DEEP LEARNING BASED OFF-ANGLE IRIS RECOGNITION

Ehsaneddin Jalilian†, Georg Wimmer†, Andreas Uhl†, Mahmut Karakaya‡

†Department of Computer Science, University of Salzburg

‡Department of Computer Science, Kennesaw State University

PARIS LODRON UNIVERSITÄT SALZBURG

WAVELAB

Multimedia Signal Processing and Security Lab

Abstract

In this work, CNNs trained with the triplet loss function are applied to extract features for iris recognition. To analyze which parts of the eye are most suited for the CNN-based recognition system, experiments are carried out using image data from different parts of the eye (full eye, eye zoomed to iris, iris only, iris normalized, eye without iris). To analyze the impact of different gaze angles on the recognition performance, experiments are applied on: (1) different gaze angles separately, (2) image data with increasing differences in the gaze angles, and (3) corrected off-angle image data. The experiment results show superior performance of the CNN trained with the triplet loss on the iris images with more lateral gaze angles ($\geq 30°$). However, higher differences in the gaze angles between images deteriorate the network performance. Also, the results are about the same for the different parts of the eye and correcting the gaze angle did not really improve the performance of the CNN.

Main Results

- The results of the proposed CNN approach did not decrease at stronger gaze angles and maintained an EER of around 2% across all gaze angles, making it a better choice when dealing with more extreme off-angle iris images ($\geq 30°$).
- Higher differences in the gaze angles between images deteriorate the results of the proposed CNN approach ($EER \approx 2\%$ at 0° difference and $EER \approx 8\%$ at 40° difference), but to a lesser extent than most of the comparison methods.
- It is not so important which parts of the eye images are used for subject recognition, as eventually the results remain similar.
- Correcting the gaze angle did not really improve the triplet loss CNN results. However, the segmentation-CNN method did clearly benefit from using rotation corrected data.

Objectives of the research

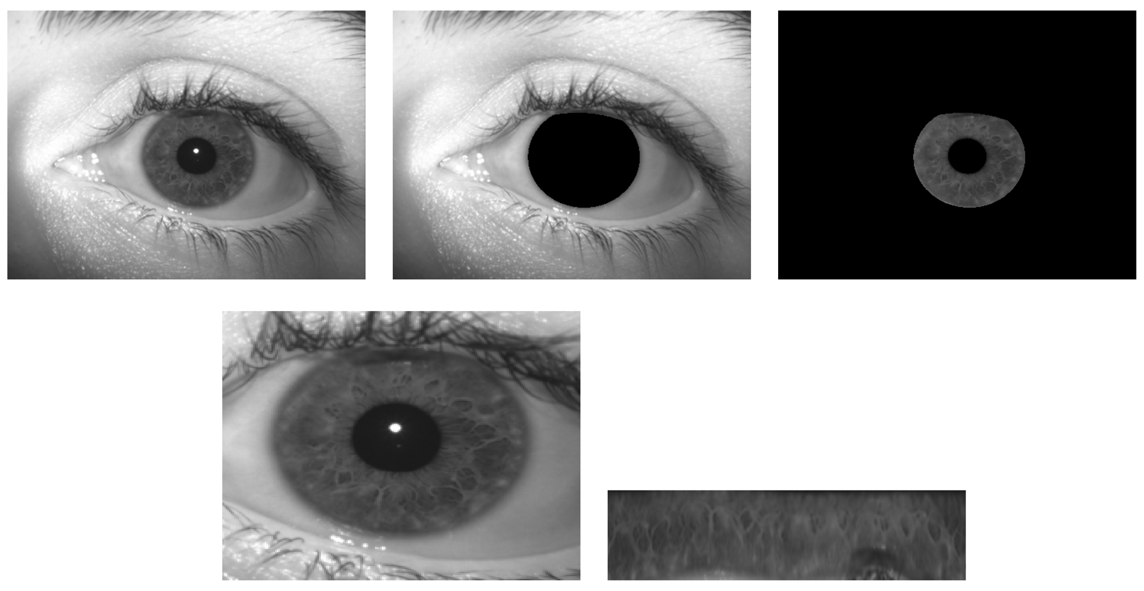
In this work we aim to address four main questions:

Q1: Are different gaze angles easier or harder for iris recognition systems? To find out if iris images with extreme gaze angles are harder to recognize than iris images that are less off-angle or entirely not off-angle (frontal view), the EER is computed separately for the images of 11 different gaze angles ($-50°, -40°, \dots, +40°, +50°$). That means for each gaze angle, only similarity scores between images of the considered gaze angle are computed for the EER.

Q2: How tolerant are iris recognition systems to off-angle iris data? To find out the impact of differences in the gaze angle between images on the results of recognition systems, we compute the EER using only similarity scores between images with a maximum gaze angle difference of θ with $\theta \in \{0°, 10°, 20°, 30°, 40°\}$.

Q3: Which parts of the eye work best for the triplet loss based CNNs? In order to find out which parts of the eye can be used for subject recognition with the triplet loss CNN, we carry out experiments using image data from different parts of the eye. We use the following image data in our experiments: (1) full eye images, (2) images zoomed to the iris, (3) images with only the iris, (4) images where the iris is removed and (5) images of the normalized iris.

Q4: Does gaze angle correction improve the results? We aim to find out if it is beneficial to correct the image gaze angles by bringing them back to the frontal view.

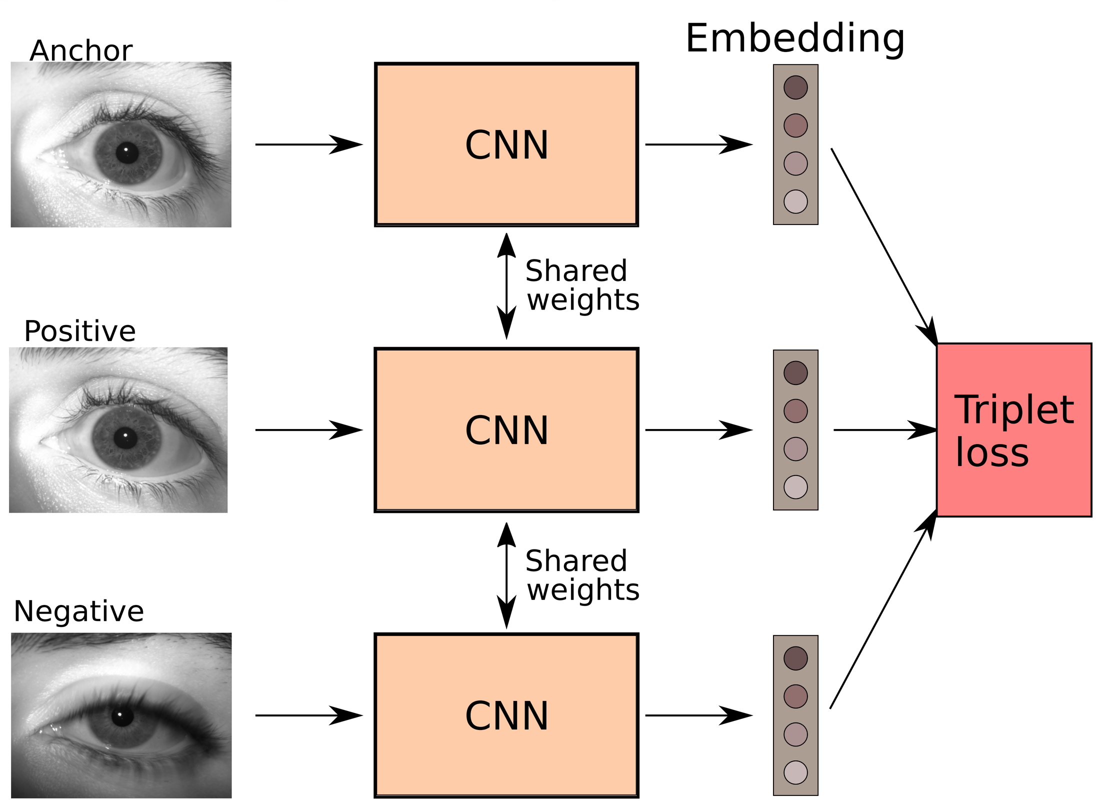


Triplet loss CNNs

The triplet loss requires three input images at once (a so called triplet), where two images belong to the same class (the so called Anchor image and a sample image from the same class, further denoted as Positive) and the third image belongs to a different class (further denoted as Negative). The triplet loss trains the network to minimize the distance between the Anchor and the Positive and maximize the distance between the Anchor and the Negative. The triplet loss using the squared Euclidean distance is defined as:

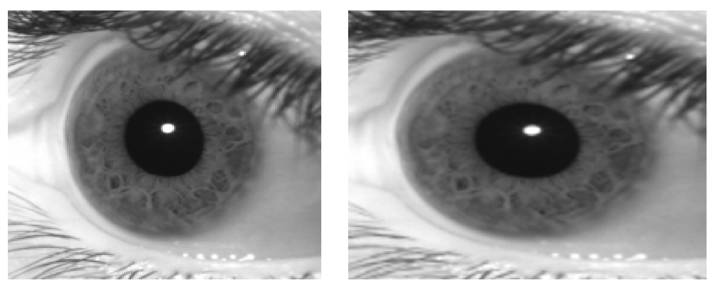
$$L_{(A,P,N)} = \max(\|f(A) - f(P)\|^2 - \|f(A) - f(N)\|^2 + \alpha, 0),$$

where A is the Anchor, P the Positive and N the Negative. α is a margin that is enforced between the positive and negative pairs and is set to $\alpha = 1$. $f(I)$ is an embedding (the CNN output) of an input image I. The CNNs are trained for 400 epochs with the ADAM optimizer, starting with a learning rate of 0.001.



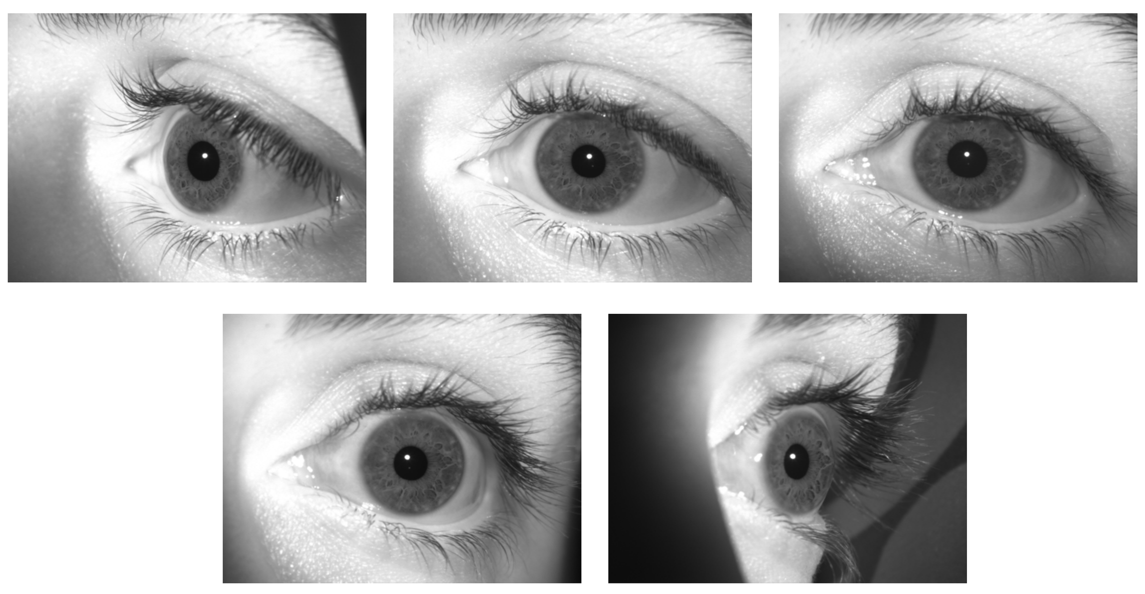
Gaze angle correction

The main issue with off-angle iris images is the related distortions such as: 3D structural changes, missing iris boundaries, and perspective and refraction distortions, which erode and deform the geometric profile of the iris. So, correcting the images (bringing them back to frontal view) may help to correct these distortions and improve the performance of iris recognition systems. To investigate this, we additionally extend our experiments to gaze angle corrected image data. The gaze angles are determined using various measures of the eye that are unique for different gaze angles and then the images are re-projected to the frontal view.



Experimental Framework

Dataset: 4400 iris images captured from 40 subjects of an off-angle iris database. Images at 0° gaze angle were captured by a frontal fixed camera, and off-angle images were captured by a frontal moving camera rotating horizontally from $-50°$ $to + 50°$ in angle with a 10° step-size. Each camera captured 10 (gray scale) iris images per stop, giving 10 frontal and 100 off-angle iris images per subject.



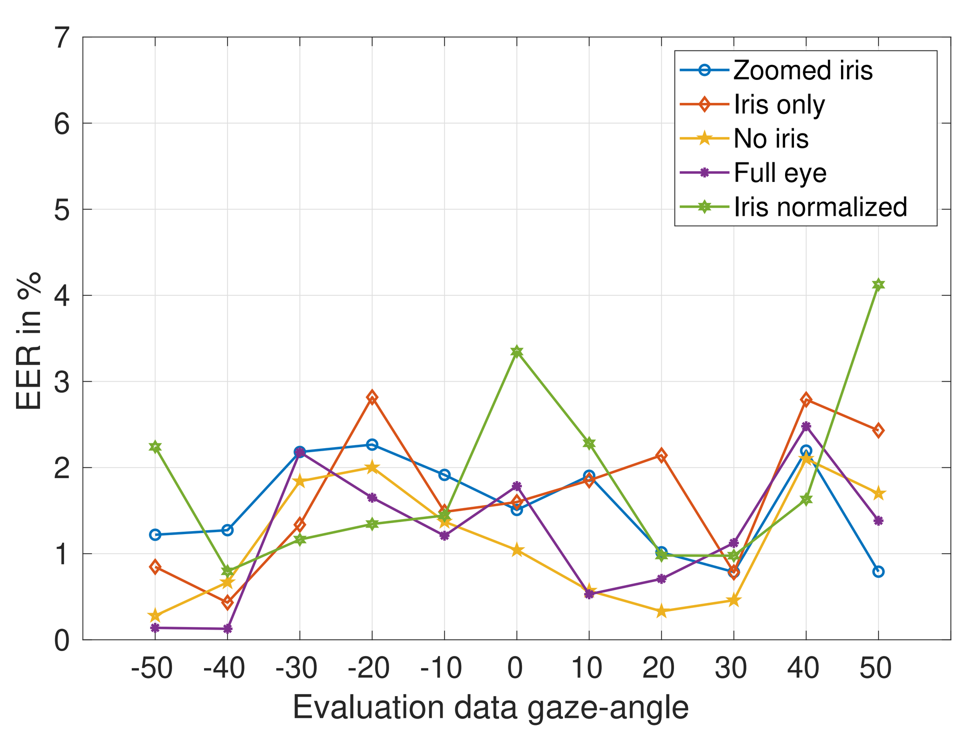
Comparison methods: The IrisSeg algorithm and the WAHET algorithm are two classical methods, and the Segmentation-CNN is a deep learning based method used in the experiments.

Training approach: We employ 2-fold cross validation to train and evaluate the CNNs. For this, we divide the whole database into two equal parts (20 subjects per fold). In the first fold, one part is used as training data and the other one as evaluation data. In the second fold, the roles are switched.

Recognition metrics: To quantify the recognition performance the Equal Error Rate (EER) is calculated (we report mean EER).

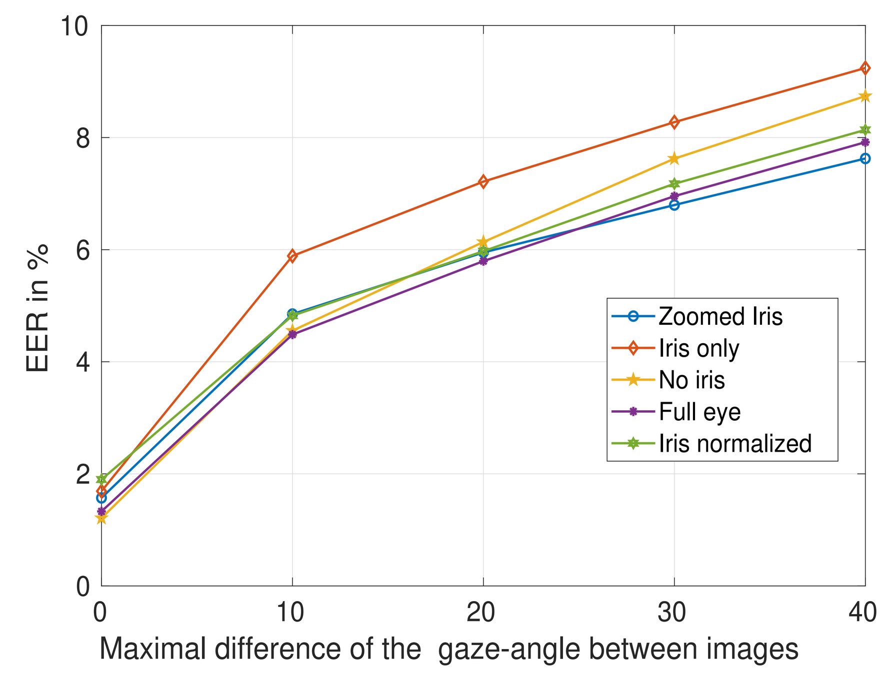
Results and discussion

Answer to Q1: More extreme gaze angles do not worsen the results compared to lower gaze angles.

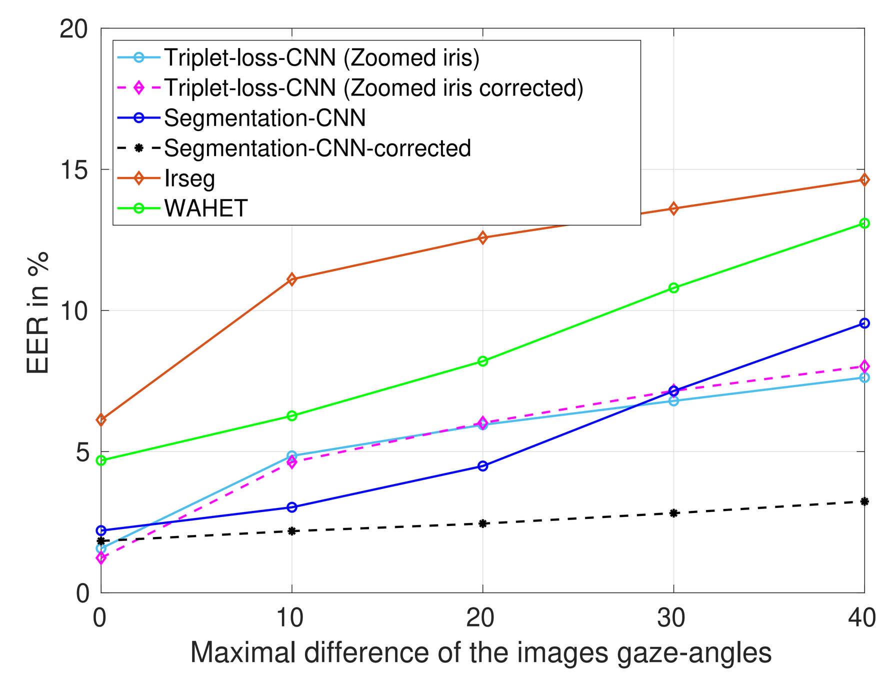


Results and discussion

Answer to Q2: Higher differences in the gaze angles between images deteriorate the results of the proposed CNN approach (EER ≈ 2% difference and EER ≈ 8% at 40°difference), but to a lesser extent than most of the comparison methods. In this case, the Segmentation-CNN combined with gaze angle correction proved to be a better choice.



Answer to Q3: It does not really matter which parts of the eye are used, the results are always about the same. Yet interestingly, it seems that the results are slightly better when removing the iris information ('No iris') than for keeping it ('Full eye'). This, together with the fact that using image data from only the iris ('Iris only') often performed worst in the experiments, indicates that iris information is less suitable for subject recognition than information from other parts of the eye, at least for the proposed CNN approach.



Answer to Q4: Correcting the gaze angles improves the results very slightly, but not consistently across all gaze angles. At more extreme gaze angles the triplet loss CNN achieves the best results. The comparison approaches perform worse at more extreme gaze angles. At lower gaze angles, Segmentation-CNN applied to the uncorrected image data achieves the best results.