Gaze-angle Impact on Iris Segmentation using CNNs

Ehsaneddin Jalilian†, Andreas Uhl†, Mahmut Karakaya‡

†Department of Computer Science, University of Salzburg

‡Department of Computer Science, University of Central Arkansas

UNIVERSITÄT SALZBURG

WAVELAB

Multimedia Signal Processing and Security Lab

Abstract

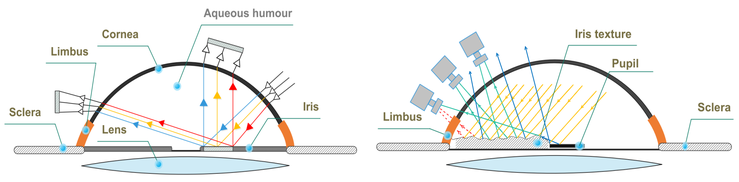
Iris images captured by emerging stand-off iris recognition systems are more subject to off-angle distortions. While deep learning techniques (*e.g.*convolutional neural networks (CNNs)) are increasingly becoming a tool of choice for iris segmentation tasks, yet there is a significant lack of information about how these distortions affect the performance of such models. We thoroughly discuss the general effect of different gaze-angles on ocular biometrics and relate the findings to off-angle iris segmentation using CNNs. We further investigate the effect of (i) increasing the quantity of iris training data in case of gaze-angles in training and test data match, and (ii) considering iris training data consisting of several distinct gaze-angles. We also compare our results to those of some classical iris segmentation algorithms, where the CNNs are found to outperform the classical algorithms.

Main Results

- Performance of the networks decline as the gaze-angles diverge
- Effect of perspective, refraction, and 3D iris structure distortions on the networks mainly appeared as missing detected iris texture
- Missing and dilated iris boundary distortions (especially in right-most gaze-angle images) cause the networks to fail to accurately extract the iris boundary pixels
- Limbus occlusion results in missing detected outer boundary iris pixels
- FCNs with shallow architecture resist more against off-angle distortions
- Increasing the quantity of training images (with heterogeneous gaze-angles) doesn't increase the networks' generalization capability
- The CNNs have better performance than classical algorithms (on off-angle iris data)

Eye structures effect on iris segmentation

3D structure of iris: There will be considerable changes in the (3D) distribution of iris features when the gaze-angle of training and testing iris images differs, and the network may not be able to spot the corresponding features.

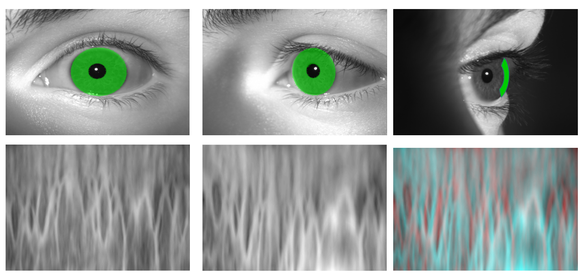


Limbus occlusion: The diameter of the cornea-limbus border is slightly smaller than the anterior chamber width. Therefore, the limbus consistently occludes the boundary region of the iris texture (especially) in extreme off-angle view.

Perspective and refraction distortion: CNNs learn scale-dependent patterns at a specific combination of image size and network architecture, and thus they are not able to spot the learned patterns in the testing data, if their geometric properties are changed with respect to the training data.

$$
\begin{aligned}
a'_{x_2} &= c_y(s_z y_1 + c_z x_1) - s_y z_1 \\
a'_{y_2} &= s_x(c_y z_1 + s_y(s_z y_1 + c_z x_1)) + c_x(c_z y_1 - s_z x_1) \\
a'_{z_2} &= c_x(c_y z_1 + s_y(s_z y_1 + c_z x_1)) - s_x(c_z y_1 - s_z x_1)
\end{aligned} \tag{1}
$$

Missing iris boundary in extreme angles: As the gaze-angle gets steeper, the sclera-iris boundary erodes and finally disappears. Thus, the network will not be able to retrieve the accurate boundary pixels (which are missing), or may spot false boundaries.



Experimental framework

Database: We used a subset (containing 4400 left eye iris images captured from 40 subjects) of an off-angle iris database. The iris images are captured by two near-infrared cameras. 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°(N50) to +50°(P50) in angle with a 10°step-size. Each camera captured 10 iris images per stop, giving 10 frontal and 100 off-angle iris images captured from each subject. For our experiments we divided the whole database into two equal parts (each containing iris images of 20 separate subjects), and used one part as our testing data and the other one as our training data.

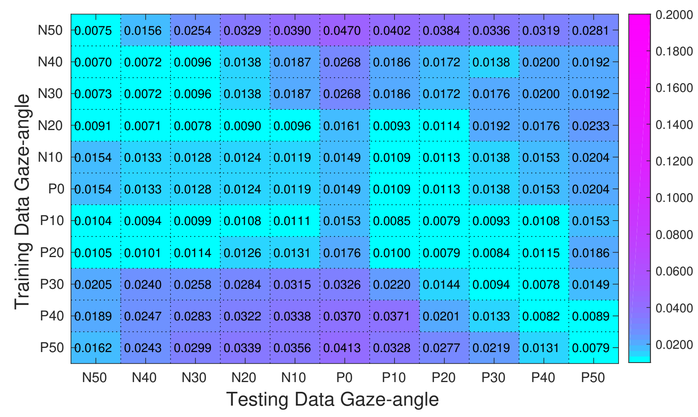
Segmentation evaluation and measures: For quantification of the resulting segmentations, we considered the *nice*1 iris segmentation error rate which is based on the NICE1 protocol. *Nice*1 for each input iris mask I_i is given by the proportion of corresponding disagreeing pixels (through the logical exclusive-or operator) over all the mask as follows:

$$nice1 = \frac{1}{c \times r} \sum_{c'} \sum_{r'} O(c', r') \otimes C(c', r') \tag{2}$$

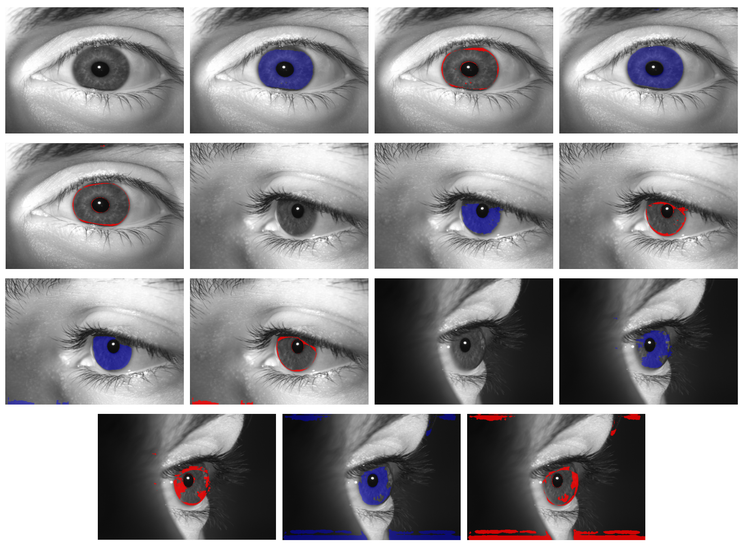
Fully convolutional neural networks (FCNs): We used two different FCN architectures: The first network architecture we used is identical to the "Basic" fully convolutional encoder-decoder network termed "SegNet". The network has a rather shallow structure, which includes an encoder, and a corresponding decoder. The second network architecture used is RefineNet. RefineNet has a very deep multi-path refinement architecture, which employs a 4-cascaded architecture with 4 Refining nets, each of which directly connects to the output of one Residual block, as well as to the preceding Refining net in the cascade.

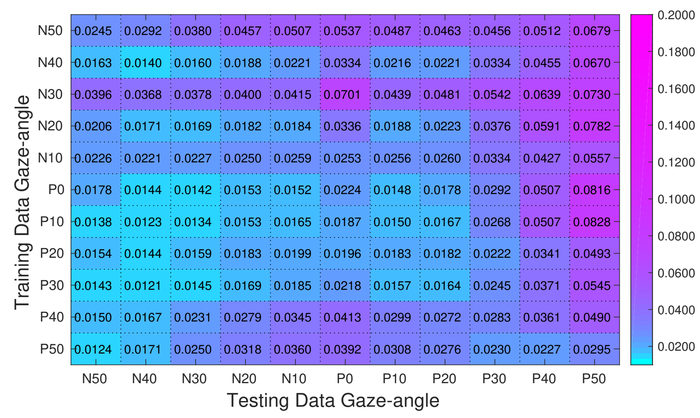
Is gaze-angle specific training required

We trained the networks on iris images with identical gaze-angles (200 images per gaze-angle as in our training data). For each available gaze-angle, a dedicated network is trained. Subsequently, we conduct segmentation on all the test data, differentiating and grouping results into the different gaze-angles available (starting from -50°(N50) to +50°(P50)).



- Missing left iris boundary and dilated right boundary (perspective and refraction distortions) in P50 gaze-angle images cause the network to fail to extract iris boundary pixels.
- The right iris boundary is also affected by the limbus occlusion, resulting in additional false-negatives detections in this area.
- The other false-negative detections visible in the iris main texture (far from the iris boundaries) seem to be mainly due to the 3D iris structure, perspective and refraction distortions.

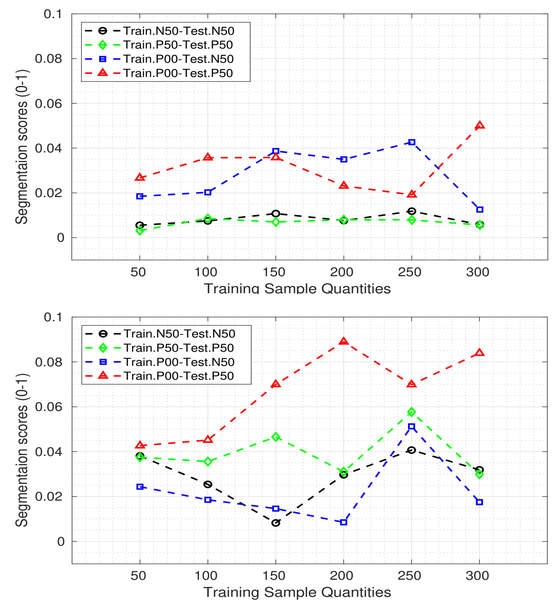




- The table (above) is less symmetric and we observe higher errors.
- The network architecture (concentrated on content information) makes it very vulnerable to the texture related distortions such as perspective, refraction and also 3D iris structure distortions.
- Correspondingly, we can observe many false-positives detections in the output masks, as well as some false-negatives detections, especially on the iris main texture, when testing the network on images acquired from steeper angles (*e.g.*P30, P40, P50). The strong decline of the network performance visible at the upper left and right side of the table is mainly due to this issue.
- Distortions affect the iris images of the left sided gaze-angles less (better results in the lower left and left side of the table).
- The effect of missing iris boundary is also visible, however the effect of limbus occlusion distortion seems not to be so severe.

Impact of the size of the training set

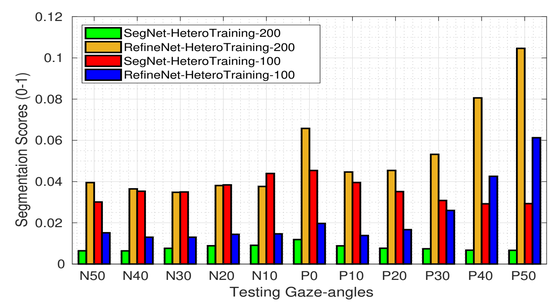
We trained the networks with: 50, 100, 150, 200, 250, and 300 pcs of N50 and P50 gaze-angle images respectively, and then tested them on the remaining (100) images of the same gaze-angles. Similarly, we trained networks with increasing quantities of frontal images (P0), and then tested on the same N50 and P50 gaze-angle images.



- Introducing more training images to the networks doesn't consistently improve the networks segmentation performance.

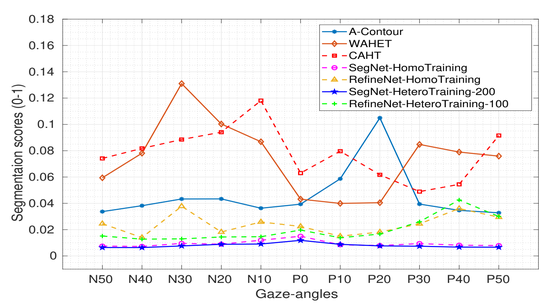
Does heterogeneous training helps, what is the optimal quantity

We trained the networks with all (200 samples per gaze-angle), and half quantity (100 samples per gaze-angle) in our training set, and then tested the networks on all images in the testing data.



- SegNet's performance gets almost identical to the results of the homogeneous training setting. This enables us to refrain from the angle-specific training strategy (and even better, from the application perspective, there is no need to determine the gaze-angles before deploying CNN-based segmentation).
- Training a (certain) network architecture (*e.g.*RefineNet) with a distinct quantity of iris images of different gaze-angles improves the network's generalization capability, while if we select an improper architecture (*e.g.*SegNet) or quantity of training data, the results may not be as expected.

Comparison with other classical iris segmentation algorithms



- FCNs have superior performance over the classical algorithms.

Acknowledgment

This project was funded by the European Union's Horizon 2020, agreements No 700259 (PROTECT), 690907 (IDENTITY) and NSF grant No CNS-1909276.