# Suitable Embedding to Find Similarity Between Left and Right Retinas of a Person

Sangeeta Biswas
*Dept. of Intelligent Systems*
*Brno University of Technology*
Brno, Czech Republic
biswas@fit.vutbr.cz

Johan Rohdin
*Dept. of Computer Graphics and Multimedia*
*Brno University of Technology*
Brno, Czech Republic
rohdin@fit.vutbr.cz

Martin Drahansky
*Dept. of Intelligent Systems*
*Brno University of Technology*
Brno, Czech Republic
drahan@fit.vutbr.cz

*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. However, in our previous work we showed that by looking at retinal images human volunteers can find some similarity and tell whether a pair of the left and right retinal images belongs to a single person or to two different persons. We also showed that using cosine similarity between the embedded vectors of two retinal images it is possible to find the correct right retinal image for a left retinal image more than what could be expected by chance. In this paper we investigate which type of embedding can give us better cosine similarity measurement between the left and right retinal images of a person.

## I. Introduction

The retina is a neurosensory tissue lining at the back of our each eye. 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 appear in different colors when they are captured by Fundus cameras. In an RGB colored retinal image, the optic disc, macula and retinal blood vessels (RBV) are spotted. Like the color of retina, variability can also be noticed in optic disc, macula and RBV. However, optic disc, macula and retina iteslf do not provide universal uniqueness as the tree structure of RBV does. 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 [1]. This uniqueness is true even for two monozygotic twins [2]. 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 it, since its location is inside of our eye [3]. Therefore, it is considered a reliable biometric when we want to ensure high security in an environment.

Contrary to the common belief among biometric researchers, we have shown in [4] that there is a recognizable similarity between the left and right retinas of an individual. We performed experiments where human volunteers were asked to judge whether a pair of the left and right retinal images displayed side-by-side belongs to the same person or to two different persons. For RGB retinal imag ets, by considering intensity of colors, optic disc, shape of RBV (i.e.,

thickness of RBV), how RBV are entangled, how RBV are branching, the part of RBV located inside the optic disc, boundary of optic disc, texture of retinal background, three human volunteers could correctly classify a pair of RGB retinal images as being from a single person or from two different persons in on average 82% of the cases. For RBV images, by focusing on features such as the curvature of RBV, how RBV are coming out from the root, the number of branches (especially the number of small branches), etc., two volunteers were on correct in on average 67% of the cases.

We also used two similarity measurements, structural similarity (SSIM) [5] and cosine similarity, to do the investigation process automatically. We noticed 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. For the cosine distance based experiments, at first we generated RGB embeddings and RBV embeddings from the encoders of U-shaped CNNs [6] trained for generating RGB retinal images and RBV images. Then we compared two embeddings of the left and right retinal images using the cosine similarity. We found that using the cosine similarity we can identify the correct right retinal image in up to 57% of the cases depending on the evaluation data and pre-processing.

This paper is an extension of our previous work described in [4]. Here we investigate which type of embedding can give us better cosine similarity measurement between the left and right retinal images of a person.

## II. Embedding

How to find similarity between the left and right retinas of a person? To answer this question we assume that if the left and right retinas of a person is similar, then the pair of the left and right retinal images will have lower cosine distance or higher cosine similarity compared to the pairs made by taking the left and right retinal images from two different persons. If we reshape two 2D RBV images or 3D RGB retinal images to vectors, then we can easily measure 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 apply an embedding technique before measuring the cosine similarity. Embedding is a mapping of a high dimensional variable into

a low dimensional variable. Figuring out which embedding is appropriate for our data sets is the purpose of this paper. There are many linear and non-linear embedding techniques. We have chosen only three non neural network based embedding techniques and two neural network based techniques.

Among the non neural network based embedding, we have chosen the most popular linear embedding named *principal component analysis* (PCA) [7]–[9] and two non-linear embedding: *locally linear embedding* (LLE) [10] and *isometric mapping* (Isomap) [11]. PCA considers the directions with the largest variances to be the most important. Therefore, it finds a low-dimensional embedding that best preserves the variance of the data points as measured in the high-dimensional input space. In PCA at first data points are centered and scaled to unit variance. Then the eigenvectors and eigenvalues of the covariance matrix are calculated. After that data points are projected on to the eigenvectors which have higher eigenvalues. LLE begins by detecting neighborhood for each data point, $X_i$. The neighborhood of $X_i$ can be a set of points that lay within some radius $\epsilon$ or the K-nearest neighbors of $X_i$. After detecting neighborhood, a set of weights, $W_{ij}$, is computed for $X_i$ so that $X_i$ can be described as a linear combination of its neighbors. At last, an eigenvector-based optimization technique is used to find the low-dimensional embedding, $Y_i$ of $X_i$, such that $Y_i$ is still described with the same linear combination of its neighbors (i.e., $Y_i$ can be reconstructed by its neighbors using $W_{ij}$). The first step of Isomap is similar to LLE. At first, neighborhood of $X_i$ is detected. After that a star graph is constructed for $X_i$ by keeping $X_i$ as the center vertex and assigning $d_X(ij)$ (i.e., the distance between $X_i$ and its neighbor, $X_j$, measured either in the standard Euclidean metric or in some domain-specific metric) to the weight of edge, $e_{ij}$. All star graphs in together form a neighborhood graph $G$. After forming $G$, the geodesic distances, $d_G(ij)$, between all the connected vertices in $G$ are estimated by computing the shortest path between any two vertices of $G$ using the Floyd-Warshall algorithm. Finally, an eigenvector-based optimization technique is applied on the matrix consisted of $d_G(ij)$ to find the low-dimensional embedding. LLE preserves the local geometry of the data whereas Isomap preserves the global structure of the data.

Along with the above linear and non-linear embedding techniques, we have also chosen the *encoder* of an autoencoder (AE) or a *U-Net* [6] as a neural network based non-linear embedding technique. Typically, AEs are symmetrical neural networks with a middle layer that is substantially narrower than the input and output layers. The narrow middle layer is known as the *bottleneck layer*. AEs can be seen as having two parts: an encoder (from the input layer to the bottleneck layer) and a decoder (from the bottleneck layer to the output layer). Encoder maps the input into the *embedding*, and decoder maps the *embedding* to a reconstruction of the original input. U-Net is, a U shaped convolutional neural network (CNN), one of the most popular CNNs in medical image processing. Like AE, it is also a symmetrical neural network with two parts: an encoder and a decoder. However, in this CNN, the output from layers of the encoder are added to the output of layers of the decoder. These are called skip connections which are absent in general AEs. U-Net performs very well for image-to-image translations specially when the data set is small.

## III. EXPERIMENTAL SETUP

For non neural network based embedding techniques we used scikit-learn [12]. We did all implementations of neural networks using TensorFlow's Keras API 2.1.6-tf 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.

### A. Data Sets

In order to train CNNs, we used four publicly available data sets named DRIVE [13], HRF [14], MESSIDOR [15] and STARE [16]. Since no information about the identity of the patients is provided, we considered images of these four data sets as unpaired images. There is a lot of varieties in the position of optic disc in these four data sets. All data sets except Messidor have manually segmented blood vessels.

As a validation set we used another publicly available data set named CHASE_DB1 [17]. For figuring out similarity, we used two private data sets named FNUSA and EBD_RET1 along with CHASE_DB1. All of these three data sets have pairs of the left and right retinal images. Except CHASE_DB1, other two data sets do not have manually segmented blood vessels. Optic discs are almost in the center in CHASE_DB1, whereas they are close to the boundary in the FNUSA and in different places in the EBD_RET1. Images of the CHASE_DB1 have deeper pigmentation than the other three data sets. The EBD_RET1 is a subset of the EBD_RET data set belonging to the STRaDe group of Faculty of Information Technology (FIT), Brno University of Technology (BUT), Czech Republic. Images of the EBD_RET data set were taken from 110 students and faculty members of BUT in the STRaDe, FIT, BUT laboratory environment. In this data set, 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. For building the EBD_RET1, we selected 24 pairs of left and right retinal images, from the EBD_RET data set, which are not underexposed or overexposed or do not contain any artifacts and optic discs have similar alignment. Images of the FNUSA data set were taken from the patients visited St. Anne's University Hospital Brno, Czech Republic. Table I and Table II show 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 $256 \times 256$ 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-Net shaped CNNs. 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 the images.

| Database | Pixels | Fundus Camera | Age | # Images |
|----------|--------|---------------|-----|----------|
| DRIVE | $565 \times 584$ | Canon CR5 3CCD | 25-90 | 40 |
| HRF | $3504 \times 2336$ | Canon CF-60UVi | * | 45 |
| MESSIDOR | $1440 \times 960$ $2240 \times 1488$ $2304 \times 1536$ | Topcon TRC NW6 3CCD | * | 1187 |
| STARE | $700 \times 605$ | TopCon TRV-50 | * | 20 |

| Database | Pixels | Fundus Camera | Age | # Pairs |
|----------|--------|---------------|-----|---------|
| CHASE_DB1 | $999 \times 960$ | Hand-held Nidek NM-200-D | 10-11 | 14 |
| FNUSA | $3608 \times 3608$ | Carl Zeiss VISUCAM 524 | 20-95 | 68 |
| EBD_RET1 | $1008 \times 982$ | Canon CR-1 Mark II NM | 25-32 | 24 |

## B. Network Architecture

We trained a U shaped CNN, UNet1, for segmenting RBV from RGB retinal images. Figure 1 shows the model architecture of UNet1 for $256 \times 256$ sized images. We set $input\_ch = 3$ and $output\_ch = 1$; *mean-squared-error* (mse) as the loss function; RMSProp [18] with a learning rate of $0.001$ as the optimizer, and $mini\_batch\_size = 8$. We used exponential linear unit (ELU) [19] 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 $strides = 1$, $kernel\_size = 3$, and $padding = same$ for all convolutional layers. For all convolutional and transposed convolutional layers, we set $kernel\_initializer = he\_normal$. For all other settings, we used the default values of TensorFlow's Keras API 2.1.6-tf. We used RGB retinal images and manually segmented RBV of DRIVE, HRF and STARE to train UNet1. Since CHASE_DB1 has manually segmented RBV along with RGB retinal images, we used it to tune the number of iterations of UNet1.

During training of UNet1 our targets was to minimize the reconstruction error of the RBV images by the decoders of UNet1. By combining three data sets, we trained five UNet1: UNet1_DRIVE, UNet1_DRIVE+HRF, UNet1_DRIVE+STARE, UNet1_HRF+STARE and UNet1_DRIVE+HRF+STARE. The mse of segmenting RBV from 28 RGB retinal images of CHASE_DB1 were 0.55, 0.51, 0.47, 0.47, 0.43 by UNet1_DRIVE, UNet1_DRIVE+HRF, UNet1_DRIVE+STARE, UNet1_HRF+STARE and UNet1_DRIVE+HRF+STARE, respectively. Since UNet1_DRIVE+HRF+STARE gave the minimum mse and slightly brighter segmented RBV for the images of CHASE_DB1 than the other UNet1s (as shown in Fig. 2), therefore, we kept it for segmenting RBV of RGB retinal

images of the EBD_RET1 and FNUSA data sets.

In order to get *embedding*, we trained one kind of U-shaped CNN, UNet2, and two kinds of autoencoders (AEs), AE1 and AE2. The architecture of UNet2 was exactly the same as the architecture of UNet1. As shown in the Fig. 3, the architecture of AE1 is the same as UNet1 except it does not have the skip connections like UNet1. As shown in Fig. 4, the architecture of AE2 is almost the same as AE1 except it has few layers than the AE1. Parameter settings of UNet2, AE1 and AE2 were the same as UNet1.

Both AEs and UNet2 were trained for both the RGB retinal images and for RBV images using the images of MESSIDOR. CHASE_DB1 was used to decide the number of iterations. Instead of targeting to achieve the minimum reconstruction error, we used accuracy to decide the number of iterations. For the RGB retinal images, we set $input\_ch = 3$ and $output\_ch = 3$ and for the RBV images and $input\_ch = 1$ and $output\_ch = 1$ for UNet2, AE1 and AE2. RBV images were segmented from RGB retinal images of MESSIDOR by using UNet1_DRIVE+HRF+STARE. Even though CHASE_DB1 has manually segmented RBV, we generated segmented RBV for it by UNet1_DRIVE+HRF+STARE to keep consistency. Both AEs and UNet2 had two parts: encoder and decoder. However, after training we kept only the encoder part to get the RGB and RBV *embedding* from the images of EBD_RET1 and FNUSA data sets.

## IV. RESULTS AND ANALYSIS

The results are shown in Table III. 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. The overall trend is that the probability to retrieve the correct right retina for a left retina is much more than by 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. Neural network based embeddings out-performs non-neural network based embeddings. AE2 based embedding performs slightly better than the UNet2 and AE1 based embeddings. Figuring out an appropriate architecture for the autoencoder can be our future work since it may improve the performance of autoencoder based embedding. Non-neural network based non-linear embeddings, i.e., LLE and Isomap, do not perform better than the linear embedding, PCA. Moreover, they take longer time than PCA.

To some extent the presence of color and the visibility of optic disc help the RGB retinal images have better results than the RBV images. 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-orange color [20]. 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
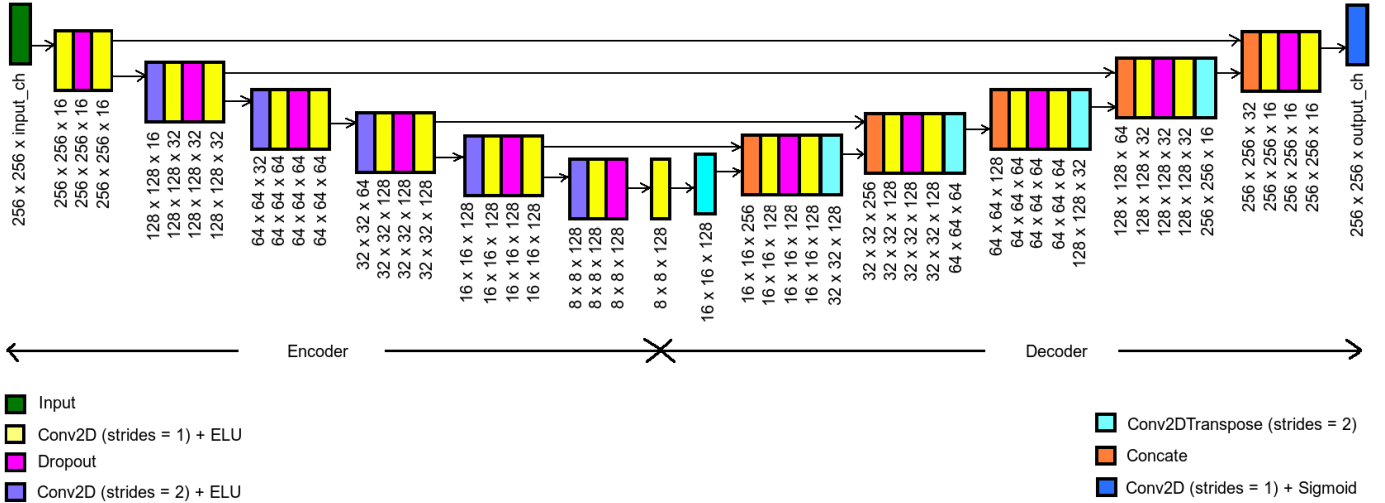
Fig. 1. Architecture of U-Nets (i.e., UNet1 and Unet2). Vertical text shows the output shape of the corresponding layer.



(a) Manually segmented RBV    (b) DRIVE    (c) DRIVE + HRF    (d) DRIVE + STARE    (e) HRF +STARE    (f) DRIVE +HRF + STARE
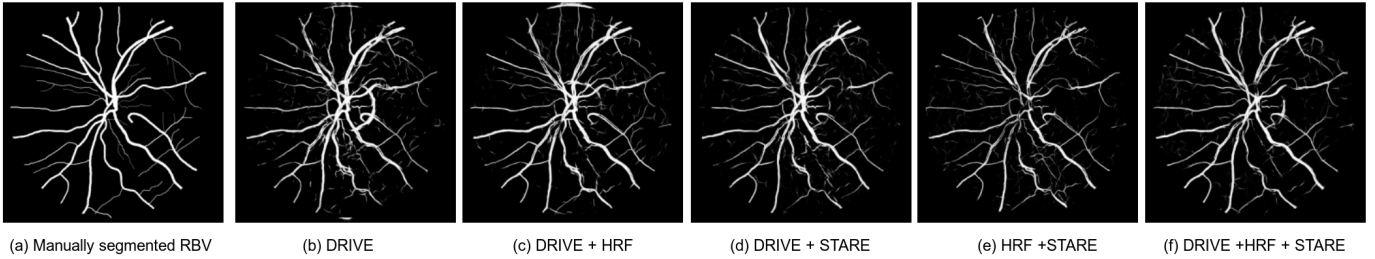
Fig. 2. RBV, of an RGB retinal image of CHASE_DB1, segmented by UNet1 trained by different data sets.
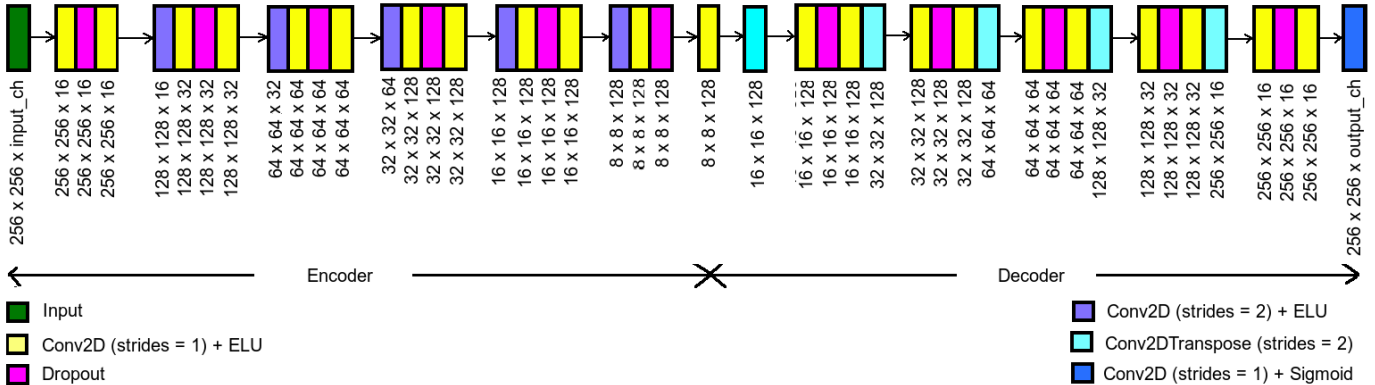


Fig. 3. Architecture of AE1. Vertical text shows the output shape of the corresponding layer.

each other) whereas not all individuals had their session on the same day. Thus it cannot 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 by chance.

As shown in Fig. 5, when the left and right retinal images of a person have the same color or their optic discs positioned almost in the same place in the retina, the identification task of all identifiers becomes easier for RGB retinal images. The

violation of any of these two conditions may confuse any approaches. For RBV images, the position of the root of RBV plays an important role in the identification task (see Fig. 6).

## V. CONCLUSIONS

Contrary to the common belief among biometric researchers, we have shown that there is a recognizable similarity between the left and right retinal images of an individual. Therefore, it is possible to find the correct right retinal image for a left retinal image by measuring the cosine similarity between the two retinal images with higher probability than by
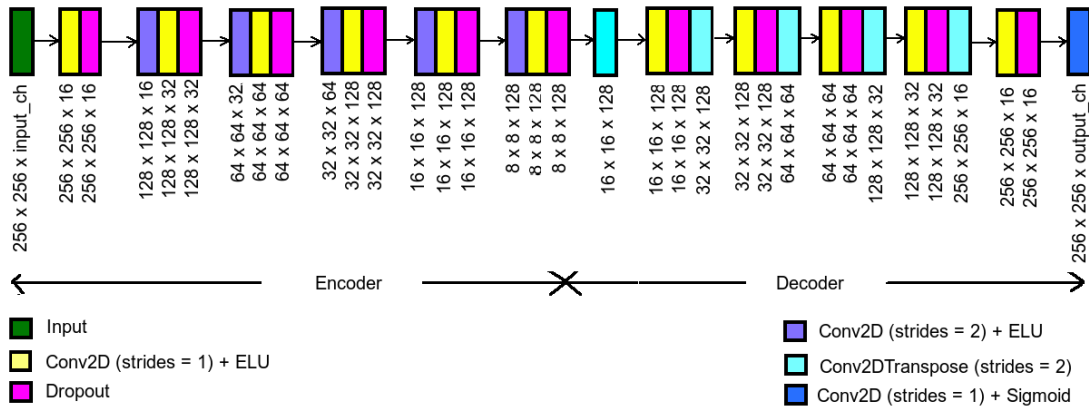
Fig. 4. Architecture of AE2 with less number of layers than the AE1 in Figure 3. Vertical text shows the output shape of the corresponding layer.

TABLE III
THE PROBABILITY OF RETRIEVING THE CORRECT RIGHT RETINA FOR A LEFT RETINA BY DIFFERENT EMBEDDINGS.

| | RGB | | | | | | | | | RBV | | | | | | | | |
| | CHASE_DB1 | | | EBD_RET1 | | | FNUSA | | | CHASE_DB1 | | | EBD_RET1 | | | FNUSA | | |
| | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| LongVector | 0.36 | 0.50 | 0.79 | 0.42 | 0.54 | 0.62 | 0.47 | 0.54 | 0.60 | 0.21 | 0.21 | 0.36 | 0.21 | 0.33 | 0.42 | 0.13 | 0.25 | 0.26 |
| PCA | 0.43 | 0.64 | 0.86 | 0.62 | 0.67 | 0.71 | 0.53 | 0.60 | 0.66 | 0.29 | 0.29 | 0.50 | 0.21 | 0.33 | 0.46 | 0.08 | 0.16 | 0.21 |
| LLE | 0.43 | 0.57 | 0.79 | 0.37 | 0.46 | 0.54 | 0.35 | 0.43 | 0.53 | 0.07 | 0.21 | 0.21 | 0.00 | 0.08 | 0.12 | 0.03 | 0.07 | 0.09 |
| Isomap | 0.14 | 0.21 | 0.36 | 0.17 | 0.42 | 0.42 | 0.29 | 0.37 | 0.44 | 0.14 | 0.14 | 0.14 | 0.08 | 0.08 | 0.08 | 0.07 | 0.11 | 0.16 |
| UNet2 | 0.64 | 0.71 | 0.86 | 0.25 | 0.54 | 0.67 | 0.47 | 0.54 | 0.59 | 0.29 | 0.36 | 0.57 | 0.21 | 0.33 | 0.50 | 0.18 | 0.25 | 0.31 |
| AE1 | 0.50 | 0.57 | 0.64 | 0.33 | 0.46 | 0.46 | 0.22 | 0.32 | 0.37 | 0.35 | 0.43 | 0.50 | 0.25 | 0.42 | 0.54 | 0.18 | 0.22 | 0.28 |
| AE2 | 0.64 | 0.64 | 0.64 | 0.41 | 0.62 | 0.67 | 0.38 | 0.51 | 0.54 | 0.36 | 0.43 | 0.50 | 0.50 | 0.54 | 0.58 | 0.13 | 0.25 | 0.34 |
| UNet2 + PCA | 0.71 | 0.79 | 0.86 | 0.25 | 0.42 | 0.54 | 0.44 | 0.54 | 0.60 | 0.29 | 0.42 | 0.50 | 0.17 | 0.33 | 0.33 | 0.13 | 0.28 | 0.31 |
| AE1 + PCA | 0.42 | 0.71 | 0.79 | 0.29 | 0.50 | 0.54 | 0.21 | 0.31 | 0.37 | 0.29 | 0.36 | 0.50 | 0.29 | 0.42 | 0.62 | 0.10 | 0.19 | 0.26 |
| AE2 + PCA | 0.50 | 0.64 | 0.71 | 0.37 | 0.50 | 0.62 | 0.29 | 0.46 | 0.51 | 0.29 | 0.36 | 0.43 | 0.41 | 0.58 | 0.62 | 0.18 | 0.25 | 0.34 |
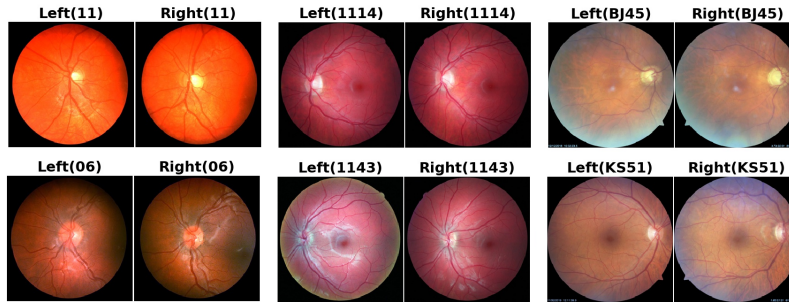


Fig. 5. Pairs of left and right RGB retinal images agreed by *all* or *nine out of ten* approaches. 1st row: pairs having strong cosine similarity claimed by all or nine out of ten approaches. 2nd row: failed to be figured out by *all* of our approaches.. 1st & 2nd cols: a pair from CHASE_DB1, 3rd & 4th cols: a pair from EBD_RET1 and 5th & 6th cols: a pair from FNUSA.
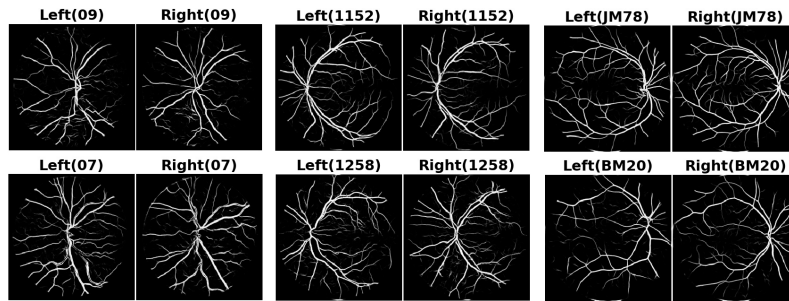


Fig. 6. Pairs of left and right RBV images agreed by *all* or *six out of ten* approaches. 1st row: pairs having strong cosine similarity claimed by six out of ten approaches. 2nd row: failed to be figured out by *all* of our approaches.. 1st & 2nd cols: a pair from CHASE_DB1, 3rd & 4th cols: a pair from EBD_RET1 and 5th & 6th cols: a pair from FNUSA.

chance. Non-linear embeddings by the neural networks such as autoencoder and U shaped CNN (i.e., U-Net) can give us better cosine similarity measurement between the left and right retinal images of a person.

## REFERENCES

[1] C. Simon and I. Golstein, "A New Scientific Method of Identification," *New York State Journal of Medicine*, vol. 35, no. 18, pp. 901–906, 1935.

[2] P. Tower, "The Fundus Oculi in Monozygotic Twins: Report of Six Pairs of Identical Twins," *Archives of Ophthalmology*, vol. 54, pp. 225–239, 1955.

[3] A. K. Jain, R. Bolle, and S. Pankanti, Eds., *Retina Identification*. Springer US, 1996, pp. 123–141.

[4] S. Biswas, J. Rohdin, T. Mnuk, and M. Drahansky, "Is there any similarity between a person's left and right retina?" in *Proceedings of the 18th International Conference of the Biometrics Special Interest Group (BIOSIG)*, ser. Lecture Notes in Informatics. GI - Group for computer science, 2019, pp. 71–82. [Online]. Available: http://www.fit.vutbr.cz/research/view_pub.php.en?id=12075

[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, April 2004.

[6] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional Networks for Biomedical Image Segmentation," *Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer, LNCS*, vol. 9351, pp. 234–241, 2015.

[7] K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine*, vol. 2, pp. 559–572, 1901.

[8] H. Hotelling, "Analysis of a complex of statistical variables into principal components," *Educational Psychology*, vol. 24, no. 6, pp. 417–441, 1933.

[9] I. T. Jolliffe, Ed., *Principal Component Analysis*. New York: Springer-Verlag, 1986.

[10] S. T. Roweis and L. K. Saul, "Nonlinear Dimensionality Reduction by Locally Linear Embedding," *SCIENCE*, vol. 290, no. 5000, pp. 2323–2326, December 2000.

[11] J. B. Tenenbaum, V. de Silva, and J. C. Langford, "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *SCIENCE*, vol. 290, no. 5000, pp. 2319–2323, December 2000.

[12] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay, "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[13] J. Staal, M. Abramoff, M. Niemeijer, M. Viergever, and B. van Ginneken, "Ridge based vessel segmentation in color images of the retina," *IEEE Transactions on Medical Imaging*, vol. 23, no. 4, pp. 501–509, 2004. [Online]. Available: https://www.isi.uu.nl/Research/Databases/DRIVE/download.php

[14] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust Vessel Segmentation in Fundus Images," *International Journal of Biomedical Imaging*, vol. 2013, 2013. [Online]. Available: https://www5.cs.fau.de/fileadmin/research/datasets/fundus-images/all.zip

[15] E. Decencire, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed database: the messidor database," *Image Analysis & Stereology*, vol. 33, no. 3, pp. 231–234, Aug. 2014. [Online]. Available: http://www.adcis.net/en/third-party/messidor/

[16] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating Blood Vessels in Retinal Images by Piece-wise Threhsold Probing of a Matched Filter Response," *IEEE Transactions on Medical Imaging*, vol. 19, no. 3, pp. 203–210, 2000. [Online]. Available: http://cecas.clemson.edu/~ahoover/stare/images/all-images.zip

[17] C. G. Owen, A. R. Rudnicka, C. M. Nightingale, R. Mullen, S. A. Barman, N. Sattar, D. G. Cook, and P. H. Whincup, "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*, vol. 31, no. 8, pp. 1933–1938, 2011. [Online]. Available: https://staffnet.kingston.ac.uk/~ku15565/CHASE\_DB1/assets/CHASEDB1.zip

[18] G. Hinton, N. Srivastava, and K. Swersky, "Neural Networks for Machine Learning Lecture 6a: Overview of mini-batch gradient descent." [Online]. Available: http://www.cs.toronto.edu/~tijmen/csc321/slides/lecture\_slides\_lec6.pdf

[19] D. Clevert, T. Unterthiner, and S. Hochreiter, "Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs)," in *4th International Conference on Learning Representations, ICLR*, 2016.

[20] C. R. Blake, W. W. Lai, and D. P. Edward, "Racial and Ethnic Differences in Ocular Anatomy," *International Ophthalmology Clinics*, vol. 43, no. 4, pp. 9–25, October 2003.