



VISOB 2.0 - The Second International Competition on Mobile Ocular Biometric Recognition

Hoang (Mark) Nguyen¹(✉), Narsi Reddy¹, Ajita Rattani²,
and Reza Derakhshani¹

¹ Department of Computer Science and Electrical Engineering,
University of Missouri at Kansas City, Kansas City, USA
`hdnf39@umsystem.edu`

² Department of Electrical Engineering and Computer Science, Wichita State
University, Wichita, USA
`ajita.rattani@wichita.edu`

Abstract. Following the success of VISOB 1.0 visible light ocular biometrics competition at IEEE ICIP 2016, we organized VISOB 2.0 competition at IEEE WCCI 2020. The aim of VISOB 2.0 competition was to evaluate and compare the performance of ocular biometrics recognition approaches in visible light using (a) stacks of five images captured in burst mode and (b) subject-independent evaluation, where subjects do not overlap between training and testing set. We received three submissions in which the authors developed various deep learning based and texture-analysis based methods. The best results were obtained by a team from Federal University of Parana (Curitiba, Brazil), achieving an Equal Error Rate (EER) of 5.25% in a subject-independent evaluation setting.

1 Introduction

Biometric user verification in mobile devices has all but won the top spot as the user access control method of choice [8, 13]. Biometrics has brought convenience and enhanced security to a wide range of applications such as user login, payments¹, and eCommerce in general². The use of biometrics in mobile devices is termed as mobile biometrics [13].

Thanks to deep learning and advanced camera technologies, mobile face biometrics has come a long way in terms of robustness, accuracy, and user experience. However, given the recent privacy concerns, especially amid the COVID-19 pandemic, and the resulting face-covering mandates, there is an intensified desire for alternate solutions to face recognition [1, 2]. According to a recent 2020 NIST

¹ <https://www.aitrends.com/financial-services/facial-recognition-making-its-way-in-banking/>.

² <https://www.ft.com/content/5d8100b6-ca6e-11e9-af46-b09e8bfe60c0>.

study [9], the presence of face masks could cause face recognition systems to fail up to 50%. Ocular biometrics offers a viable alternative to mobile face recognition given that similar to face, the ocular band can be acquired using the front-facing RGB camera of the mobile device. Ocular biometrics in and of itself has attracted exceeding attention from the research community thanks to its accuracy, security, and robustness against many facial expressions [12, 16]. The ocular regions that have been studied for their biometric utility include the iris [5], conjunctival and episcleral vasculature [4], and the periocular region [7]. Several datasets have been published capturing ocular images in the visible spectrum under various conditions, including UBIRIS [11] (241 subjects), MICHE-I [3] (92 subjects), and VISOB [10]. The last one offers the largest number of subjects (550) captured in mobile environment. Part of this dataset was used for VISOB 1.0 ICIP 2016 ocular biometric recognition competition.

Following the success of our previous VISOB ICIP 2016 competition [14], we organized VISOB 2.0 competition [10] as a part of the IEEE WCCI 2020 conference using a different subset of the VISOB database. The differences between VISOB dataset used in WCCI 2020 compared to ICIP 2016 version are given in Table 1. In VISOB 2.0 competition, we extended the region of interest from the tight eye crop (mainly iris, conjunctival, and episcleral vasculature) to larger periocular (a region encompassing the eye and the surrounding skin). The evaluation protocol for VISOB 2.0 is subject-independent (akin to open-set for identification), in which the subjects in the training and testing set do not overlap. This is compared to the less challenging subject-dependent evaluation used in ICIP VISOB 1.0 competition. More specifically, in VISOB 1.0 the 150 subjects in the testing set overlapped with the 550 identities in the training set; while there are no such overlapping identities between training and testing sets in VISOB 2.0. Further, instead of single frame eye captures of VISOB 1.0, VISOB 2.0 samples are comprised of stacks of five images captured in rapid succession (burst mode), opening the door for multi-frame enhancements.

Table 1. Differences between VISOB 1.0 and VISOB 2.0 competition.

	VISOB 1.0	VISOB 2.0
Devices	iPhone, OPPO, Note 4	OPPO, Note 4
ROI	Tight eye crops	Larger periocular region
Data type	Single frame image	Stack of five images
Train-test identities	Overlapping	Independent
Training set	550 subjects in Visit 1	150 subjects in Visit 1 and 2
Testing set	290 subjects in Visit 2	100 subjects in Visit 1 and 2

We note that *multi-frame* ocular biometrics in the visible spectrum has not attracted much attention in the research community [15], which could be in part due to a lack of public multi-frame datasets, something that VISOB 2.0 strives



Fig. 1. Example eye images from VISOB 2.0, WCCI 2020 competition edition.

to overcome. Single-frame mobile captures from the front-facing “selfie” camera may unexpectedly introduce degradation due to illumination variations, noise, blur, and user to camera distance; all adversely affecting matching performance. One way to mitigate this problem is by capturing multiple frames of the eye in burst-mode, followed by multi-frame image enhancement. Frames may be fused at the input level (e.g., using multi-frame image enhancement and super-resolution techniques) or at the feature or score level for enhanced matching performance (e.g. a multi-match system) (Fig. 1).

2 VISOB 2.0 Dataset and Protocol

VISOB 2.0 Dataset: WCCI 2020 VISOB 2.0 competition VISOB Dataset is publicly available³, and consists of stacks of eye images captured using the burst mode by two mobile devices: Samsung Note 4 and Oppo N1. During the data collection, the volunteers were asked to take their selfie images in two visits, 2 to 4 weeks apart from each other. The selfie-like images were captured with the participant holding the phone naturally, using front-facing camera of the mobile devices under three lighting conditions: daylight, indoor (office) lighting, and dim indoors in two sessions (about 10 to 15 min apart). The ocular burst stacks were cropped from full face frames. The burst sequences were selected if correlation coefficient between the center frame and the remaining four images was greater than 90% (i.e. no excessive motion). We detected the face and eye landmarks using Dlib library [6]. The eye crops were generated such that the width and height of the crop is $2.5\times$ that of the eye’s corner to corner width.

Protocol: VISOB 2.0, WCCI 2020 edition, consists of captures from 150 identities. Both left and right eyes from two visits were provided to the participants. Data characteristics is given in Table 2. Also, we provided images from visit 1 and visit 2 (2–4 weeks apart) under earlier mentioned three lighting conditions in order to keep the focus on the long-term verification and cross-illumination

³ <https://sce.umkc.edu/research-sites/cibit/dataset.html/>.

Table 2. Number of VISOB 2.0 training images provided to the challenge participants.

Device	Lighting condition	Visit 1	Visit 2
Note 4	Office light	4,542	6, 138
	Dim indoors	4,788	6, 158
	Daylight	4,868	6, 148
OPPO	Office light	7,848	10, 546
	Dim indoors	5,138	7, 076
	Daylight	5,314	6, 864

comparisons. No image enhancement was applied to the data so that the participants could perform end-to-end learning to obtain the best fusion of biometrics information and multi-frame image enhancement from the burst of input images. In order to evaluate the submissions according to real-life scenarios, we set up this competition in a subject independent environment. For the competition, the participants were simply asked to submit a model that generates the match score from a pair of images (simple reference-probe comparison). Table 3 shows 18 experiments with 3.6M comparisons across different lighting conditions at the evaluation stage. We used Equal Error Rate (EER), ROC Area Under the Curve (AUC), and Genuine Match Rates (GMR) at 10^{-2} , 10^{-3} , and 10^{-4} False Match Rates (FMR) to evaluate accuracies.

3 Summary of Participants' Algorithms

Department of Informatics, Federal University of Parana (UFPR), Curitiba, PR, Brazil: Zanlorensi et al.'s submitted model is an ensemble of five ResNet-50 models pre-trained on the VGG-Face dataset proposed in [17]. Each ResNet-50 was fine-tuned using a softmax loss through 30 epochs on the periocular images from VISOB 2.0 training subset. The last fully connected layer from the original architecture was removed and replaced by two fully connected layers. The first layer is the feature layer containing 256 neurons, and the last one is the prediction layer consisting of 300 neurons as the number of classes in the training set (left and right eyes from 150 subjects). Eventually, the prediction layer was removed, and the output of the feature layer was taken as the deep feature vector for each input image. For each stack of five images, the five ResNet-50 ensemble generates a combined feature vector of length 1280 (5×256). The authors used cosine distance similarity to generate a match score and compare template-test ocular image pairs.

Bennett University, India: Ritesh Vyas' submission employed hand-crafted features, namely directional threshold local binary patterns (DTLBP) and a wavelet transform for feature extraction. This was the only non-deep learning approach submitted to the competition. The authors used Daubechies, an orthogonal

Table 3. Data distribution for the 18 experiments performed on the test set, as used by the organizers to evaluate the submitted methods.

Device	Experiment number	Enrollment	Verification	# of Comparison
Note 4	1	Dim	Dim	82,322
	2	Dim	Daylight	116,028
	3	Dim	Office	123,168
	4	Daylight	Dim	105,614
	5	Daylight	Daylight	153,416
	6	Daylight	Office	184,512
	7	Office	Dim	99,684
	8	Office	Daylight	143,716
	9	Office	Office	169,890
OPPO	10	Dim	Dim	184,360
	11	Dim	Daylight	190,444
	12	Dim	Office	303,686
	13	Daylight	Dim	178,418
	14	Daylight	Daylight	184,438
	15	Daylight	Office	294,356
	16	Office	Dim	290,082
	17	Office	Daylight	332,696
	18	Office	Office	494,492
Total				3,631,322

Table 4. Details of the algorithm submitted to the IEEE WCCI VISOB 2.0 competition.

Participant	Feature extraction	Matcher/Classifier
Team 1	ResNet-50	Cosine
Team 2	DTLBP	Cosine
Team 3	GoogleNet	Euclidean distance + LSTM

wavelet, to facilitates the multi-resolution analysis. The local texture representation operator captures the unique intensity variations of the periocular image. DTLBP is more robust to noise and is able to extract more distinctive feature representation than the local binary pattern (LBP). Chi-square distance was utilized to compare features from two stacks of images, followed by score normalization.

Anonymous Participant: The authors used a GoogleNet pre-trained on the ImageNet dataset to extract the representation features. Euclidean distance was employed to calculate the similarity between pairs of periocular images.

Following the distance calculation, the scores were used to train Long Short Term Memory (LSTM) model to predict if the pair of images belong to the same individual.

4 Result and Discussion

Table 4 shows the details of the three algorithms submitted to the competition. Experiments were setup as subject independent (open-set-like). All the algorithms consisted of a feature extractor and a similarity-based matcher. The former extracts the feature representation of the image, and the latter computes the match score between two data samples (enrollment and verification). Two out of the three submissions employed deep learning based approaches.

Table 5 shows the EER and AUC of the competition's 18 experiments using Note4 and OPPO N1 challenge data for the three submitted algorithms (note that OPPO N1 has a better camera). Figure 2 shows the average GMRs at different FMRs in 18 experiments. These values are calculated by taking the average of GMRs from the 18 experiments. It can be easily seen that team 1 outperformed the other two teams by a large margin. The best result obtained by team 1 for Note 4 is 5.256% EER and 0.988 AUC for the 9th experiment (office versus office), shown in the result table. For OPPO N1, team 1 achieved the highest performance for dim light versus dim light condition with 6.394% EER and 0.984 AUC. Three experiments with enrollment and verification under the same lighting condition (experiment 10, 14, and 18) generally obtained slightly better performance than the other experiments. This implies cross illumination comparison degrades the performance of the model submitted by team 1.

As shown in Table 5, team 2 achieved the 2nd best place in our competition. Using a similar cosine matcher as team 1, team 2 utilized a non-deep learning based textural feature extractor, DTLBP. The lowest EER for team 2 was 27.05% for Note 4 and 26.208% for OPPO N1 device in the office versus office lighting

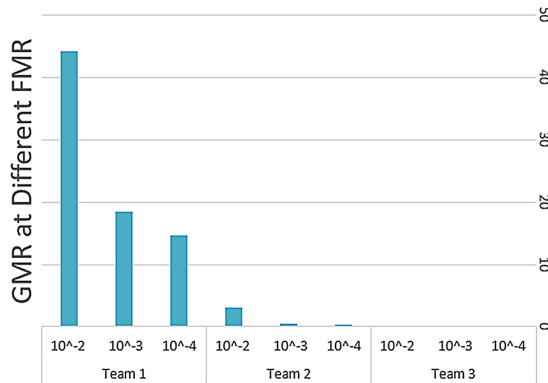


Fig. 2. GMR% at 10^{-2} , 10^{-3} , and 10^{-4} FMR of the three submissions.

Table 5. EER and AUC of the 9 experiments for three submissions, Note 4 device.

	Team 1		Team 2		Team 3	
	AUC	EER (%)	AUC	EER (%)	AUC	EER (%)
1	0.98	7.462	0.715	35.014	0.597	42.074
2	0.952	10.025	0.625	40.468	0.57	44.688
3	0.978	6.659	0.614	42.153	0.583	43.435
4	0.955	11.456	0.615	41.499	0.568	44.41
5	0.971	7.763	0.767	30.679	0.604	40.685
6	0.985	6.722	0.716	34.403	0.585	42.514
7	0.951	12.102	0.605	43.651	0.557	46.085
8	0.968	8.063	0.722	34.309	0.582	42.686
9	0.988	5.256	0.804	27.05	0.629	39.772
10	0.984	6.394	0.732	34.334	0.61	40.301
11	0.961	9.397	0.628	40.362	0.549	44.943
12	0.974	8.082	0.622	40.898	0.568	43.705
13	0.971	8.282	0.623	41.993	0.551	45.411
14	0.965	8.112	0.776	29.697	0.573	42.46
15	0.979	6.672	0.745	31.911	0.548	45.137
16	0.959	9.757	0.603	42.945	0.535	46.679
17	0.961	8.654	0.746	31.785	0.541	45.698
18	0.978	6.487	0.819	26.208	0.581	42.047

setting. However, the model performance degraded significantly for other experiments with EER fluctuating from 30% to 43%. It appears that the non-deep learning features from DTLBP are not as robust against changes in illumination. Team 3’s model did not obtain satisfactory results for any of the experiments.

5 Conclusion

Ocular biometric is becoming an attractive alternative to face recognition in the mobile environment, especially due to occlusion caused by masks worn during the COVID-19 pandemic. We organized the VISOB 2.0 competition at IEEE WCCI 2020 conference to further advance the state-of-the-art in such ocular recognition methods, with a focus on multi-frame captures. We performed a thorough evaluation of three ocular recognition algorithms submitted to our VISOB 2.0 Challenge Competition. VISOB 2.0 dataset consists of stacks of five ocular images captured in burst mode using the front-facing camera from two different smartphones. From the obtained test results, it is obvious that the deep learning approach could obtain better results in our more challenging subject-independent evaluation settings. The comparison across different illumination

settings showed adverse effects on the performance of all three submissions. These results can serve as a reference for future research and development in multi-frame RGB ocular recognition.

Acknowledgement. This work was funded in part by a gift from EyeVerify, Inc. (dba ZOLOZ) an affiliate of Ant Group Co., Ltd., and its affiliates. Dr. Derakhshani is also a consultant for ZOLOZ.

References

1. Carlaw, S.: Impact on biometrics of covid-19. *Biometric Technol. Today* **2020**(4), 8–9 (2020)
2. Damer, N., Grebe, J.H., Chen, C., Boutros, F., Kirchbuchner, F., Kuijper, A.: The effect of wearing a mask on face recognition performance: an exploratory study. arXiv preprint [arXiv:2007.13521](https://arxiv.org/abs/2007.13521) (2020)
3. De Marsico, M., Nappi, M., Riccio, D., Wechsler, H.: Mobile iris challenge evaluation (miche)-i, biometric iris dataset and protocols. *Pattern Recogn. Lett* **57**, 17–23 (2015)
4. Derakhshani, R., Ross, A., Crihalmeanu, S.: A new biometric modality based on conjunctival vasculature. In: *Proceedings of Artificial Neural Networks in Engineering*, pp. 1–8 (2006)
5. Gangwar, A., Joshi, A.: Deepirisnet: deep iris representation with applications in iris recognition and cross-sensor iris recognition. In: *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 2301–2305. IEEE (2016)
6. King, D.E.: Dlib-ml: a machine learning toolkit. *J. Mach. Learn. Res.* **10**, 1755–1758 (2009). <http://dl.acm.org/citation.cfm?id=1577069.1755843>
7. Kumari, P., Seeja, K.: Periocular biometrics: a survey. *J. King Saud Univ.-Comput. Inf. Sci.* **38**, 945–951 (2019)
8. Lovisotto, G., Malik, R., Sluganovic, I., Roeschlin, M., Trueman, P., Martinovic, I.: *Mobile biometrics in financial services: A five factor framework*. University of Oxford, Oxford, UK (2017)
9. Ngan, M.L., Grother, P.J., Hanaoka, K.K.: Ongoing face recognition vendor test (frvt) part 6a: Face recognition accuracy with masks using pre-covid-19 algorithms (2020)
10. Nguyen, H., Reddy, N., Rattani, A., Derakhshani, R.: Visob dataset wcci2020 challenge version (2020). <https://sce.umkc.edu/research-sites/cibit/dataset.html#visob-data-descrip>
11. Proença, H., Alexandre, L.A.: UBIRIS: a noisy iris image database. In: Roli, F., Vitulano, S. (eds.) *ICIAP 2005. LNCS*, vol. 3617, pp. 970–977. Springer, Heidelberg (2005). https://doi.org/10.1007/11553595_119
12. Rattani, A., Derakhshani, R.: Ocular biometrics in the visible spectrum: a survey. *Image Vis. Comput.* **59**, 1–16 (2017)
13. Rattani, A., Derakhshani, R., Ross, A. (eds.): *Selfie Biometrics*. ACVPR. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-26972-2>
14. Rattani, A., Derakhshani, R., Saripalle, S.K., Gottemukkula, V.: Icip 2016 competition on mobile ocular biometric recognition. In: *2016 IEEE International Conference on Image Processing (ICIP)*, pp. 320–324. IEEE (2016)
15. Reddy, N., Fahim Noor, D., Li, Z., Derakhshani, R.: Multi-frame super resolution for ocular biometrics. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops* (2018)

16. Reddy, N., Rattani, A., Derakhshani, R.: Generalizable deep features for ocular biometrics. *Image Vis. Comput.* **103**, 103996 (2020)
17. Zanlorensi, L.A., Lucio, D.R., Junior, A.D.S.B., Proença, H., Menotti, D.: Deep representations for cross-spectral ocular biometrics. *IET Biometrics* **9**(2), 68–77 (2019)