Deep Multi-class Eye Segmentation for Ocular Biometrics

Peter Rot*, Žiga Emeršič*, Vitomir Štruc†, Peter Peer*
*Faculty of Computer and Information Science, University of Ljubljana, Slovenia
E-mail: peter.rot93@gmail.com, ziga.emersic@fri.uni-lj.si, peter.peer@fri.uni-lj.si
†Faculty of Electrical Engineering, University of Ljubljana, Slovenia
E-mail: vitomir.struc@fe.uni-lj.si

***Abstract*—Segmentation techniques for ocular biometrics typically focus on finding a single eye region in the input image at the time. Only limited work has been done on multi-class eye segmentation despite a number of obvious advantages. In this paper we address this gap and present a deep multi-class eye segmentation model build around the SegNet architecture. We train the model on a small dataset (of 120 samples) of eye images and observe it to generalize well to unseen images and to ensure highly accurate segmentation results. We evaluate the model on the Multi-Angle Sclera Database (MASD) dataset and describe comprehensive experiments focusing on: *i)* segmentation performance, *ii)* error analysis, *iii)* the sensitivity of the model to changes in view direction, and *iv)* comparisons with competing single-class techniques. Our results show that the proposed model is viable solution for multi-class eye segmentation suitable for recognition (multi-biometric) pipelines based on ocular characteristics.**

***Index Terms*—Sclera, segmentation, deep learning, convolutional neural networks (CNN), iris, pupil, ocular biometrics**

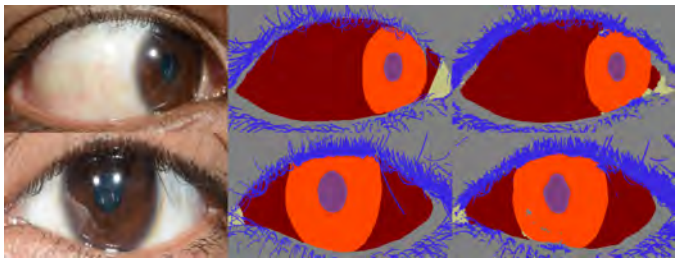

Fig. 1: Illustration of multi-class eye segmentation. The input image is segmented into multiple regions that correspond to different parts of the eye. The results of the segmentation procedure are useful for ocular biometric pipelines. The figure shows: input RGB images (left column), reference annotations (middle column), prediction by our approach (right column). The segmentation results are color-coded.

I. Introduction

Automated recognition systems relying on ocular biometric traits such as the iris, retina, sclera or the peri-ocular region have received significant attention from the research community in recent years [1]–[16]. The interest in these traits is fueled by the significant market potential of ocular biometrics, but also from a number of desirable characteristics, such as high recognition accuracy, reliability, non-invasive data acquisition and considerable user acceptance [1].

Ocular biometrics are today still dominated by iris recognition technology, which is used widely for authentication purposes, forensics applications and surveillance software. Applications relying on other ocular traits, on the other hand, are less frequent, but the research community is increasingly looking into the blood vessels of the sclera [17], [18] and the peri-ocular region [19]–[21] to extract identity cues that can either be utilized in stand-alone recognition systems or complement the information commonly used for iris recognition.

A key component of ocular recognition systems is an efficient segmentation procedure that extracts the region-of-interest (ROI) from the input images and provides the basis for the subsequent feature extraction and matching steps. While significant research efforts have been directed towards segmentation techniques for different ocular traits (e.g., [22]–[32]), these efforts have typically been limited to techniques that are able to segment only a single trait at the time from the inputs (e.g., only the iris, sclera or eye region). Remarkably, little work has been done on multi-class eye segmentation, where multiple eye parts need to be segmented from the input images with a single segmentation approach, despite some expected advantages, such as:

- Segmenting the eye image into multiple classes with a single segmentation approach makes it possible to apply recognition techniques to different parts of the image efficiently and devise (ocular) multi-biometric systems without computational overhead.
- The target classes of the segmentation procedure act as sources of contextual information for each other and are expected to improve the segmentation performance for all (or most) traits, especially in challenging conditions. For example, information about the boundaries of the sclera region provides contextual information about the bounderies of the iris area, the location of the eye leashes acts as a constraint for the sclera region and so on.

In this paper we try to address the outlined gap and present an approach to multi-class eye segmentation (as shown in Fig. 1) that is able to segment the input images in to six distinct classes that correspond to the pupil, iris, sclera, eyelashes, medial canthus and the peri-ocular region using a single segmentation model. We build on the success of deep learning models for semantic image segmentation, specifically on

convolutional encoder-decoder networks [33], and design our approach around the recently introduced SegNet architecture. To facilitate training, we manually annotate a small dataset of (visible light) eye images in a per-pixel manner, as shown in the middle column of Fig 1. We report results on the Multi-Angle Sclera Database (MASD, [29]) dataset and provide comparative experiments with competing techniques from the literature. Our experiments suggest that the proposed model is viable solution for multi-class eye segmentation that could be integrated in the future into recognition pipelines exploiting ocular biometrics.

The rest of the paper is structured as follow: In Section II we present a brief overview of the related work on eye segmentation. In Section III we describe the multi-class eye segmentation approach and corresponding training procedure. We evaluate the proposed approach in comprehensive experiments in Section IV and conclude the paper with some final comments in Section V.

II. Related work

In this section we present a brief overview of existing techniques for ROI segmentation for ocular biometrics. The reader is referred to [1] for a more complete and comprehensive coverage of the topic.

A. Iris segmentation

Iris segmentation techniques have received considerable attention from the research community [1]. The interest in these techniques is spurred mainly by the popularity of iris recognition technology and the corresponding market potential. Existing techniques for iris segmentation range from Daugman's integro-differential operator [34], active contour models [35] and clustering algorithms [36] to methods exploiting gradient (edge) information [37]–[39], variants of the Hough transform [40], [41] and others [1].

More recently, researchers started looking at classification-based methods for iris segmentation e.g., [23], [24], [42]. These methods try to assign each pixel from the input image a label from a predefined set of target labels and are thus similar in spirit to our work. Nevertheless, most of the existing methods only consider binary segmentation problems (e.g., iris vs. non-iris) or resort to separate segmentation models to find other eye-regions, e.g., most commonly the sclera [43], that can contribute towards more efficient iris segmentation. Our approach, on the other hand, segments the not only the iris, but also five other eye-parts from the image using a single segmentation model. Because all six classes are segmented jointly, there is limited computational overhead and the model is forced to learn inter-class-relationships during training.

B. Sclera segmentation

As emphasized in [31], the problem of sclera segmentation is typically studied as a sub-problem of other tasks, such as iris recognition or gaze estimation. Nonetheless, research into sclera segmentation techniques is gaining traction within the biometric community, owing mostly to the series of sclera segmentation competitions, organized in the scope of major biometric conferences [29], [30], [32], [44] and the potential of sclera recognition techniques for stand-alone biometric systems or or supportive recognition technology for other ocular traits.

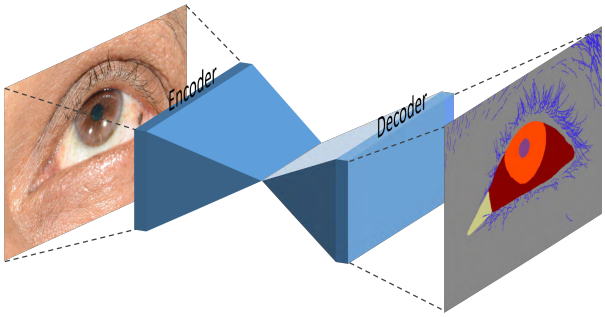

Fig. 2: We use the SegNet [49] convolutional encoder-decoder (CED) architecture for multi-class eye segmentation. Our approach is able to segment six classes from the image, i.e., the pupil, the iris, the sclera region, the eyelashes, the medial canthus and the periocular-region, using a single segmentation model.

The literature on sclera segmentation is split between *i)* techniques that aim to extract only information about the sclera vasculature from images and *ii)* methods that try to segment the entire sclera region from the input eye-images. An overview of techniques from the first group can be found in [45]. Among the techniques from the second group (which are also of relevance to this work) convolutional decoder-encoder (CED) architectures have recently been shown to ensure state-of-the-art performance, as they represented the winning techniques from the 2017 and 2018 sclera segmentation competitions, where they were able to outperform competing approaches by a large margin - see [29], [44] for details. In this paper we build on the success of CED networks for sclera segmentation, but present an approach capable of segmenting multiple classes from the input images.

C. Other traits

Segmentation techniques for other ocular traits are limited in the literature. Some research has been conducted on eyelid and eyelash segmentation with the goal of masking out occluded regions during iris recognition, e.g., [46]–[48]. However, these attempts are again limited to segmentation/detection models trained and applied separately from other segmentation approaches and can, hence, make only limited use of the available contextual information provided by other ocular traits.

III. Methodology

We now describe our approach to multi-class eye segmentation. As illustrated in Fig. 2, we use a convolutional encoder-decoder (CED) network architecture to segment the input eye images into six distinct classes, i.e., the pupil, the iris, the sclera region, the eyelashes, the medial canthus and

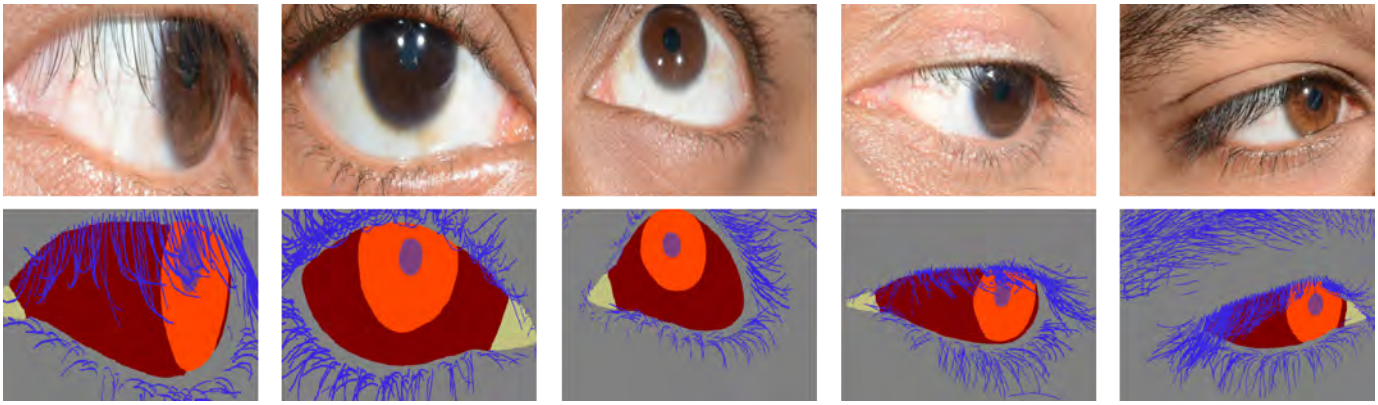

Fig. 3: Examples of eye images (top row) and corresponding pixel-level annotations (bottom row) used in this work. The annotations are color coded and cover six distinct eye parts, i.e., the pupil, the iris, the sclera region, the eyelashes, the medial canthus and the periocular-region. The images were taken from the Multi-Angle Sclera Database (MASD, [29]). The figure is best viewed in color.

the periocular-region. Our network is based on the recently introduced SegNet architecture from [49], which was shown to ensure state-of-the-art results for various segmentation tasks and was already successfully applied for the problem of sclera segmentation [50].

In the remainder of this section, we discuss (in detail) the SegNet architecture, the training procedure and our multi-class segmentation approach.

A. The SegNet architecture

The SegNet [49] (semantic-segmentation) network architecture consists of two high level building blocks: an encoder and a decoder. The goal of the encoder is to compress the semantic content of the image and generate a descriptive image representation that can then be fed to the decoder, which then produces the final segmentation output [51].

SegNet's encoder is inspired by the VGG-16 [52] architecture, but unlike VGG-16, the encoder uses only convolutional layers and no fully-connected layers. This is what makes SegNet a fully-convolutional network [53] (FCN). The SegNet encoder consists of 13 convolutional layers (followed by batch normalization and ReLU activations) and 5 pooling layers. The decoder is an additional, inverted VGG-16 model without fully-connected layers, which has a pixel-wise soft-max layer at the top. The soft-max layer generates a probability distribution for each pixel that can be exploited to classify pixels into one of the predefined semantic target classes. Thus, during training the encoder needs to learn to produce low-resolution semantically-meaningful feature maps, whereas the decoder has to learn filters that can generate high-resolution segmentation masks from the low-resolution feature maps produced by the encoder.

A unique aspect of the SegNet model are so-called skip-connections that connect the pooling layers of the encoder with the corresponding up-sampling layers of the decoder and propagate spatial information (pooling indices) from the encoder to the decoder, which helps avoid information loss through the network. Consequently, SegNet's output probability masks have the same dimensions (i.e., width and height) as the input images, which enables relatively precise segmentation. The number of output probability masks is equal to the number of semantic target classes.

B. Multi-class eye segmentation and network training

To be able to train the SegNet model for multi-class eye segmentation, we generate the needed training data by manually annotating a (small) dataset of eye images at the pixel level with $C = 6$ target-class labels $\Omega = \{\omega_i\}_{i=1}^{C}$ that correspond to six distinct eye regions, i.e., the pupil, the iris, the sclera region, the eyelashes, the medial canthus and the periocular-region, as shown in Fig. 3. We use the annotated images and learn the parameters of the SegNet model using categorical cross-entropy as our training objective. Since certain classes are represented more frequently in the images than others (i.e., the images are dominated by peri-oculuar, iris and sclera pixels) we use median frequency balancing to force the network to efficiently segment less frequent classes as well. Such an approach is also advocated in [49].

Once the model is trained, it takes an RGB eye image $I(x, y) \in \mathbb{R}^{w \times h \times 3}$, consisting of wh pixels, as input and returns a probability distribution over the $C = 6$ target classes for each pixel location. This is, for each location $s = [x, y]^T$ in the input image $I(x, y)$, the network outputs a distribution $\mathbf{p}_s = [p_{\omega_1}, \dots p_{\omega_C}]^T \in \mathbb{R}^{C \times 1}$, where p_{ω_i} denotes the probability that the pixel at location s belongs to the i-th target class and $\sum_{i=1}^{C} p_{\omega_i} = 1$. To generate the final segmentation result, we consider two strategies:

- *Max-out strategy (MOS):* We predict the final class label $\hat{\omega}_k \in \Omega$ for each pixel location s by finding the maximum of the generated probability distribution, i.e.,

$$\hat{\omega}_k = \arg\max_{\omega_i} (\mathbf{p}_s). \tag{1}$$

  This strategy assigns the class label with the highest probability to the pixel at location s regardless of the

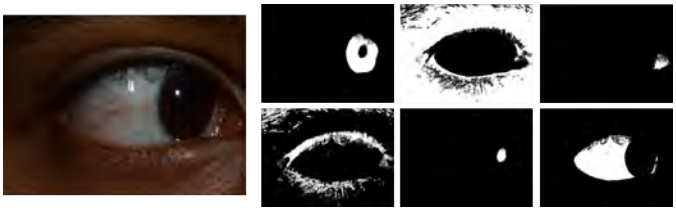

Fig. 4: Example probability maps $P_{\omega_i}(s)$ generated by the softmax layer of the SegNet model. The figure shows the input image (on the left) and corresponding probability maps (from left to right and top to bottom): the iris, the peri-ocular region, the medial canthus, the eyelashes, the pupil and the sclera region.

actual value of the probability (i.e., even if the class probability is low or comparable to the probability of another class). As a result, each pixel in the input images is assigned one (and only one) class label.

- *Thresholding strategy (THS):* We construct six probability maps (one per class label) from the probability distributions $\mathbf{p}_s$ of all pixel locations, i.e.,

$$P_{\omega_i}(s) = p_{\omega_i}^{(s)}, \tag{2}$$

where the superscript $^{(s)}$ indicates that the probability p_{ω_i} corresponds to pixel location $s = [x, y]^T$ and $x = 1, \ldots, w$, $y = 1, \ldots, h$. An example of the probability maps $P_{\omega_i}(s)$ generated for a sample eye image are shown in Fig. 4. The final segmentation result (i.e., the segmentation masks) is generated by thresholding the probability maps using some class-specific threshold Δ_{ω_i}, i.e., $P_{\omega_i}(s) > \Delta_{\omega_i}$. Because we tune the thresholds independently for each class, not all pixels may be assigned a class label, while some may be assigned multiple labels.

IV. Experiments and Results

In this section we evaluate the proposed multi-class eye segmentation approach. We start the section by describing the experimental database, proceed by elaborating on the network training procedure and then present the final experimental results.

A. Database description

For the evaluation, we use part of the Multi-Angle Sclera Database (MASD, [29]), which in total features 2624 eye images of variable size and corresponding masks for the sclera region. The images were taken in the visible spectrum in completely unconstrained settings. The variability of the images is mostly across view directions, as the subjects look either to the left, to the right, up or down.

Because we are interested in multi-class eye segmentation, we manually annotate 120 images (belonging to 30 distinct subjects) in a per-pixel manner as shown in Fig. 3. This makes it possible to train our SegNet model for multi-class eye segmentation and to evaluate it's performance.

B. Experimental protocol and performance metrics

We use a 4-fold cross validation experimental protocol for our experiments. During each fold 90 images are used as the training set and 30 as the test set. The training images are utilized to learn the parameters of the segmentation models, and the test images are used for performance evaluation.

We measure the performance of the multi-class segmentation approach on a per-class basis and, therefore, compute precision, recall and $F1$ scores for all classes in each experiment. The performance metrics are defined as follows:

$$precision = \frac{TP}{TP + FP}, \tag{3}$$

$$recall = \frac{TP}{TP + FN}, \tag{4}$$

$$F1\text{-}score = 2 \cdot \frac{precision \cdot recall}{precision + recall}, \tag{5}$$

where TP denotes the number of true positive pixels, FP stands for the number of false positive pixels, and FN represents the number of false negative pixels.

Among the above measures, precision measures the proportion of correctly segmented pixels with respect to the overall number of true pixels of a given class (i.e., it provides information about how many segmented pixels are relevant). Recall measures the proportion of correctly segmented pixels with respect to the overall number of pixels assigned to a certain class (i.e., it provides information about how many relevant pixels are segmented for a given class). It is often convenient to combine precision and recall into a single metric called F_1-score, in particular if a simple way to compare two segmentation models is required.

Note that with the MOS strategy, we obtain fixed precision and recall values for the segmentation output, while for the THS strategy complete precision-recall curves can be generated for each of the six target classes by varying the corresponding decision threshold Δ_{ω_i}.

C. Training details

Prior to SegNet training we resize all training images to a fixed size of 640×480 pixels. Since we only have 90 images available during each experimental fold for learning the parameters of the segmentation models, we use heavy data augmentation as suggested in [54]. Specifically, we generate 100 variants of each training image by flipping images left-right, generating sub-crops, Gaussian blurring, brightness changes, and applying affine transformations such as scaling, rotation or shear. Our experiments suggest that the small number of images together with data augmentation provides sufficient training data to learn segmentation models that perform robustly in difficult conditions and generalize well to unseen subjects.

For the training phase, we use the Caffe implementation of SegNet from [49]. We train the model on a GTX 1080 Ti graphics card with $30,000$ iterations using a batch size of 4, learning rate of 0.001 and stohastic gradient descent

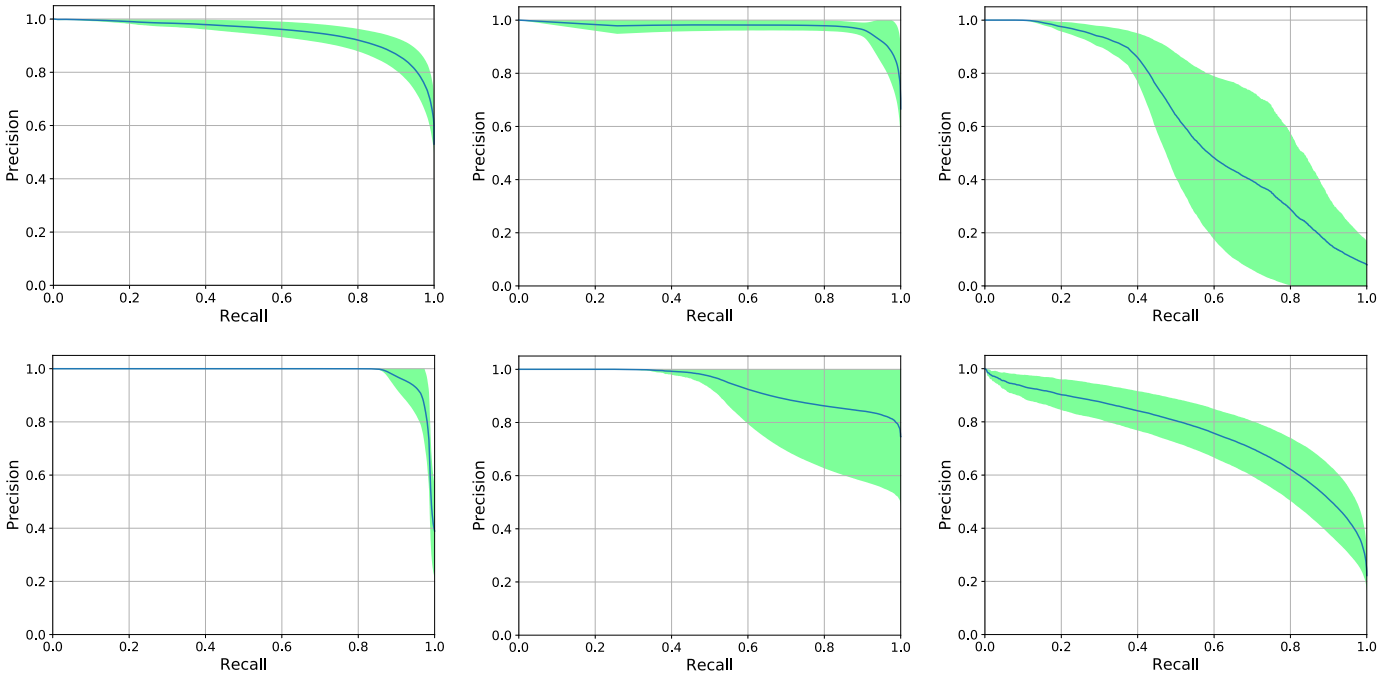

Fig. 5: Average precision-recall curves (with standard deviation) generated during the evaluation of the SegNet model using the THS strategy and a 4-fold cross validation scheme. The graphs show results for (from left to right and top to bottom): the peri-ocular region, the sclera, the canthus, the iris, the pupil and the eyelashes region.

TABLE I: Multi-class segmentation results for the MOS strategy. Average performance metrics and the corresponding standard deviations are reported.

Class (eye region)	Precision	Recall	F1-score
Iris	0.94 ± 0.05	0.89 ± 0.12	0.91 ± 0.10
Sclera	0.92 ± 0.04	0.90 ± 0.09	0.91 ± 0.07
Pupil	0.89 ± 0.16	0.84 ± 0.21	0.85 ± 0.19
Peri-ocular	0.92 ± 0.05	0.89 ± 0.06	0.90 ± 0.05
Eyelashes	0.57 ± 0.10	0.72 ± 0.12	0.63 ± 0.10
Canthus	0.67 ± 0.28	0.45 ± 0.25	0.49 ± 0.24

(SGD). All four view directions are represented equally in all training sets. At test time the model takes approximately 150 milliseconds on average to make a prediction for a single input image (calculated over 100 input images) when run on the GPU.

D. Multi-class segmentation results

In our first experiments we examine the performance of the SegNet model for multi-class eye segmentation using 4-fold cross validation outlined in Section IV-B. We report results in the form of average precision, recall and $F1$-scores and corresponding standard deviations for the MOS strategy in Table I and in the form of precision-recall curves for the THS strategy in Fig. 5. note that the use of the class-specific threshold in the THS strategy allows us to show complete precision-recall curves, instead of a single operating point as this is the case with the MOS strategy.

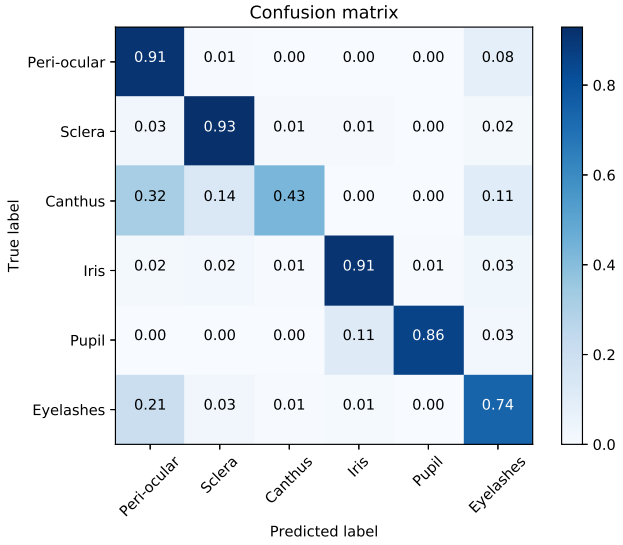

Fig. 6: Confusion matrix generated for the MOS strategy. The matrix shows the type of errors produced by the SegNet model.

The results show that the iris, sclera and peri-ocluar region are segmented by the model quite accurately. The segmentation performance for the pupil is slightly worse, but exhibits large deviations across the experimental folds. The segmentation performance for the eyelashes and canthus is considerably lower, which is understandable given the fine structure of the eyelashes that is difficult to segment and the poor contrast with respect to the surrounding region in the canthus area.

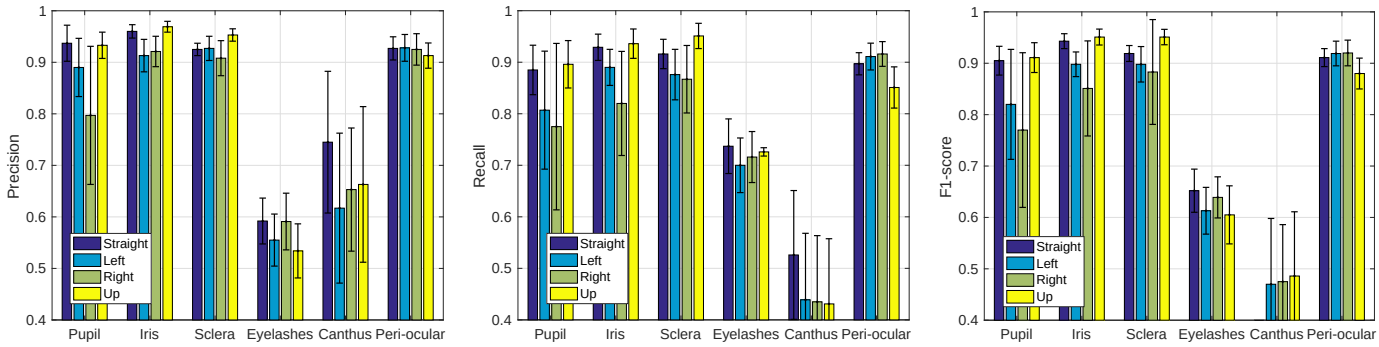

Fig. 7: Sensitivity analysis: the graphs show the average precision (left), recall (middle) and $F1$-scores (right) with respect to the view direction of the subjects, i.e., straight, left, right or up. Standard errors for the average values are also shown at the top of the bars ($n = 4$).

To get better insight into the types of errors that are produced by the SegNet model, we present a class-normalized confusion matrix for the MOS strategy in Fig. 6. The matrix show which classes are confused most often and is normalized for each eye part separately - the sum of all values in one row equals one. We see that the peri-ocular region is most often confused with the eyelashes, cantus pixels are mixed with peri-ocular pixels and (to a lesser extent) also with sclera pixels, the eyelashes are often classified as belonging to the peri-ocular region, and the pupil is confused with the iris. The sclera and iris regions have more or less evenly distributed errors among the other classes.

E. Sensitivity analysis and qualitative results

Next, we investigate the sensitivity of the SegNet model to changes in view direction. The MASD dataset, used in our experiments contains eye images of four different views (i.e., looking straight, left, right and up), so it is of interest to examine whether the performance of the model varies with respect to the view direction. Since the MOS strategy ensures a reasonable operating point with comparable precision and recall values, we use this strategy for the sensitivity analysis. The result of our experiments are presented in Fig. 7.

When comparing recall values for the pupil or iris by view direction, we see that the values are higher when the eye is looking straight/up and lower when looking left or right. The reason for such an result can likely be attributed to the fact that in the dataset eyes are not as widely open when looking left/right as they are when looking straight/up. Because of this there are more occlusions by eyelashes. The second reason is that all irises in the dataset are brown, and when the eye is looking right/left the brown iris color becomes darker in the corners of an eye and the model often confuses pupil pixels with iris pixels.

Sclera segmentation is on average the most successful when the eye is looking up (both precision and recall exceed 0.95 with the MOS strategy). We believe the main reason here is that there are usually no eyelashes covering the sclera region in these images. The second reason is that on images where the eye is looking up, there is only one continuous sclera region, compared to other cases where the sclera region is often split into two parts because of the iris. When looking left/right there is also often a small region of the sclera next to the medial/lateral canthus (depending on eye and the direction of view) which is frequently undetected or confused for the canthus itself.

Some of the best and worst per-pixel results for each view direction are presented in Fig. 8. Here, the first column of image triplets represent examples of the best segmentation results and the second column of triplets represents examples of the worst segmentation results. In each image triplet, the first image is the original input image, the second is the manually annotated ground truth and the third is the SegNet prediction. Example $a)$ presents a straight view, $b)$ a view to the left, $c)$ a view to the right and $d)$ a view up. The sclera, the pupil, the iris and the canthus are classified (segmented) very accurately in these examples. With example $a)$ we can see that the segmented pupil is a little smaller than the ground truth. In examples $b)$ and $c)$ we can also see that the model detects the majority of eyelashes which are located close to one another, but usually misses the separate ones.

On the right side of Fig. 8 some of the worst per-pixel results are presented for each direction of view. With example $e)$ we can see that a lot of the iris is misclassified as the medial canthus. The eyes marked as $f)$ and $g)$ are almost closed. In $f)$ we can see that there is almost no canthus detected and the bottom part of the iris is misclassified as peri-ocular skin. With $g)$ there is significant occlusions present due eyelashes and thus almost no iris pixels are correctly classified. In $h)$ we can see that the right part of the sclera is not detected due to the dark sclera pattern, and it is misclassified as peri-ocular skin. Also the top most part of iris is classified as eyelashes because of the very contrasting transition from the iris to the upper eyelid.

F. Comparison with single-class segmentation techniques

In our last experiment, we compare the performance of the multi-class SegNet model to the performance of the SegNet model trained only for a single class. We train the single-class model with the same set of parameters as the multi-

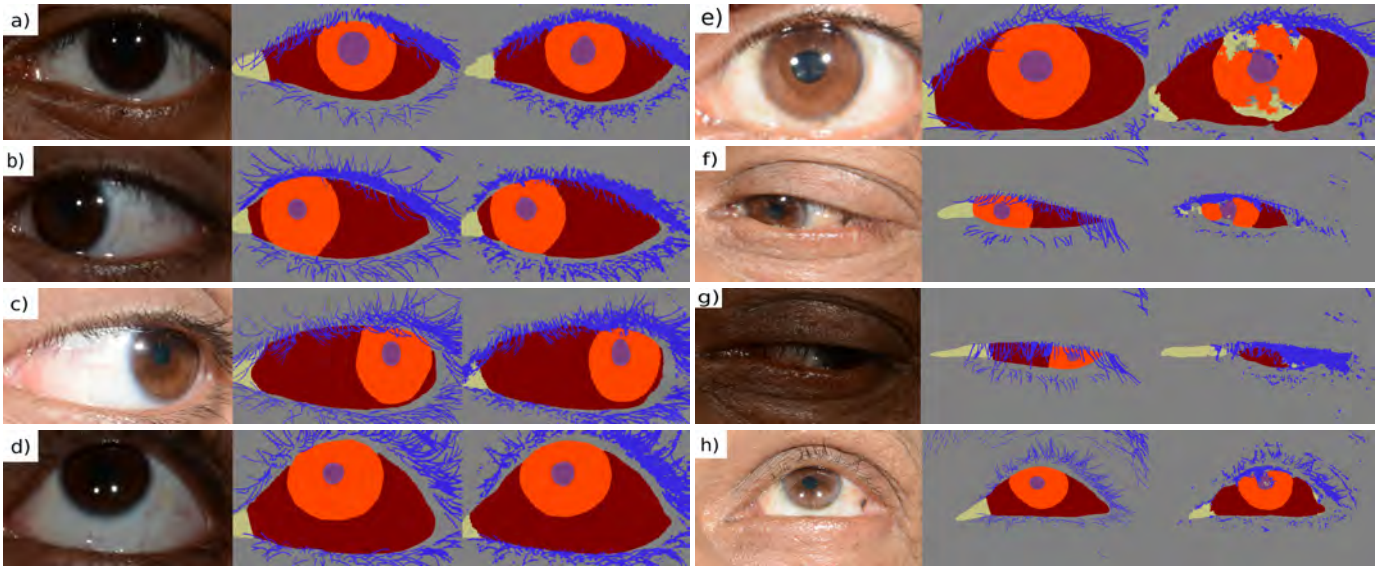

Fig. 8: Examples of some of the best and worst SegNet multi-class predictions. The first column of image triplets represents examples of the best SegNet predictions and the second column represents examples of the worst predictions. In each triplet the first image is the original RGB image, the second is the ground truth and the third is the segmentation result.

TABLE II: Comparison of multi-class and single-class segmentation results for the sclera region and the MOS strategy using SegNet.

Performance metric	Precision	Recall	F1-score
Sclera (single-class)	0.88 ± 0.08	0.93 ± 0.07	0.90 ± 0.07
Sclera (multi-class)	0.92 ± 0.04	0.90 ± 0.09	0.91 ± 0.07

class model and observe approximately the same training time as with the multi-class model. We focus on the sclera region in this experiment, where the single-class SegNet model currently represents the state-of-the-art on the MASD dataset as reported in recent sclera-segmentation competitions [44], [50]. We report result for the MOS strategy in Table II.

As we can see from the presented results, both model perform comparably and no statistically significant differences can be observed between the results. This results suggests that multi-class segmentation has no disadvantage (in terms of performance) in comparison to the single-class approach despite the fact that it is able to segment multiple eye parts form the input images. It also needs to be noted that the computational complexity of both models is identical.

A qualitative comparison of the segmentation masks obtained with both methods (multi-class and single-class) is presented in Fig. 9. We can see that both models predict the majority of pixels the same way and that the largest difference is close to the borders with other eye parts, e.g., close to the eyelashes, the canthus, etc.

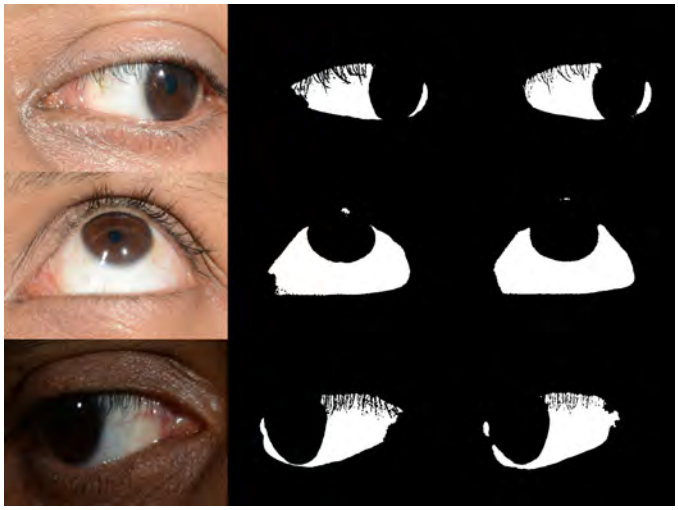

Fig. 9: Qualitative results for the sclera region obtained with the multi-class and single-class SegNet models and the MOS strategy. The figure shows (from left to right): the input image, the multi-class results, the single-class result.

V. Conclusion

We have presented a multi-class approach to eye-image segmentation based on the SegNet model. Our results show that the model is able to segment eye images into a number of target classes with high accuracy. As part of our future work we plan to integrate the segmentation model into a multi-modal ocular biometric pipeline and evaluate the multi-class segmentation approach on other iris datasets with larger variability in eye colors.

Acknowledgements

This research was supported in parts by ARRS (Slovenian Research Agency) Research Program P2-0250 (B) Metrology and Biometric Systems, ARRS Research Program P2-0214 (A)

Computer Vision, and the RS-MIZŠ and EU-ESRR funded GOSTOP. The Titan Xp GPU used for this research was donated by the NVIDIA Corporation.

REFERENCES

[1] I. Nigam, M. Vatsa, and R. Singh. Ocular biometrics: A survey of modalities and fusion approaches. *Information Fusion*, 26:1–35, 2015.

[2] A. Sequeira, L. Chen, J. Ferryman, P. Wild, F. Alonso-Fernandez, J. Bigun, K. Raja, R. Raghavendra, C. Busch, T. Freitas Pereira, et al. Cross-eyed 2017: Cross-spectral iris/periocular recognition database and competition. In *IJCB*, 2017.

[3] H. Proença and J. Neves. Deep-prwis: Periocular recognition without the iris and sclera using deep learning frameworks. *IEEE TIFS*, 13(4):888–896, 2018.

[4] Z. Zhao and A. Kumar. Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network. *IEEE TIFS*, 12(5):1017–1030, 2017.

[5] F. Alonso-Fernandez and J. Bigun. Near-infrared and visible-light periocular recognition with gabor features using frequency-adaptive automatic eye detection. *IET Biometrics*, 4(2):74–89, 2015.

[6] T. R. Borah, K. K. Sarma, and P. H. Talukdar. Retina recognition system using adaptive neuro fuzzy inference system. In *IC4*, pages 1–6, 2015.

[7] A. Das, P. Mondal, U. Pal, M. Blumenstein, and M. Ferrer. Sclera vessel pattern synthesis based on a non-parametric texture synthesis technique. In *CVIP*, pages 241–250. Springer, 2017.

[8] S. Alkassar, W. L. Woo, S. Dlay, and J. Chambers. Sclera recognition: on the quality measure and segmentation of degraded images captured under relaxed imaging conditions. *IET Biometrics*, 6(4):266–275, 2016.

[9] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers. Robust sclera recognition system with novel sclera segmentation and validation techniques. *IEEE TSMC-A*, 47(3):474–486, 2017.

[10] D. Yadav, N. Kohli, J. Doyle, R. Singh, M. Vatsa, and K. Bowyer. Unraveling the effect of textured contact lenses on iris recognition. *IEEE TIFS*, 9(5):851–862, 2014.

[11] M. De Marsico, A. Petrosino, and S. Ricciardi. Iris recognition through machine learning techniques: A survey. *PRL*, 82:106–115, 2016.

[12] J. Daugman. Iris recognition at airports and border crossings. *Encyclopedia of Biometrics*, pages 998–1004, 2015.

[13] K. B. Raja, R. Raghavendra, V. K. Vemuri, and C. Busch. Smartphone based visible iris recognition using deep sparse filtering. *PRL*, 57:33–42, 2015.

[14] K. Grm, V. Štruc, A. Artiges, M. Caron, and H. Ekenel. Strengths and weaknesses of deep learning models for face recognition against image degradations. *IET Biometrics*, 7(1):81–89, 2017.

[15] J. K. Pillai, V. Patel, R. Chellappa, and N. Ratha. Robust and secure iris recognition. In *Handbook of Iris Rec.*, pages 247–268. Springer, 2016.

[16] K. Nguyen, C. Fookes, R. Jillela, S. Sridharan, and A. Ross. Long range iris recognition: A survey. *PR*, 72:123–143, 2017.

[17] V. Patil and A. M. Patil. Human Identification Method: Sclera Recognition. *IJCSN*, 6(1):24–29, 2017.

[18] S. Alkassar, W. L. Woo, S. S. Dlay, and J. A. Chambers. Robust sclera recognition system with novel sclera segmentation and validation techniques. *IEEE TSMC-A*, 47(3):474–486, 2017.

[19] Z. Ali, U. Park, J. Nang, J.S. Park, T. Hong, and Sungjoo Park. Periocular recognition using uMLBP and attribute features. *TIIS*, 11(12):6133–6151, 2017.

[20] J. M. Smereka and V. Kumar. What is a" good" periocular region for recognition? In *CVPRW*, pages 117–124, 2013.

[21] C. Tan and A. Kumar. Towards online iris and periocular recognition under relaxed imaging constraints. *IEEE TIP*, 22(10):3751–3765, 2013.

[22] M. Abdullah, S. Dlay, W. Woo, and J. Chambers. Robust iris segmentation method based on a new active contour force with a noncircular normalization. *IEEE TSMC-A*, 47(12):3128–3141, 2017.

[23] E. Jalilian, A. Uhl, and R. Kwitt. Domain adaptation for cnn based iris segmentation. *BIOSIG*, 2017.

[24] E. Jalilian and A. Uhl. Iris segmentation using fully convolutional encoder–decoder networks. In *Deep Learning for Biometrics*, pages 133–155. Springer, 2017.

[25] N. Liu, H. Li, M. Zhang, J. Liu, Z. Sun, and T. Tan. Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In *ICB*, pages 1–8, 2016.

[26] X. Liu, K. W. Bowyer, and P. J. Flynn. Experiments with an improved iris segmentation algorithm. In *AIAT*, pages 118–123, 2005.

[27] S. Shah and A. Ross. Iris segmentation using geodesic active contours. *IEEE TIFS*, 4(4):824–836, 2009.

[28] J. Daugman. How iris recognition works. In *The essential guide to image processing*, pages 715–739. Elsevier, 2009.

[29] A. Das, U. Pal, M. A. Ferrer, M. Blumenstein, D. Štepec, P. Rot, Ž. Emeršič, P. Peer, V. Štruc, S.A. Kumar, et al. SSERBC 2017: Sclera segmentation and eye recognition benchmarking competition. In *IJCB*, pages 742–747, 2017.

[30] A. Das, U. Pal, M. A. Ferrer, and M. Blumenstein. SSBC 2015: sclera segmentation benchmarking competition. In *BTAS*, pages 1–6, 2015.

[31] P. Radu, J. Ferryman, and P. Wild. A robust sclera segmentation algorithm. In *BTAS*, pages 1–6, 2015.

[32] A. Das, U. Pal, M. A. Ferrer, and M. Blumenstein. Ssrbc 2016: sclera segmentation and recognition benchmarking competition. In *ICB*, pages 1–6, 2016.

[33] Y. Guo, Y. Liu, T. Georgiou, and M. S. Lew. A review of semantic segmentation using deep neural networks. *International Journal of Multimedia Information Retrieval*, 2017.

[34] J. Daugman. High confidence visual recognition of persons by a test of statistical independence. *IEEE TPAMI*, 15(11):1148–1161, 1993.

[35] J. Daugman. New methods in iris recognition. *IEEE TSMC-B*, 37(5):1167–1175, 2007.

[36] T. Tan, Z. F. He, and Z. Sun. Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. *IVC*, 28(2):223–230, 2010.

[37] M. De Marsico, M. Nappi, and R. Daniel. Is_is: Iris segmentation for identification systems. In *ICPR*, pages 2857–2860, 2010.

[38] G. Sutra, S. Garcia-Salicetti, and B. Dorizzi. The viterbi algorithm at different resolutions for enhanced iris segmentation. In *ICB*, pages 310–316, 2012.

[39] H. Li, Z. Sun, and T. Tan. Robust iris segmentation based on learned boundary detectors. In *ICB*, pages 317–322, 2012.

[40] J. Koh, V. Govindaraju, and V. Chaudhary. A robust iris localization method using an active contour model and hough transform. In *ICPR*, pages 2852–2856, 2010.

[41] A. Uhl and P. Wild. Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. In *ICB*, pages 283–290, 2012.

[42] M. Arsalan, H. Gil Hong, R.A. Naqvi, M. B. Lee, M. C. Kim, D. S Kim, C. S. Kim, and K. R. Park. Deep learning-based iris segmentation for iris recognition in visible light environment. *Symmetry*, 9(11):263, 2017.

[43] H. Proenca. Iris recognition: On the segmentation of degraded images acquired in the visible wavelength. *TPAMI*, 32(8):1502–1516, 2010.

[44] A. Das, U. Pal, M. A. Ferrer, M. Blumenstein, D. Štepec, P. Rot, P. Peer, and V. Štruc. SSBC 2018: Sclera Segmentation Benchmarking Competition. In *ICB*, pages 303–308, 2018.

[45] A. Das, U. Pal, M. Blumenstein, and M. A. Ferrer Ballester. Sclera recognition-a survey. In *ACPR*, pages 917–921, 2013.

[46] M. Adam, F. Rossant, F. Amiel, B. Mikovicova, and T. Ea. Reliable eyelid localization for iris recognition. In *ACIVS*, pages 1062–1070. Springer, 2008.

[47] W. Aydi, L. Kamoun, and N. Masmoudi. A fast and accurate eyelids and eyelashes detection approach for iris segmentation. In *WCCIT*, pages 1–6, 2013.

[48] M. J. Aligholizadeh, S. Javadi, R. Sabbaghi-Nadooshan, and K. Kangarloo. Eyelid and eyelash segmentation based on wavelet transform for iris recognition. In *CISP*, volume 3, pages 1231–1235, 2011.

[49] V. Badrinarayanan, A. Kendall, and R. Cipolla. SegNet: A Deep Convolutional Encoder–Decoder Architecture for Image Segmentation. *IEEE TPAMI*, 39(12):2481–2495, 2017.

[50] A. Das, U. Pal, M. A. Ferrer, M. Blumenstein, D. Štepec, P. Rot, Ž. Emeršič, P. Peer, and V. Štruc. SSRBC 2017: Sclera Segmentation and Recognition Benchmarking Competition. In *IJCB*, pages 1–6, 2017.

[51] Ž. Emeršič, L. Gabriel, V. Štruc, and P. Peer. Convolutional encoder–decoder networks for pixel-wise ear detection and segmentation. *IET Biometrics*, 7(3):175–184, 2018.

[52] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *ICRL*, 2015.

[53] J. Long, E. Shelhamer, and T. Darrell. Fully Convolutional Networks for Semantic Segmentation. In *CVPR*, pages 3431–3440, 2015.

[54] Ž. Emeršič, D. Štepec, V. Štruc, and P. Peer. Training convolutional neural networks with limited training data for ear recognition in the wild. *B-Wild (FG Workshop)*, 2017.