris-V4.

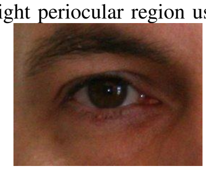

Fig. 4. Image of the left periocular region using the UBIRIS.V2.

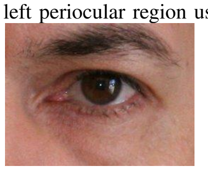

Fig. 5. Image of the right periocular region using the UBIRIS.V2.

B. Technologies

In this research was utilized the *ML.NET* Framework from *Microsoft's .Net* platform of development. Through this development framework we had access to predetermined algorithms to perform multi-class image classification. Such algorithms operate on features computed from image inputs and categorize them in prescribed classes, in this case which class would refer to a person.

Following the description presented in the documentation of the framework, the *L-BFGS Maximum Entropy* algorithm is better suited trainer to deal with large datasets who possess a large number of features than the *Naive Bayes* algorithm, aside from this, both algorithms can perform the required task of classifying an image in a multi-class context directly. The *OneVersusAll* algorithm perform the multi-class classification using a binary classification algorithm as base, which means that the *OneVersusAll*, different from the other two algorithms, train one classifier for every possible class in the context, and then utilizes each classifier to identify the image, resulting in a more complex approach, with a larger consume of computing resources. The chosen binary classification algorithm to be utilized as base for the *OneVersusAll* was the logistic regression version of the *L-BFGS* algorithm.

C. Methods

As proposed in this research, was utilized the methodology of transfer learning to develop new convolutional neural networks with the purpose of performing multi-class classification in order to recognize different subjects.

Transfer Learning is a process that permits a neural network already trained to be adapted and applied in different, and perhaps more specific, tasks [5]. In the architecture of a convolutional neural network, the last layer is responsible for determining the output of the classification based on the vector of features generated by the previous layers [5]. The Transfer Learning method allows us to use such previous layers to generate a new classification layer to our specific task. The use of two different pre-trained models of convolutional neural networks as base should demonstrate whether or not different architectures of convolutional neural network employ an impactful difference in the training of a new model.

IV. Results

A. Discussion

The development involving the combination of the previously mentioned datasets, pre-trained networks and multi-class classification algorithms generated 30 new models. An evaluation of the success rate of every model was done through the process of submitting to classification images of the same dataset that the model was trained on, but that was not used in the process of training. Two images from every subject were submitted for classification, the artificial intelligence should, in case of success, return a value representing the subject which the image corresponds to, or, in case of failure, return a value representing another subject.

Considering that every possibility of union between dataset, pre-trained network and classification algorithm was used to develop a new model, a comparison between the results obtained should demonstrate the influence and effect of each material utilized.

An analysis over the success rate attained over the evaluation of new models shows a certain influence that a constrained environment in the capture of the images exert in the effectiveness of the classification.

The networks developed using the *CASIA-Iris-V4* dataset (Table. I, Table. II, Table. III) produced higher success rates than the one developed using the *UBIRIS.V2* dataset (Table. IV, Table. V), while the CASIA-Iris-V4 lowest success rate reached 79%, the highest success rate reached by a network trained over the UBIRIS.V2 dataset was around 62%.

Another influential aspect was the classification algorithm utilized in the training of the models. The OneVersusAll algorithm obtained the higher success rates. The highest success rate, 95%, was scored by a model trained with the OneVersusAll algorithm over the CASIA-Iris-V4 dataset of the left side's periocular region images. The L-BFGS algorithm obtained an inferior success rate than the OneVersusAll, but superior overall to the Naive Bayes algorithm. The difference between the L-BFGS algorithm to the OneVersusAll algorithm ranged over less than 1 point to 5 points in percentage, while the difference between the Naive Bayes algorithm to the L-BFGS was more sparse, getting even to 14 points in percentage. While the OneVersusAll algorithm scored a better success rate it also had demanded more computing resources, while the L-BFGS and Naive Bayes algorithms trained each of its networks between 2 and 10 minutes, the OneVersusAll algorithm took more than 30 minutes overall to complete the task.

The influence caused by the pre-trained networks is inconclusive, considering that both of them had better success rate than the other in different scenarios of dataset and algorithm.

Reference [11] presents iris at a distance recognition, a different methodology capable of performing the task of person recognition proposed in this research. The CASIA-Iris-V4 and UBIRIS.V2 are among the number of datasets utilized in [11], allowing a comparison between both methods of recognition. The results obtained in [11] using the CASIA-Iris-V4 (Table. VI) dataset are superior to the ones obtained with the UBIRIS.V2 dataset (Table. VII), similar to the results produced in this research. Comparing the best results obtained using the CASIA-Iris-V4 (Table. VI) in [11] with the ones obtained in this research using the same dataset (Table. I,Table. II,Table. III) we can see a higher success rate in the methodology used in this research, being 95.42% of success rate using the InceptionV1 network with the OneVersusAll training algorithm (Table. II), while the best result obtained using iris recognition was 89.88% using images with the highest resolution and the WLD texture descriptor. With the UBIRIS.V2 dataset the methodology used in [9] was superior, with a success rate of 79.88% using images with the highest resolution and the WLD texture descriptor, while the methodology of periocular recognition used in this research reached the best success rate of 62.90%.

B. Tables

TABLE I
TAXA DE SUCESSO OF PERIOCULAR RECOGNITION (CASIA-IRIS-V4, BOTH SIDES)

Pre-trained Network	CASIA-Iris-V4 — Both sides		
	Algorithm	**Quantity of Images**	**Success Rate**
InceptionV1	OneVersusAll	284	91.90%
InceptionV1	L-BFGS	284	90.85%
InceptionV1	Naive Bayes	284	83.10%
MobilenetV2.7	OneVersusAll	284	88.38%
MobilenetV2.7	L-BFGS	284	86.62%
MobilenetV2.7	Naive Bayes	284	80.99%

TABLE II
ACCURACY OF PERIOCULAR RECOGNITION (CASIA-IRIS-V4, LEFT SIDE)

Pre-trained Network	CASIA-Iris-V4 — Left side		
	Algorithm	**Quantity of Images**	**Success Rate**
InceptionV1	OneVersusAll	284	95.42%
InceptionV1	L-BFGS	284	93.66%
InceptionV1	Naive Bayes	284	87.32%
MobilenetV2.7	OneVersusAll	284	91.90%
MobilenetV2.7	L-BFGS	284	90.49%
MobilenetV2.7	Naive Bayes	284	83.80%

TABLE III
ACCURACY OF PERIOCULAR RECOGNITION (CASIA-IRIS-V4, RIGHT SIDE)

Pre-trained Network	CASIA-Iris-V4 — Right side		
	Algorithm	**Quantity of Images**	**Success Rate**
InceptionV1	OneVersusAll	284	89.79%
InceptionV1	L-BFGS	284	86.26%
InceptionV1	Naive Bayes	284	79.92%
MobilenetV2.7	OneVersusAll	284	94.36%
MobilenetV2.7	L-BFGS	284	94.01%
MobilenetV2.7	Naive Bayes	284	86.61%

TABLE IV
ACCURACY OF PERIOCULAR RECOGNITION (UBIRIS.V2, LEFT SIDE)

Pre-trained Network	UBIRIS.V2 — Left side		
	Algorithm	**Quantity of Images**	**Success Rate**
InceptionV1	OneVersusAll	186	62.90%
InceptionV1	L-BFGS	186	58.06%
InceptionV1	Naive Bayes	186	54.84%
MobilenetV2.7	OneVersusAll	186	61.29%
MobilenetV2.7	L-BFGS	186	58.60%
MobilenetV2.7	Naive Bayes	186	44.09%

TABLE V
ACCURACY OF PERIOCULAR RECOGNITION (UBIRIS.V2, RIGHT SIDE)

Pre-trained Network	UBIRIS.V2 — Right side		
	Algorithm	**Quantity of Images**	**Success Rate**
InceptionV1	OneVersusAll	186	62.37%
InceptionV1	L-BFGS	186	57.53%
InceptionV1	Naive Bayes	186	52.69%
MobilenetV2.7	OneVersusAll	186	61.83%
MobilenetV2.7	L-BFGS	186	58.60%
MobilenetV2.7	Naive Bayes	186	52.69%

TABLE VI
ACCURACY OF IRIS RECOGNITION (CASIA-IRIS-V4)

Texture descriptor	Iris Sample Resolution				
	20x240	**30x336**	**40x480**	**50x600**	**60x720**
Daugman	54.08%	54.13%	54.44%	54.78%	55.32%
MBP	58.19%	52.21%	54.40%	54.77%	58.90%
Median-LMP	69.33%	72.14%	77.55%	80.33%	82.77%
MM-LMP	74.08%	77.50%	81.20%	84.55%	88.77%
LMP	58.73%	61.25%	64.99%	66.04%	70.25%
LMP2	56.99%	60.98%	63.99%	64.73%	69.75%
WLD	52.77%	66.14%	71.33%	78.44%	89.88%

Success rate performance of Iris recognition obtained by de Souza JM.

TABLE VII
ACCURACY OF IRIS RECOGNITION (UBIRIS.V2)

Texture descriptor	Iris Sample Resolution				
	20x240	**30x336**	**40x480**	**50x600**	**60x720**
Daugman	54.08%	54.13%	54.44%	54.78%	55.32%
MBP	48.27%	49.04%	49.16%	57.36%	58.27%
Median-LMP	55.80%	60.32%	61.77%	64.88%	68.21%
MM-LMP	60.18%	66.22%	68.63%	72.33%	75.88%
LMP	47.71%	52.16%	53.28%	54.75%	59.27%
LMP2	46.38%	49.80%	51.64%	53.94%	56.70%
WLD	37.48%	40.21%	58.77%	65.14%	79.88%

Success rate performance of Iris recognition obtained by de Souza JM.

V. CONCLUSION

The methods employed in this research demonstrate that the utilization of deep learning artificial intelligence can produce promising results in the task of person recognition based on periocular region features. Still the performance of this technology may be incremented with the use of more complex techniques, such as the development of convolutional neural networks from scratch with the purpose to perform biometric classification, using a larger amount of data with better definition and resolution, and applying different classification algorithms. The quantity of classes that such models are trained to identify also compromise its quality, a smaller scope of classes should help the definition of more accurate discriminant features, a larger number of images in such classes could as well assist in generating more accurate features. The definition of a solid standard on the images that are used in the training can also improve the quality of the model, so images captured by nowadays devices and manipulated to ensure a desired pattern can also improve the success rate in a real life situation of the use of this technology.

Tests with more pre-trained networks are needed to help ascertain if a deeper convolutional neural network could have an impactful effect in biometric tasks.

ACKNOWLEDGMENT

The financial support provided by the Instituto Federal de São Paulo through the PIBIFSP program is gratefully acknowledged.

REFERENCES

[1] Vielhauer, Claus. *Biometric User Authentication for IT Security: From Fundamentals to Handwriting.* 2006th ed, *Springer*, 2005.

[2] Park, U; Ross, A.; Jain, A. K. *Periocular biometrics in the visible spectrum: A feasibility study*. 2009. *IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems.* pp. 1-6.

[3] C. Rathgeb and C. Busch, Eds. *Iris and Periocular Biometric Recognition.* Stevenage, England: *Institution of Engineering and Technology*, 2017.

[4] R. M. Bodade and S. Talbar, *Iris Analysis for Biometric Recognition Systems*, 2014th ed. New Delhi, India: *Springer*, 2014.

[5] Khan, S., Rahmani, H., Shah, S. A. A., & Bennamoun, M. (2018). *A Guide to Convolutional Neural Networks for Computer Vision.* Morgan Claypool.

[6] CASIA-IrisV4. Chinese Academy of Sciences Institute of Automation. http://biometrics.idealteste.org Tech. Rep. 2005.

[7] PROENÇA, H. et al. The UBIRIS.v2: A Database of Visible Wavelength Imagens captured On-The-Move and At-A-Distance. IEEE Trans. PAMI, 2010. v. 32, n. 8, p. 1529-1536. 2010.

[8] Howard, Andrew G. et al. *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*. ArXiv abs/1704.04861. 2017.

[9] Szegedy, Christian et al. *Going deeper with convolutions*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2015.

[10] P. Viola and M. Jones, Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001, pp. I-I, doi: 10.1109/CVPR.2001.990517

[11] Souza, Jones & Gonzaga, Adilson. (2019). *Human iris feature extraction under pupil size variation using local texture descriptors*. Multimedia Tools and Applications. 78. 10.1007/s11042-019-7371-4.