

Detection of Glasses in Near-infrared Ocular Images

P. Drozdowski^{*†}, F. Struck^{*}, C. Rathgeb^{*} and C. Busch^{*}

^{*} da/sec - Biometrics and Internet Security Research Group, Hochschule Darmstadt, Germany

[†] Norwegian Biometrics Laboratory, NTNU, Gjøvik, Norway

{pawel.drozdowski, christian.rathgeb, christoph.busch}@h-da.de
florian.struck@stud.h-da.de

Abstract—Eyeglasses change the appearance and visual perception of facial images. Moreover, under objective metrics, glasses generally deteriorate the sample quality of near-infrared ocular images and as a consequence can worsen the biometric performance of iris recognition systems. Automatic detection of glasses is therefore one of the prerequisites for a sufficient quality, interactive sample acquisition process in an automatic iris recognition system. In this paper, three approaches (i.e. a statistical method, a deep learning based method and an algorithmic method based on detection of edges and reflections) for automatic detection of glasses in near-infrared iris images are presented. Those approaches are evaluated using cross-validation on the CASIA-IrisV4-Thousand dataset, which contains 20000 images from 1000 subjects. Individually, they are capable of correctly classifying 95-98% of images, while a majority vote based fusion of the three approaches achieves a correct classification rate (CCR) of 99.54%.

Keywords-Biometrics; Iris Recognition; Glasses Detection;

I. INTRODUCTION

In recent years, iris recognition has become a popular modality for biometric systems and is used in many large-scale deployments (e.g. the Indian National ID project [22]). The technology is also increasingly being used in automatic (without human operator supervision) systems, such as smart border/airport gates and mobile devices [14]. Operational systems typically capture iris images in the near-infrared light spectrum, in which the iris patterns are much more pronounced than in the visible light spectrum, even for darkly pigmented irides [6]. According to recent reports [20], [21], over 50% of adult population in the developed world wear eyeglasses. The pervasiveness of short-sightedness (myopia) has been on an extreme rise in Eastern Asia and around the world in general; a recent report in Nature News [7] states:

East Asia has been gripped by an unprecedented rise in myopia, also known as short-sightedness. Sixty years ago, 10-20% of the Chinese population was short-sighted. Today, up to 90% of teenagers and young adults are. In Seoul, a whopping 96.5% of 19-year-old men are short-sighted. Other parts of the world have also seen a dramatic increase in the condition, which now affects around half of young adults in the United States and Europe

- double the prevalence of half a century ago. By some estimates, one-third of the world's population - 2.5 billion people - could be affected by short-sightedness by the end of this decade.

Due to specular reflections, blur, scratches and other factors, glasses tend to decrease the biometric sample quality and consequently often the biometric performance of the systems. While several researchers have investigated the impact of glasses on face recognition systems, the scientific literature on iris recognition contains very little related work on this subject, except for a paper in which a small-scale quantification of the effects of glasses on iris image pre-processing is presented [13] and glasses being mentioned as a significant noise factor (e.g. [3], [1], [9]). ISO/IEC 29794-6 biometric sample quality standard [10] specifically recommends to instruct data subjects to remove glasses during acquisition or to perform the acquisition with additional care.

Therefore, and due to the prevalence of glasses in the world population, automatic detection of glasses is an important matter in iris recognition (as will also be substantiated by the experiments described in section III). It is of particular interest for automatic sample acquisition systems, where such a detection module would enable an interactive sample acquisition and thus facilitate higher sample quality. While this is a well-researched topic in systems working with images of the facial region (e.g. [23], [2]), doing so in images of ocular region alone has not received enough attention. In this paper, three methods for accomplishing said task are presented and benchmarked.

This paper is organised as follows: in section II, the used dataset and experimental setup are described. Section III provides an overview of the impact of glasses on iris recognition. In section IV the three proposed automatic glasses detection approaches are presented and evaluated. Concluding remarks are given in section V.

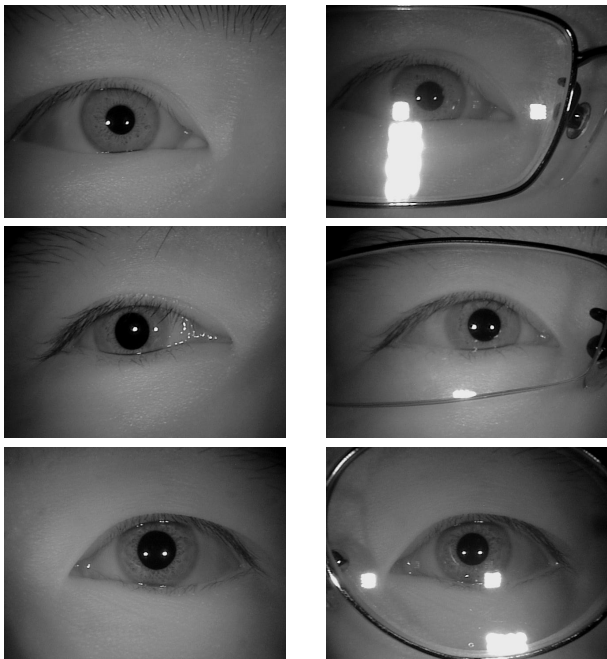
II. EXPERIMENTAL SETUP

The Thousand subset of the CASIA-IrisV4 database [5] (henceforth referred to as "CASIA-Thousand dataset") was chosen for the experiments performed for this paper. Said dataset contains near-infrared iris images of size 640×480 pixels and, due to its size, is suitable for large-scale testing.

Additionally, for subjects who are glass-wearers, it contains images both with and without glasses, thus enabling a direct biometric performance benchmark. Figure 1 shows example images from the dataset, while table I summarises its properties¹. Observe the high fraction of subjects who are glass-wearers coinciding with the statistics mentioned in section I. The groundtruth labels (with/without glasses) had to be assigned to all the images, which was done manually by a single researcher via visual inspection.

Table I: Overview of the CASIA-Thousand dataset

| | Samples | Subjects | Instances |
|------------------------|---------|----------|-----------|
| Total | 20000 | 1000 | 2000 |
| Without glasses | 14664 | 1000 | 2000 |
| With glasses | 5336 | 617 | 1193 |



(a) Without glasses

(b) With glasses

Figure 1: Example images from the CASIA-Thousand dataset. Samples (a) and (b) are captured from the same eye instance.

The images were processed with commonly used methods (specifically, Viterbi algorithm for segmentation [19], Daugman’s rubber sheet model for normalisation, LogGabor wavelet for feature encoding and fractional Hamming distance for template comparison [6]) implemented by the open-source OSIRIS [15] and USIT [16] frameworks. Subsequently, two evaluations took place:

¹Observe, that since for every subject/instance there is at least one sample without glasses in the dataset, the numbers for subjects/instances *seemingly* do not add up.

- The impact of glasses on sample quality (some metrics from ISO/IEC 29794-6 standard [10]) and thereby on iris recognition in terms of biometric performance measured in equal error rate (EER). (section III)
- The classification accuracy of the proposed detection approaches using cross-validation over 4 folds (i.e. 15000 training and 5000 test images), measured in correct classification rate (CCR). (section IV)

III. IMPACT OF GLASSES ON IRIS RECOGNITION

The topic of glasses in iris recognition systems has often been mentioned in the scientific literature (e.g. [3], [1], [9]) and presentations [18]. It is commonly agreed that they can have detrimental effect on sample quality due to specular reflections, dirt, optical distortions and shadows. A decrease in sample quality in turn negatively affects the segmentation accuracy and/or biometric performance. Furthermore, as shown in figure 2, they introduce potential for explicit failures, where the reflections or frame can be misunderstood as pupillary or limbic boundaries by the segmentation algorithm (the red blobs in the images represent areas masked out by the algorithm as eyelids and noise). Those assertions notwithstanding, with an exception of a small investigation [13], studies quantifying the effects glasses have on the biometric performance of iris recognition systems are lacking in the scientific literature.

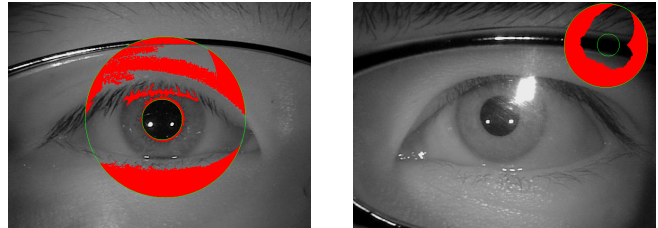


Figure 2: Segmentation failures caused by glasses

The results of a biometric verification experiment on the CASIA-Thousand dataset, shown in table II, demonstrate the negative impact of glasses on an iris recognition system. In addition to the data shown in the table, the motion blur in images with glasses was calculated to be twice as high as in images without glasses, which in turn can negatively affect other iris image quality metrics, such as the iris-pupil and iris-sclera contrast.

Table II: Impact of glasses on iris recognition

| Metric | Without glasses | With glasses |
|--|-----------------|--------------|
| EER, all images | 6.86% | 12.16% |
| EER, no segmentation failures | 3.79% | 10.67% |
| Images with usable iris area $\geq 70\%$ | 59.19% | 51.63% |

The aforementioned issues are also mentioned in the ISO/IEC 29794-6 biometric sample quality standard [10], where it is recommended to perform data acquisition so

that the specular reflections on the iris are minimised or even to instruct the data subject to remove their glasses. In some, particularly automatic systems, doing so would require automatically detecting the glasses. In the next section, methods of automatic detection of glasses in near-infrared iris images are described and evaluated.

IV. AUTOMATIC DETECTION APPROACHES

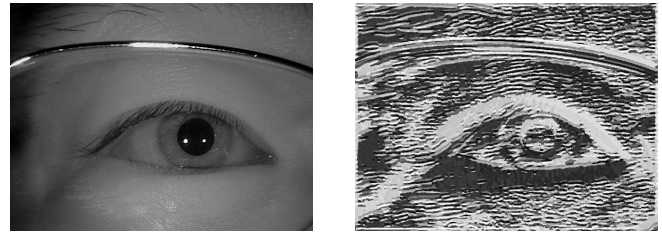
As discussed earlier, automatic detection of glasses appears to be an overseen or underappreciated issue in the scientific literature. However, based on the sheer numbers of glass-wearers in the population (section I) and the significant impact of glasses on the biometric performance (section III), it is abundantly clear that methods for automatic glasses detection are beneficial for iris recognition systems. With it in place, such systems would be enabled to provide actionable feedback to the capture subject - meaning to ask the subject to take off the glasses and to subsequently initiate a re-capture. In this section, three such approaches are presented and evaluated on near-infrared iris image data.

A. Texture Descriptor

Binarized statistical image features (BSIF) [12] is a generic texture descriptor, which uses filters learned from patches of natural images. Pre-trained filters made available as part of the above publication are used. The process of using BSIF to detect glasses in iris images is as follows:

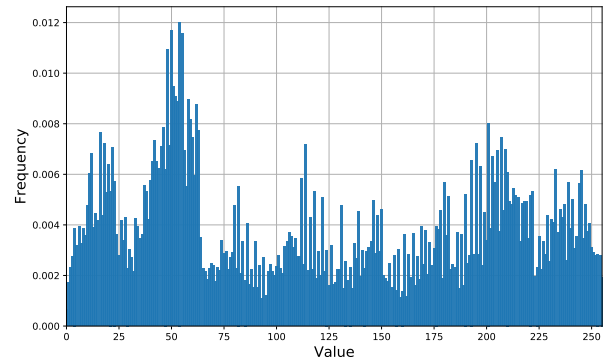
- 1) An input image (figure 3a), is convolved with 8 stacked linear filters of size 15×15 pixels; the sign of each filter response is used to binarise it (such that negative responses become 0 and positive responses become 1), resulting in a binary string of length 8 for each pixel of the image. The integer representation of those binary strings lies in range (in range 0 to $2^8 - 1$), and can be thus displayed as a 256-bit grayscale image, as shown in figure 3b.
- 2) The aforementioned integer values for the whole image are stored in a histogram, as shown in figure 3c.
- 3) Using the previously (section II) mentioned cross-validation loop for training and testing, the classification decision is obtained by passing the histogram as an input to a support vector machine (SVM). It uses a linear kernel, which is suitable for high-dimensional vectors. A lightweight implementation provided by the libsvm (version 3.22) [4] library was used; for training the SVM, 1000000 was used as cost parameter and 0.001 was used for termination tolerance. The parameters were estimated empirically on a small, disjoint training set.

As shown in figure 3d, subtle differences in the BSIF-histogram values' frequency distribution are perceivable. The SVM is capable of using those to distinguish between and correctly classify images with and without glasses. To

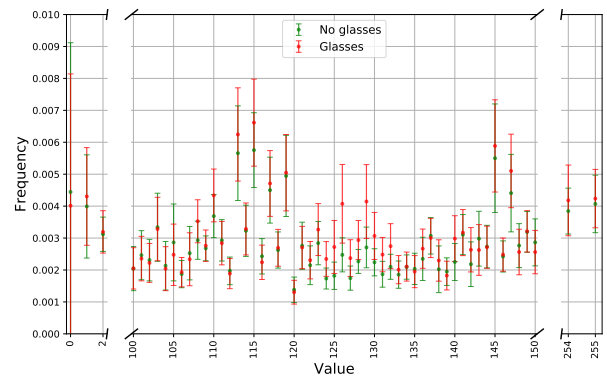


(a) Image

(b) BSIF applied



(c) BSIF grayscale values histogram for one image



(d) BSIF histogram value frequencies (mean and one standard deviation as errorbars) for the entire CASIA-Thousand dataset

Figure 3: BSIF-based approach

facilitate reproducible research, the trained SVM is released publicly at [8].

B. Deep Learning

Deep neural networks convince by successful application with a huge variety of tasks [17], including classification specifically [24]. The problem of classifying images with and without glasses falls well within the areas in which deep neural networks are commonly applied.

Using the Caffe framework (version 1.0) [11], a deep convolutional neural network for classification of images has been created; its topology can be seen in table III. The neural network is trained and tested using the previously (section II) mentioned cross-validation loop. The images are resized to

320×240 pixels and the training is run over 20000 iterations, with batch size of 32 images and 20000 as step size. Using 15000 input images and 5 steps, the network was trained for about 213 epochs. The learning rate is set to 0.0001 and gets multiplied by 0.25 after every step.

Multiple other architectures, which differed mostly in the input dimensions and the size of the convolution layers, were tested. It turned out that input dimensions larger than 320×240 (e.g. 640×480) are not necessary to achieve good classification results. Thus, for computational performance reasons the relatively small network was chosen in order to attain an acceptable trade-off between classification accuracy and throughput. The dimensions of the convolution layers were determined by the size of potential feature blocks, which are effected by glasses being present in an image. To facilitate reproducible research, the trained network is released publicly at [8].

Table III: Topology of the DNN-based approach

| Part | Layer | Iterations | Details |
|--------------------|-----------------|------------|---------------------------------------|
| Feature Extraction | Convolution | 2 | 1. 17 × 17 pixels, 48 filters |
| | ReLU | | 2. 7 × 7 pixels, 96 filters |
| | Pooling | | — |
| | | | Max, 3 × 3 pixels, 1 filter, stride 2 |
| Classification | Fully connected | 2 | 96 neurons |
| | ReLU | | — |
| | Dropout | — | |
| | Fully connected | 1 | 2 neurons |
| | Decision | | linear classifier |

C. Edge and Reflection Detection based Algorithm

Two key differences between images with and without glasses are more pronounced specular reflections and stronger edges due to the frames of glasses. The classification approach described in this subsection is based on detection and quantification of those image features.

1) *Reflections*: The image is divided into blocks of equal size (chosen empirically to be 30 × 30 pixels and the brightness of each block is computed relative to the brightness of the entire image, thus producing a map of relative brightness deviation. The block size filters out small, natural reflections (figure 4a), whereas large, artificial reflections are very well pronounced (figure 4b).

2) *Edges*: The process of detecting and measuring edges for glasses detection in an iris image is described below and shown in figure 5 and described below.

- 1) The image is convolved with a simple kernel which detects horizontal edges. This process is independent of the average brightness of the image, since only the local brightness gradients are computed. (figure 5b)
- 2) The grayscale image is transformed into a black and white image. This is done by applying a brightness threshold (usually between 128 and 129, estimated empirically on a small disjoint training set), which only accepts sharp brightness transitions and ignores blurred edges. (figure 5c)

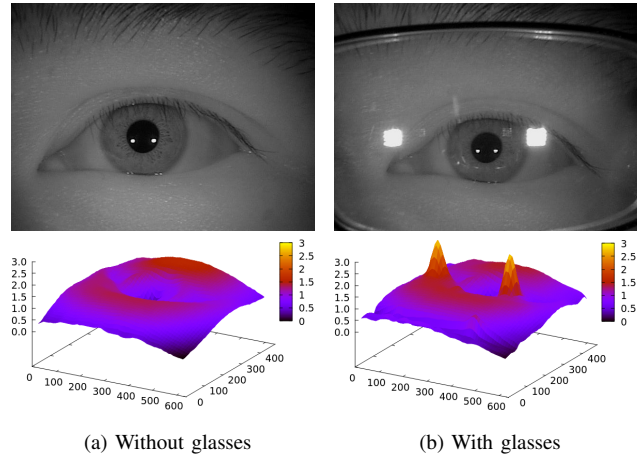
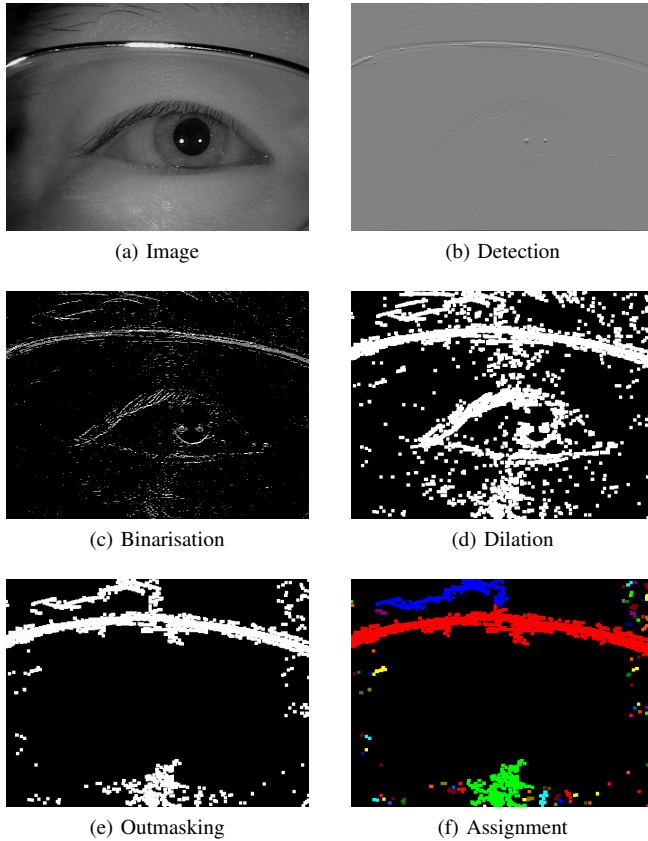


Figure 4: Reflection detection with a relative brightness measure. The two specular reflections caused by the glasses are clearly observed by this proposed metric.

- 3) Due to illumination artefacts or image compression many edges have small gaps. A dilation filter of size 7 × 7 pixels (estimated empirically on a small disjoint training set) is used to fill those gaps. (figure 5d)
- 4) The edges in the middle of the image are masked out, since they tend to be natural eye edges. (figure 5e)
- 5) To distinguish between individual edges, the flood fill algorithm with 8 directions is applied. This algorithm finds connected pixels and represents them with different colours. (figure 5f)
- 6) The width and height of the found edges is calculated using the leftmost and rightmost, and topmost and bottommost pixels. Very small edges are discarded (e.g. the small points on the right top corner in figure 5f) are discarded because they do not contain information. Subsequently, then the ratio between widths and heights of the remaining edges is computed. (figure 5g)

3) *Classification*: The reflection and edge detection methods described in subsections IV-C1 and IV-C2, respectively, are applied to an iris sample. Using the previously (section II) mentioned cross-validation loop, tuples containing the values of largest relative brightness block and the edge with the highest width-to-height ratio are passed to a SVM, which performs the classification decisions. A radial basis function (RBF) kernel was chosen, since it is well suited for low-dimensional vectors. For training the SVM, 10000 was used as cost parameter, which was estimated using a small, disjoint training set. As shown in figure 6, the two metrics (reflection and edge scores) contain sufficient discriminative power to distinguish quite accurately between images with and without glasses, albeit some overlap (and thereby classification errors) is still present.



(g) Measurement

Figure 5: Edge detection and measurement

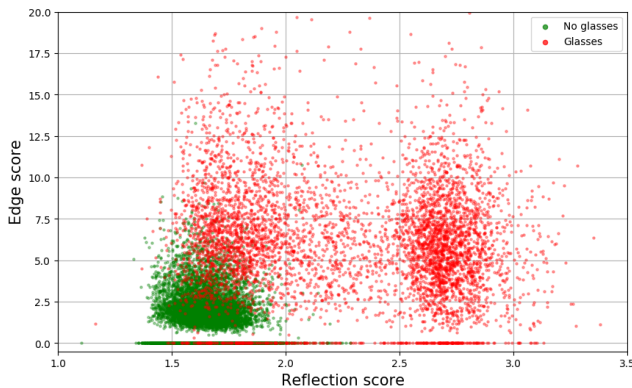


Figure 6: A scatter plot of edge and reflection scores for all images from the CASIA-Thousand dataset, which shows significant separation between the two image classes

D. Results

The classification accuracy of the proposed methods is estimated by performing cross-validation over 4 folds. The results are shown in table IV. All three proposed approaches perform well, with overall accuracy ranging between 95-99%. Notice, however, that the CCR for images with and without glasses vary - for instance, the neural network classifies more images without glasses correctly, whereas the statistical approach does so for images with glasses. This suggests a possibility of fusing the decisions of the approaches, so that their individual weaknesses are compensated for. Performing a majority vote of all three approaches was able to significantly increase the CCR. A conjunction based fusion of all three or different configurations of two approaches was also tried, but was found to be less successful than the majority vote (albeit still improving upon the accuracy of the individual approaches).

Table IV: Results of the evaluation (with 95% CI)

| Approach | CCR (in%) | | |
|------------------------------|-----------------|--------------|--------------|
| | Without Glasses | With glasses | Overall |
| Texture Descriptor (IV-A) | 97.79 ± 0.95 | 98.54 ± 0.69 | 98.08 ± 0.44 |
| Deep Learning (IV-B) | 99.28 ± 0.22 | 97.33 ± 1.60 | 98.97 ± 0.29 |
| Edges and Reflections (IV-C) | 97.18 ± 0.38 | 92.37 ± 2.23 | 95.43 ± 0.36 |
| Majority vote | 99.72 ± 0.08 | 98.79 ± 0.66 | 99.54 ± 0.12 |

E. Classification Errors

It is of interest to investigate what types of images were incorrectly classified by the proposed approaches. Figure 7 shows such example images and corresponding error reasons. With a larger dataset and hence more training data, the classification errors could potentially be further reduced.

V. CONCLUSION

Glasses make iris recognition more challenging, since they can have a detrimental effect on sample quality and thereby biometric performance of a system. In section III, it has been shown that on the CASIA-Thousand dataset, the equal error rate on the subset of images with glasses is twice that of the subset of images without glasses. It is therefore of interest to automatically detect glasses in iris images in order to handle such images separately or re-acquire once the data subject has been asked to remove their glasses. In this paper, three approaches for automatically detecting glasses in near-infrared ocular images have been presented. They achieve classification accuracy in range of 95-98%, which can be further improved on by a decision-level fusion. A majority vote of all three approaches achieved an overall 99.54% correct classification rate, whereas slightly lower (but still above 99%) correct classification rate was achieved with an conjunction-based fusion of two approaches. In contrast to other approaches for glasses detection, the proposed methods require only a single-frame image and work with the ocular area alone instead of whole face. They could be seamlessly

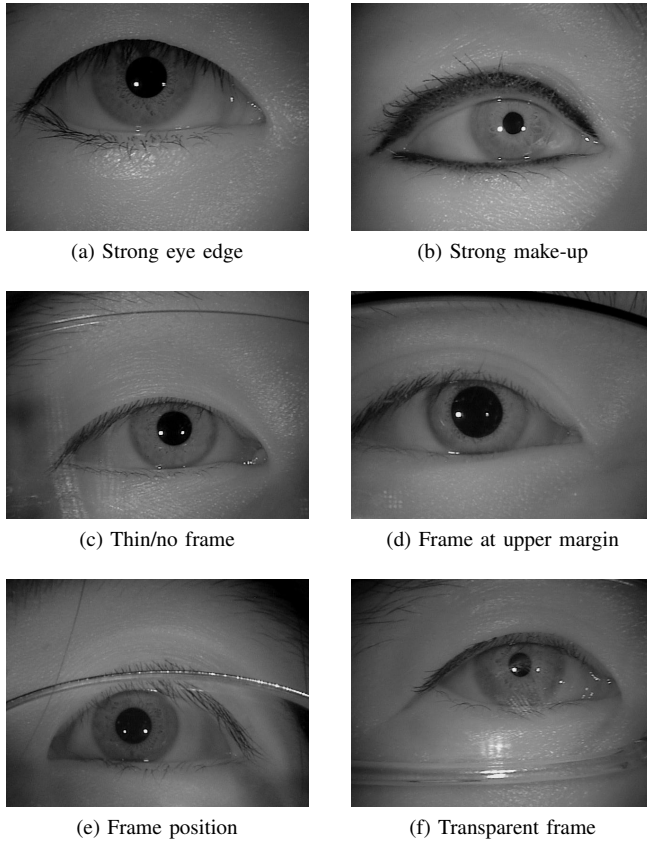


Figure 7: Examples of incorrectly classified images from all 3 methods. Figures (a)-(b) falsely classified as glasses, figures (c)-(f) falsely classified as non-glasses.

integrated into operational automatic systems, for instance to facilitate interactive image acquisition, where the data subjects would be required to take off glasses if detected. Furthermore, such systems often capture images of both eyes simultaneously, thus the accuracy of glasses detection could be further improved by performing a multi-instance fusion, i.e. a conjunction of the decisions from both eyes.

ACKNOWLEDGEMENTS

This work was partially supported by the German Federal Ministry of Education and Research (BMBF) as well as by the Hessen State Ministry for Higher Education, Research and the Arts (HMWK) within Center for Research in Security and Privacy (CRISP).

REFERENCES

[1] S. Bharadwaj, H. S. Bhatt, M. Vatsa, and R. Singh. Periocular biometrics: When iris recognition fails. In *Intl. Conf. on Biometrics: Theory, Applications and Systems*, pages 1–6, September 2010.

[2] D. Borza, A. S. Darabant, and R. Danescu. Eyeglasses lens contour extraction from facial images using an efficient shape description. *Sensors*, 13(10):13638–13658, October 2013.

[3] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn. Image understanding for iris biometrics: A survey. *Computer Vision and Image Understanding*, 110(2):281–307, May 2007.

[4] C. Chang and C. Lin. LIBSVM: A library for support vector machines. *Trans. on Intelligent Systems and Technology*, 2:27:1–27:27, 2011. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.

[5] Chinese Academy of Sciences’ Institute of Automation. CASIA iris image database. <http://biometrics.idealtest.org/>, December 2010. Last accessed: 2017-05-18.

[6] J. Daugman. How iris recognition works. *Trans. on Circuits and Systems for Video Technology*, 14(1):21–30, January 2004.

[7] E. Dolgin. The myopia boom. *Nature*, 519(7543):276–278, March 2015.

[8] P. Drozdowski, F. Struck, C. Rathgeb, and C. Busch. Glasses detection models. <https://github.com/dasec/glasses-detection-models>, 2017.

[9] H. Proença and L. A. Alexandre. Iris recognition: Analysis of the error rates regarding the accuracy of the segmentation stage. *Image and vision computing*, 28(1):202–206, 2010.

[10] ISO/IEC JTC1 SC37 Biometrics. *ISO/IEC 29794-6:2015. Information technology – Biometric sample quality – Part 6: Iris image data*. International Organization for Standardization and International Electrotechnical Committee, July 2015.

[11] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, et al. Caffe: Convolutional architecture for fast feature embedding. In *Intl. Conf. on Multimedia*, pages 675–678, November 2014.

[12] J. Kannala and E. Rahtu. BSIF: Binarized statistical image features. In *Intl. Conf. on Pattern Recognition*, pages 1363–1366, November 2012.

[13] S. Lim, K. Lee, O. Byeon, and T. Kim. Efficient iris recognition through improvement of feature vector and classifier. *ETRI Journal*, 23(2):61–70, June 2001.

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

[15] N. Othman, B. Dorizzi, and S. Garcia-Salicetti. OSIRIS: An open source iris recognition software. *Pattern Recognition Letters*, 82(2):124–131, September 2016.

[16] C. Rathgeb, A. Uhl, P. Wild, and H. Hofbauer. Design decisions for an iris recognition SDK. In *Handbook of Iris Recognition*. Springer, 2nd edition, 2016.

[17] J. Schmidhuber. Deep learning in neural networks: An overview. *Neural networks*, 61:85–117, 2015.

[18] Z. Sun. Lecture: Recent progress of iris recognition. http://www.comp.hkbu.edu.hk/wsb17/lecturer_details.php?lect_id=8, January 2017. Last accessed: 2017-09-06.

[19] G. Sutra, S. Garcia-Salicetti, and B. Dorizzi. The Viterbi algorithm at different resolutions for enhanced iris segmentation. In *Intl. Conf. on Biometrics*, pages 310–316, March 2012.

[20] The European Council of Optometry and Optics. ECOO Blue Book: Data on optometry and optics in Europe, May 2017.

[21] The Vision Council. VisionWatch: The Vision Council member benefit report, September 2016.

[22] Unique Identification Authority of India (UIDAI). Aadhaar issued summary. <https://portal.uidai.gov.in/uidwebportal/dashboard.do>, 2010. Last accessed: 2017-05-18.

[23] C. Wu, C. Liu, H. Shum, Y. Xy, and Z. Zhang. Automatic eyeglasses removal from face images. *Trans. on pattern analysis and machine intelligence*, 26(3):322–336, March 2004.

[24] G. P. Zhang. Neural networks for classification: A survey. *Trans. on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 30(4):451–462, 2000.