

Is There Any Similarity Between a Person's Left and Right Retina?

Sangeeta Biswas¹ Johan Rohdin² Tomáš Mňuk³ Martin Drahanický⁴

Abstract: It is often argued among biometric researchers that the left and right retinas of the same person are as different as the retinas of two different persons. In this paper we investigate to what extent this is true. We perform experiments where human volunteers are asked to judge whether a pair of the left and right retinal images displayed side-by-side belongs to the same person or two different persons. We also use two similarity measurements, structural similarity (SSIM) and cosine similarity, to do the investigation process automatically. Our experiments show that there is recognizable similarity in the left and right retina of a person. For a verification task done by human volunteers, the average accuracy was 82%. For identification tasks, automatic systems using cosine similarity were correct in up to 57%.

Keywords: Retina, Convolutional Neural Network, Cosine Similarity.

1 Introduction

After the interesting discovery of Dr. Carleton Simon and Dr. Isodore Golstein in 1935, we are aware that the tree structure of retinal blood vessels (RBV) of our eyes is unique [SG35]. This uniqueness is true even for two monozygotic twins [To55]. Unless being affected by severe retinal diseases such as proliferate retinopathy, retinopathy of prematurity, etc., the tree structure of RBV remains unchanged during the lifetime of a person. The external environment cannot effect on it, since its location is inside of our eye [JBP96]. Therefore, it is considered a reliable biometric when we want to ensure high security in an environment.

In the literature, *retina biometrics* generally refers biometrics based solely on the RBV which is a bit misleading. The retina is a neurosensory tissue lining at the back of our each eye. In an RGB colored retinal image the optic disc and macula are spotted along with RBV. Depending on the amount of light entered through the pupil, skin color, quantity of pigments, pathology such as cataracts, retinopathy, etc., retinas of different individuals reflect different colors when they are captured by Fundus cameras. Moreover, like the color of retina, variability can also be noticed in the optic disc and macula. However, they do not provide universal uniqueness as the tree structure of RBV does. Perhaps for this reason, pure retina based or whole retinal RGB image based person authentication studies are almost absent in the literature. Moreover, all RBV based biometric systems found in

¹ Brno University of Technology, Czech Republic, biswas@fit.vutbr.cz and University of Rajshahi, Bangladesh, sangeeta.cse@ru.ac.bd

² Brno University of Technology, Czech Republic, rohdin@fit.vutbr.cz

³ St. Anne's University Hospital Brno, Czech Republic, tomas.mnuk@fnusa.cz

⁴ Brno University of Technology, Czech Republic, drahan@fit.vutbr.cz

the literature are based on only one side RBV. One reason for this could be that the tree structures of the left and right RBV of a person are also unique.

With the help of human observers, Hollingsworth et al., in [Ho11] showed that there is recognizable similarity in the left and right irises of an individual and in the irises of identical twins. To the best of our knowledge it has not been investigated before whether there is any similarity between a person's left and right RBV as well as retinal RGB images. And if there is any similarity, is it measurable? Is it large enough to develop a side-independent person authentication system? These and similar questions are investigated in this paper by the help of human volunteers as well as Convolutional Neural Networks (CNNs).

2 Manual Verification

The first question we have faced is how to find similarity between the left and right retina. Based on which features we can say that both retina belong to the same person? Which part of the retina we need to focus on? To find answers of these questions, we have done manual verification with the help of four human volunteers who did not have much experience about retinal images before. There were two kinds of tests: RGB test and RBV test. One volunteer participated in both RGB and RBV tests and three volunteers participated either in the RGB or in the RBV test. In the RGB test, we showed 120 pairs of RGB retinal images to three human volunteers and in the RBV test we showed 120 pairs of segmented RBV to two human volunteers. RGB retinal images were taken from the EBD_RET data set consisting of retinal images of 107 different people. In this data set, there are at least three images for each side of retina of each person. Therefore, at least three pairs of left and right images can be created while using each image only once. Images were taken using a Fundus Camera Canon CR-1 in the STRaDe, FIT, BUT laboratory environment. There is not any specific alignment of the optic disc and macula, i.e., retinal images of the same retina vary to some extent due to the movement of the optic disc and macula. Among 107 people's retinal images, we have randomly picked 20 people's retina for the RGB test and 20 people's retina for the RBV test. Only four people's retinal images were overlapped in the two tests. By three sets of left and right retinal images of 20 people, even though it was possible to prepare 180 (i.e., $20 \times 3 \times 3$) positive pairs and 3420 (i.e., $20 \times 3 \times 3 \times 19$) negative pairs, we assumed that for human volunteers it would be a quite time consuming, tiring and boring task to give decision about 3600 pairs of images. Therefore, we prepared 60 positive pairs (i.e., left and right images were from the same person) and 60 negative pairs (i.e., left and right images were from two different people). To have as much variability as possible among the pairs, we created three positive pairs per person by using each of the three right and three left images once. For negative pairs, we carefully prepared a table (see Table 1) to ensure that each person can contribute equally to the negative pairs. RBV were segmented from RGB retinal images by a U-Shaped CNN.

As shown in Figure 1, pairs of images were shown side-by-side. Right side images were flipped to make comparison task of human volunteers easier. The tasks of the human volunteers were to determine which pairs of images belong to the same person and which pairs of images are from two different people and click either on the 'Same' or on the

L \ R	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t
a			✓	✓															✓	
b					✓	✓														✓
c	✓						✓	✓												
d	✓								✓	✓										
e		✓									✓	✓								
f		✓											✓	✓						
g			✓												✓	✓				
h			✓														✓	✓		
i				✓															✓	✓
j				✓							✓	✓								
k					✓					✓		✓								
l					✓					✓	✓									
m						✓								✓	✓					
n						✓							✓		✓					
o							✓						✓	✓						
p								✓									✓	✓		
q									✓								✓		✓	
r									✓								✓	✓		
s	✓									✓										✓
t		✓								✓										✓

Tab. 1: Combination of left and right retinal images to make negative pairs. L: Left, R: Right, {a,b,c,...,t}: person's ID.

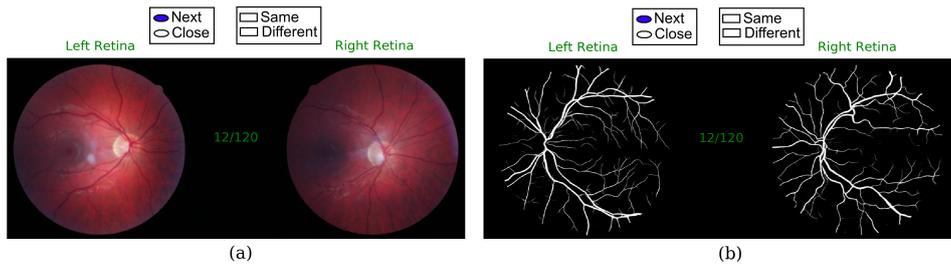


Fig. 1: Interfaces used for manual verification, (a): RGB test and (b): RBV test

'Different' box according to their decision. Even though they were allowed to stop a test any time by clicking on the 'Close' button, all volunteers completed viewing 120 pair of images. Volunteers were instructed that if they are completely unsure about any pair of images they can skip that pair by clicking on the 'Next' button. Four volunteers participated in five separate sessions in five days. We did not give any hint which feature they need to look for to find the similarity. They were free to find features by themselves. They also did not influence each other by their own assumptions. None of them were aware about the true answers.

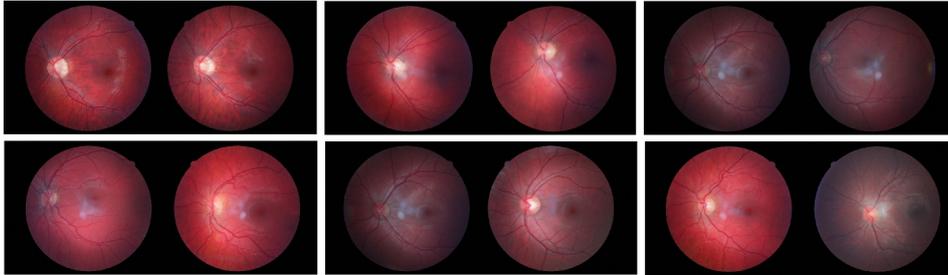


Fig. 2: Pairs of RGB retinal images of the left and right retina agreed by three volunteers. 1st row: three positive pairs recognized correctly by three volunteers and 2nd row: three negative pairs recognized correctly by three volunteers.

Perhaps surprisingly we found that in the RGB test, the three volunteers classified 95-99 out of 120 pairs of RGB images correctly which is well above the result of random guesses. Thirty three out of 60 positive pairs and 35 out of 60 negative pairs were correctly recognized by all three volunteers. There was no negative pair which all three of them wrongly chose as positive pair. Only one positive pair, they all wrongly decided as negative pair. Figure 2 shows three positive and three negative pairs correctly recognized by three volunteers. The volunteers figured out different features to find the similarity between the left and the right retina. Some of them are: intensity of colors, optic disk, shape of RBV, i.e., thickness of RBV, how RBV are entangled, how RBV are branching, the part of RBV located inside the optic disk, boundary of optic disk, texture of retinal background. The RBV test was harder than the RGB test because of the absence of color and optic disk. However, the two volunteers were able to perform better than the random guesses. As shown in Figure 3, by focusing on different features such as the curvature of RBV, how RBV are coming out from the root, how much RBV are spread over, the number of branches specially the number of small branches, etc., the two volunteers were able to recognize correctly the same 28 positive pairs and the same 29 negative pairs. However, their features also mislead them in some cases. Both of them were wrong for 10 positive and 7 negative pairs, respectively. To reduce confusion of human volunteers we need a data set having good quality images with optic discs having the same alignment. If the optic discs are aligned differently in the left and the right retinal images, the orientation of RBV becomes different in the left and right retinal images, for which positive pairs may look like as negative pairs. Underexposed images are problematic for the both RGB and RBV tests, since volunteers cannot see true color and all RBV. Individual performance, shown in Table 2, reveals that the left and the right retinas of a person are not completely different from each other. Both of them bear some common information which is not completely known to us at this moment.

3 Similarity Measure

Four human volunteers without any prior knowledge was able to figure out some features which show some similarity between the left and the right retina of an individual.

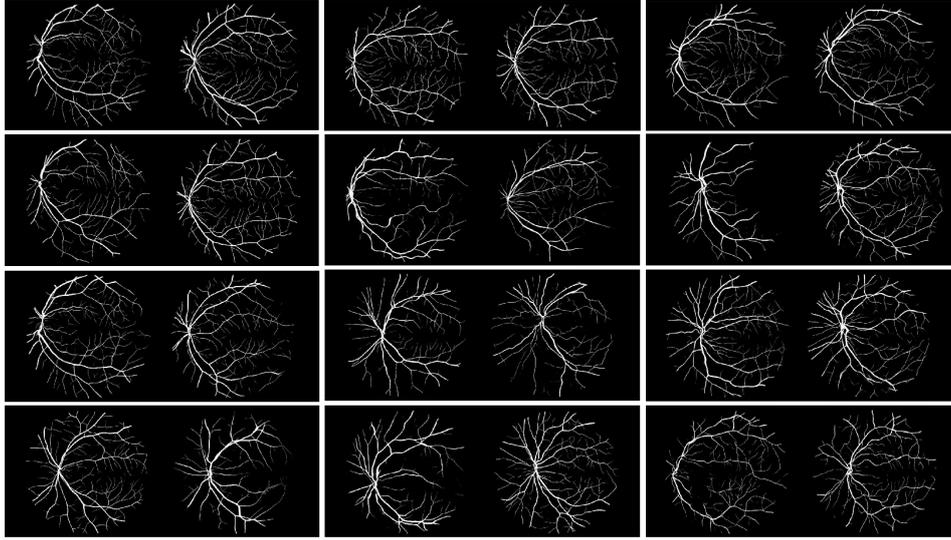


Fig. 3: Pairs of RBV of the left and right retina agreed by two volunteers. 1st row: three positive pairs recognized correctly by two volunteers; 2nd row: three negative pairs recognized correctly by two volunteers; 3rd row: three negative pairs mistakenly recognized as positive pairs by two volunteers and 4th row: three positive pairs mistakenly recognized as negative pairs by two volunteers.

Verifier ID	RGB				RBV		
	1	2	3	Avg.	1	4	Avg.
Accuracy	0.82	0.83	0.79	0.82	0.71	0.63	0.67
Precision	0.80	0.75	1.00	0.85	0.78	0.61	0.69
Recall	0.87	0.98	0.58	0.81	0.58	0.71	0.65
F1 Score	0.83	0.85	0.74	0.81	0.67	0.66	0.66

Tab. 2: Results of manual verification.

Many more features can be considered to figure out that common information. However, for that we need to investigate many more pairs of retinal images which is tiring, boring and error-prone for humans. Therefore, we have prepared two systems which can do this investigation without human. We have considered two popular similarity measurements: structural similarity (SSIM) and cosine similarity.

Structural Similarity (SSIM) [Wa04] is a well accepted similarity measurement between two images. It is based on three comparison measurements: luminance, contrast and structure. It is mainly proposed for predicting image quality of a distorted image comparing with the distortion free uncompressed image. However, because of its third component it also suits to measure structural similarity between two images. The range of SSIM value is $[-1, 1]$. When two images are same there SSIM is 1. When two images are completely different, their SSIM value is -1 .

If we reshape two 2D or 3D retinal images to vectors, then we can compare them by measuring their cosine similarity. However, these vectors will be so long that the cosine similarity will suffer from the curse of high dimensionality. Therefore, it is better to reduce the dimension before measuring the cosine similarity. Comparing to the most popular dimensionality reduction technique, principal components analysis (PCA), we have chosen a non-linear approach using the U shaped CNN (U-Net) [RFB15] which is one of the most popular CNNs in medical image processing. Typically, it is a symmetrical neural network with a middle layer that is substantially narrower than the input and output layers. In this CNN, the output from layers at earlier stages are added to the output of layers at later stages. It performs very well for image-to-image translations specially when the data set is small. The output of the middle layer can be used as *embedding* or *code* of the input.

4 Experimental Setup

We did all implementations using TensorFlow’s Keras API and Python. We used a standard PC with Intel(R) Core(TM) i9-9900K having 8 Cores and 31 GB memory, and with two NVIDIA GeForce GTX 1080 GPUs having 8 GB Memory per GPU.

In order to train U-Net [RFB15] shaped CNNs, we used a publicly available data set named DRIVE [St04]. As a validation set we used another publicly available data set named CHASE_DB1 [Ow11]. For figuring out similarity, we used two private data sets named FNUSA and EBD_RET1 along with CHASE_DB1. Since no information about the identity of the patients is provided, we considered 40 images of DRIVE as unpaired images. Other three data sets have pairs of the left and right retinal images. There is a lot of varieties in the position of optic disc in DRIVE. Optic discs are almost in the center in CHASE_DB1, whereas they are close to the boundary in FNUSA and in different places in EBD_RET1. Images of CHASE_DB1 have deeper pigmentation than the other three data sets. For building EBD_RET1, we selected 24 pair of images, which are not underexposed or overexposed or do not contain any artifacts, from EBD_RET data set. Among 24 pairs of images of EBD_RET1, only three pairs of images were used in manual verification experiments described in Section 2. Table 3 shows some details of the data sets we used in our experiments.

Because of different sizes of different data sets, at first we re-sized all images to 256×256 by bicubic interpolation. Then we re-scaled pixel values of re-sized images to the range of the *sigmoid* activation function $[0, 1]$, since the sigmoid function was used as the activation function of the output layer of the U-Nets. We then flipped only the right-side retina images to align them with the left-side retinas. Except that no other pre-processing was applied to any images.

We trained three U-Nets: UNet1 for segmenting RBV from RGB retinal images, UNet2 for getting *embedding* for RGB retinal images, and UNet3 for getting *embedding* for segmented RBV. All U-Nets had two parts: encoder and decoder (as shown in Figure 4). During training of UNet1 and UNet3, our targets were to minimize the reconstruction error of the RBV images by the decoders of UNet1 and UNet3. On the other hand, during

Database	Type	Pixels	Fundus Camera	Age	# Pairs	Manually Segmented RBV
DRIVE	Public	565×584	Canon CR5 3CCD	25-90	40 (*)	Available
CHASE_DB1	Public	999×960	Hand-held Nidek NM-200-D	10-11	14	Available
FNUSA	Private	3608×3608	Carl Zeiss VISUCAM 524	20-95	68	Not-Av.
EBD_RET1	Private	1008×982	Canon CR-1 Mark II NM	25-32	24	Not-Av.

Tab. 3: Data sets for training U-Nets and checking similarity between the left and right retinas. (*) Images of DRIVE were considered as unpaired.

training of UNet2, our target was to minimize the reconstruction error of the RGB retinal images by the decoder of UNet2. After training, the encoders of UNet2 and UNet3 were used to generate RGB embedding and RBV embedding, respectively, for measuring the cosine similarity.

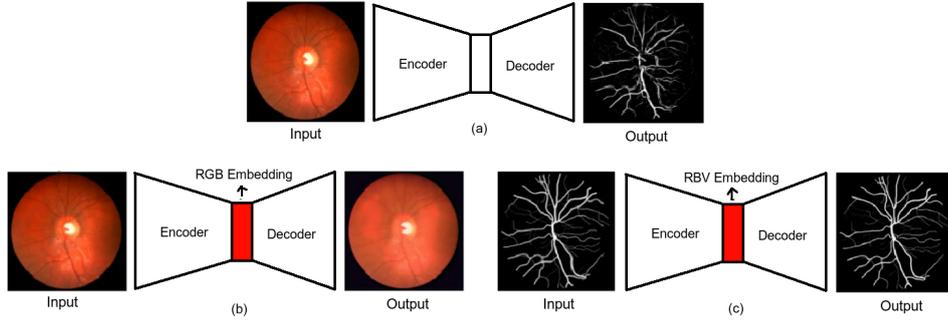


Fig. 4: (a) UNet1 for segmenting RBV from RGB retinal images, (b) UNet2 for getting *embeddings* for RGB retinal images, and (c) UNet3 for getting *embeddings* for segmented RBV.

All these three networks had the same architecture and settings except for different number of channels in the input and output layers. Figure 5 shows the model architecture for 256×256 sized images. We set $input_ch = 3$ and $output_ch = 1$ for UNet1, $input_ch = 3$ and $output_ch = 3$ for UNet2 and $input_ch = 1$ and $output_ch = 1$ for UNet3. We also set mean-squared-error as the loss function; RMSProp [HSS] with a learning rate of 0.001 as the optimizer. We set $mini_batch_size = 8$. We used exponential linear unit (ELU) [CUH16] as the activation function for all convolutional layers except the last layer of the decoder. In the last layer of the decoder *sigmoid* function was used. We set $stride = 1$, $kernel_size = 3$, and $padding = same$ for all convolutional layers. For all convolutional and transposed convolutional layers, we set $kernel_initializer = he_normal$. The number of filters was increased from 16 to 256 for the encoder and decreased from 128 to 16 for the decoder. For all other settings, we used the default values of TensorFlow's Keras API. From the output layer of the encoder part, we got *code* for retinal images and

segmented RBV. Since CHASE_DB1 also has manually segmented RBV, we used it to tune the number of iterations.

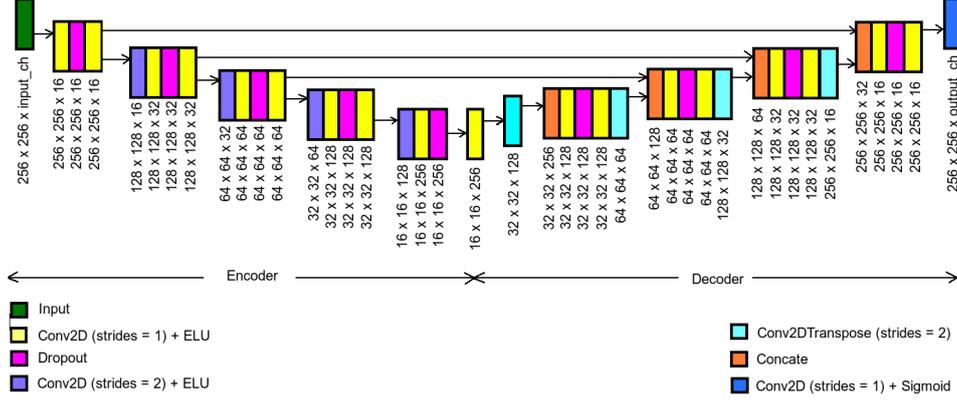


Fig. 5: Architecture of U-Net. Vertical text shows the output shape of the corresponding layer.

As shown in Figure 6, UNet1 segmented RBV quite resemble with the manually segmented RBV of the training set, DRIVE. However, the segmented RBV become a bit noisy for the validation set, CHASE_DB1. Since it gave us a clear idea about the tree structure of RBV, we were satisfied with the performance of UNet1. Getting rid of the noise and having clear RBV by UNet1, we put in our future work list.

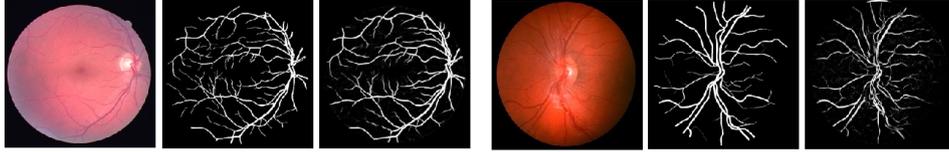


Fig. 6: RGB retinal image and segmented blood vessels. 1st col: sample image from DRIVE, 2nd col: manually segmented BV, 3rd col: generated by UNet1, 4th col: sample image from CHASE_DB1, 5th col: manually segmented, 6th col: generated by UNet1.

5 Results and Analysis

We did experiments at first by comparing each left retinal image with all right retinal images and then checking whether the similarity score for the correct right retinal image is the highest or among the top 2 or 3 largest scores. We did four experiments: SSIM_RBV and Cos_RBV for RBV and SSIM_RGB and Cos_RGB for RGB retinal images. For RBV based experiments, we generated RBV by passing RGB retinal images through UNet1. Even though CHASE_DB1 has manually segmented RBV, we generated RBV for it too in order to keep consistency for the RBV based experiments. For SSIM based experiments we simply calculated SSIM between the two RGB or RBV images. For the cosine distance based experiments, at first we generated RGB embeddings and RBV embeddings from the encoders of UNet2 and UNet3, respectively. Then we compared two embeddings of the left and right retinal images using the cosine distance.

The results are shown in Table 4. Note that the probability to retrieve the correct right retina by chance is $1/14 = 0.07$ for CHASE_DB1, $1/68 = 0.01$ for FNUSA and $1/24 = 0.04$ for EBD_RET1. Except SSIM_BV, the overall trend is that the probability to retrieve the correct right retina for a left retina is much more than a chance. It indicates that the left and right retinas of a person have more similarity than the left and right retinas from two different persons. Note that these results are not comparable with the results of manual verification tasks, since they are two different kinds of tasks: one is identification and the other is verification. For training an automatic verifier instead of using cosine distance, we needed a larger data set, with many pairs of images, which we did not have. We noticed

	CHASE_DB1			FNUSA			EBD_RET1		
	Top1	Top2	Top3	Top1	Top2	Top3	Top1	Top2	Top3
SSIM.RBV	0.07	0.14	0.21	0.03	0.04	0.09	0.12	0.21	0.21
SSIM.RGB	0.29	0.43	0.64	0.26	0.34	0.34	0.37	0.42	0.50
Cos.RBV	0.29	0.36	0.43	0.19	0.32	0.40	0.37	0.54	0.54
Cos.RGB	0.57	0.71	0.79	0.46	0.57	0.69	0.29	0.5	0.62

Tab. 4: The probability of retrieving the correct right retina for a left retina by our four approaches.

that the cosine distance based approach was much better than the SSIM based one and that the RGB input was much easier to identify than the RBV input. When the left and right retinal images of a person had the same color and their optic discs positioned almost in the same place in the retina, the identification task of all identifiers became easier (see Figure 7). The violation of any of these two conditions might confuse any approaches. As shown in Figure 8 and Fig. 9, all approaches failed when there was a large mismatch in color or when the optic disc was in a position in the left retina different from the right one or when any part of an image was underexposed. This problem can be solved to some extent by collecting a more carefully designed database. As shown in Fig. 9, the top scorers may vary from approaches to approaches. This suggests that fusion of the four approaches may improve the results further. This we have, however, left for future work.

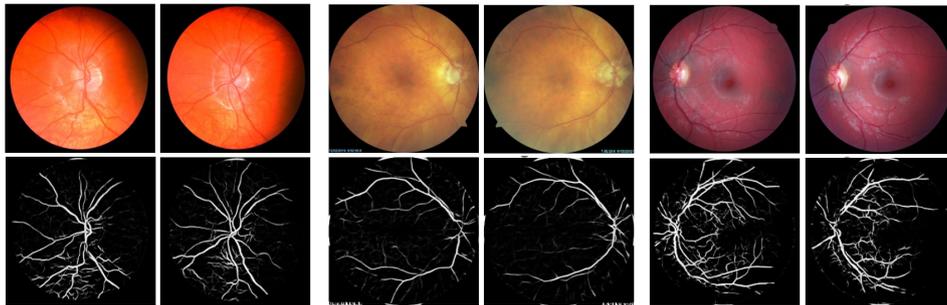


Fig. 7: Left and right retinas having strong similarity claimed by all of our four approaches. 1st & 2nd cols: a pair from CHASE_DB1, 3rd & 4th cols: a pair from FNUSA and 5th & 6th cols: a pair from EBD_RET1.

To some extent the color of the retinal images depends on biological factors such as skin color, quantity of pigments and age. Generally, lighter skinned people's retinas are more of a reddish-orange color, whereas darker skinned people's retinas are more of a darker-

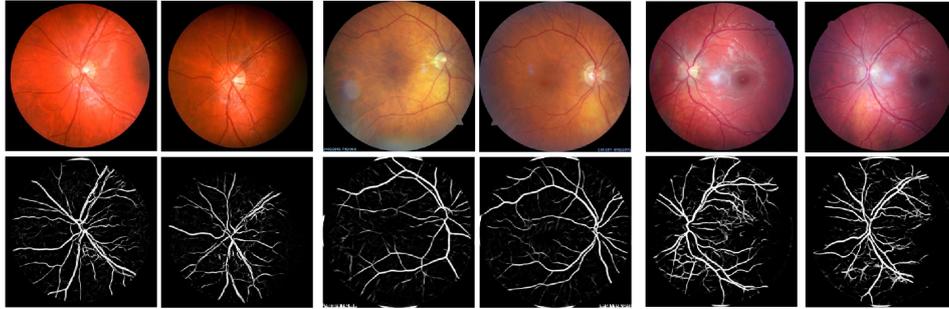


Fig. 8: Pairs of left and right retinas failed to be figured out by all of our four approaches. 1st & 2nd cols: a pair from CHASE_DB1, 3rd & 4th cols: a pair from FNUSA and 5th & 6th cols: a pair from EBD_RET1.

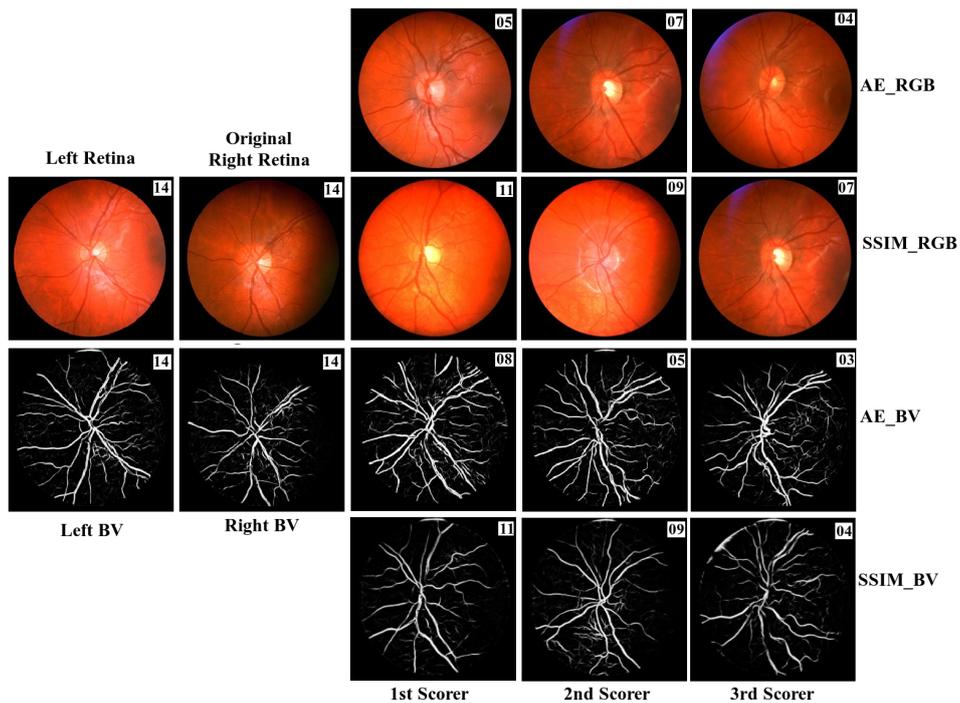


Fig. 9: A left with its original right retina along with 3 top scorer right retinas. Because the right part of the right retinal image was underexposed, our four approaches failed to recognize it.

orange color [BLE03]. Different pathology such as cataracts, retinopathy, etc., also have effect on retina's color and texture. Moreover, retina's color to some extent also depends on *session factors* such as illumination conditions etc. In all the explored databases, the right and left retina image of an individual were collected in one session (i.e. right after each other) whereas not all individuals had their session on the same day. Thus it can-

not be excluded that the good results for the RGB images partly should be attributed to session factors. To properly evaluate this, a more carefully designed database needs to be collected. Nevertheless, the result for the RBV images (for which session factors should be negligible) with the cosine distance is clearly better than chance.

6 Discussion

Even though our experiments show evidence that there is recognizable similarity between a person's left and right retina, since we did experiments on three small data sets, we are still not in the position to claim anything firmly. For that we need bigger data sets and better automatic systems. However, we believe our findings encourage future research in some new directions.

A side independent retina based person authentication system is a system where one side retina can be used to access a system which is developed for the opposite side retina. Maybe this kind of system does not add any extra point to security. However, it might increase user flexibility. Our findings also highlights both risks and opportunities of other two-sided retinal image processing applications. The common belief among the biometric researchers is that two side retinas are unique. If it is completely true, then we could naively assume that an authentication system using both side retina is *two times stronger* than an authentication system using one side retina. However, our finding goes against this naive assumption. Since both of our retinas posses some similarity, an authentication system using two side retinas will not be two times stronger than an authentication system using one side retina.

A potential application of knowing the expected similarities of the left and the right retinas could be to detect anomalies in one of them. However, for this we need a more complete understanding of the similarity between the left and right retinas.

7 Conclusion

Contrary to the common belief among biometric researchers, we have shown that there is a similarity between the left and right retinal images of an individual. We let human volunteers try to recognize whether a left and a right retinal image are from the same individual or not. Their accuracy ranged from 95 to 99 correct classifications out of 120 image pairs. Further, we showed that automatic systems could, given the left retina image of an individual, identify the correct right retinal image in up to 57% of the cases depending on the evaluation data and pre-processing. In this work we used simple automatic systems as a proof of concept and future work will include the development of better automatic systems.

Acknowledgement

This work received funding from the Ministry of Education, Youth and Sports, Czech Republic (Project No: 0008371 - International mobility of researchers at the Brno University of Technology), and the European Union's Horizon 2020 Research and Innovation Programme under the Marie Skłodowska-Curie Actions co-financed by the South Moravian Region under grant agreement No. 665860. This work was also supported by the Ministry of Education, Youth and Sports, Czech Republic, from the National Programme of Sustainability (NPU II) project "IT4Innovations Excellence in Science - LQ1602 "

References

- [BLE03] Blake, C. Richard; Lai, Wico W.; Edward, Deepak P.: Racial and Ethnic Differences in Ocular Anatomy. *International Ophthalmology Clinics*, 43(4):9–25, October 2003.
- [CUH16] Clevert, Djork-Arné; Unterthiner, Thomas; Hochreiter, Sepp: Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). In: 4th International Conference on Learning Representations (ICLR). 2016.
- [Ho11] Hollingsworth, Karen; Bowyer, Kevin W.; Lagree, Stephen; Fenker, Samuel P.; Flynn, Patrick J.: Genetically Identical Irises Have Texture Similarity That is Not Detected by Iris Biometrics. *Computer Vision and Image Understanding*, 115(11):1493–1502, November 2011.
- [HSS] Hinton, Geoffrey; Srivastava, Nitish; Swersky, Kevin: , *Neural Networks for Machine Learning Lecture 6a: Overview of mini-batch gradient descent*.
- [JBP96] Jain, Anil K.; Bolle, Ruud; Pankanti, Sharath, eds. *Retina Identification*. In (Jain, Anil K.; Bolle, Ruud; Pankanti, Sharath, eds): *Biometrics: Personal Identification in Networked Society*. Springer US, pp. 123–141, 1996.
- [Ow11] Owen, Christopher G; Rudnicka, Alicja R; Nightingale, Claire M; Mullen, Robert; Barman, Sarah A; Sattar, Naveed; Cook, Derek G; Whincup, Peter H: Retinal arteriolar tortuosity and cardiovascular risk factors in a multi-ethnic population study of 10-year-old children; The child heart and health study in England (CHASE). *Arterioscler Thromb Vasc Biol*, 31(8):1933–1938, 2011.
- [RFB15] Ronneberger, Olaf; Fischer, Philipp; Brox, Thomas: U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, LNCS, 9351:234–241, 2015.
- [SG35] Simon, Carleton; Golstein, Isodore: A New Scientific Method of Identification. *New York State Journal of Medicine*, 35(18):901–906, 1935.
- [St04] Staal, J.J.; Abramoff, M.D.; Niemeijer, M.; Viergever, M.A.; van Ginneken, B.: Ridge based vessel segmentation in color images of the retina. *IEEE Transactions on Medical Imaging*, 23(4):501–509, 2004.
- [To55] Tower, Paul: The Fundus Oculi in Monozygotic Twins: Report of Six Pairs of Identical Twins. *Archives of Ophthalmology*, 54:225–239, 1955.
- [Wa04] Wang, Zhou; Bovik, Alan Conrad; Sheikh, Hamid Rahim; Simoncelli, Eero P.: Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, April 2004.