© IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.

Mobile NIR Iris Recognition: Identifying Problems and Solutions

Heinz Hofbauer[1] • Ehsaneddin Jalilian[1] • Ana F. Sequeira[2] • James Ferryman[2] • Andreas Uhl[1]

[1]Multimedia Signal Processing and Security Lab, University of Salzburg, Austria, {hofbauer, ejalilian, uhl}@cs.sbg.ac.at
[2]Computational Vision Group, University of Reading, United Kingdom, {a.f.p.sequeira,j.m.ferryman}@reading.ac.uk

Abstract

The spread of biometric applications in mobile devices handled by untrained users opened the door to sources of noise in mobile iris recognition such as larger extent of rotation in the capture and more off-angle imagery not found so extensively in more constrained acquisition settings. As a result of the limitations of the methods in handling such large degrees of freedom there is often an increase in segmentation errors. In this work, a new near-infrared iris dataset captured with a mobile device is evaluated to analyse, in particular, the rotation observed in images and its impact on segmentation and biometric recognition accuracy. For this study a (manually annotated) ground truth segmentation was used which will be published in tandem with the paper. Similarly to most research challenges in biometrics and computer vision in general, deep learning techniques are proving to outperform classical methods in segmentation methods. The utilization of parameterized CNN-based iris segmentations in biometric recognition is a new but promising field. The results presented show how this CNN-based approach outperformed the segmentation traditional methods with respect to overall recognition accuracy for the dataset under investigation.

Contents

1 Introduction

Mobile devices have become ubiquitous and the typical cell phone is much more than merely a phone, it is frequently used to checking emails and bank accounts, accessing web services and enterprise resources and even monitoring home appliances. As a result, vulnerabilities that grant access of devices to unauthorized users have a huge impact [1]. PIN/Password, biometrics and unlock patterns are all equally accepted forms authentication for mobile devices [2]. However, PIN/Password and unlock patterns are subject to a number of attacks due to the mobile natures of the device, e.g. "shoulder surfing" [3], compared to devices located in a secure environment, e.g. personal computers. Biometric attacks are more robust to such attacks since they are harder to clone and offer high security.

Among the possibilities of performing authentication using biometrics, iris is a strong candidate due to its proven reliability among the other traits available [4, 5]. However, the use of iris recognition on mobile phones poses other challenges when compared to conventional iris recognition such as restricted computational power, suited Near Infrared (NIR) illumination and less constrained thus more difficult acquisition conditions. Mobile iris recognition systems can be divided into three main categories namely [6]: Systems using dedicated devices to perform iris recognition, systems connecting additional hardware to the mobile device, and systems attaching NIR cameras with illuminators.

Despite its promises and the already extensive literature, performing accurate iris recognition in mobile devices is still considered a challenge. The first published work on iris recognition on mobile phones was by Cho et al. [7] who proposed a novel pupil and iris localization technique and implemented those on low processing power mobile phones. Jeong et al. proposed a method where feature extraction was carried out using 1-D adaptive Gabor filter (adapted to four different image acquisition scenarios indoor, no-blur; indoor, blurred images; outdoor, no-blur images; and outdoor, blurred images)[8]. In a following work, Park et al. [9] added dual IR illuminators to Jeong et al.'s modified mobile device introducing an operating range of 35-40 cm and the capture of dual eye regions. In 2006, a major breakthrough was made by the announcement of iris OKI scanners using OKI's original iris recognition algorithm for Symbian and Windows mobile operating systems. Kurkovsky, Carpenter and McDonald proposed an approach which required no additional hardware to adapt iris recognition on resource constrained mobile phones [10]. Kang proposed to use two magnetic rings to hold the lens in mobile phones instead of using close-up lenses [11]. Lu et al. developed a smartphone iris recognition system assuming that all eye images are captured at the same illumination with same stand-off distance from the camera [12] by providing an 'eye cup' with a luminous diode attached. The common characteristics of the so-far mentioned techniques are the fact that image acquisition is constrained and there is an irreversible modification of the smartphone camera for iris recognition. The path forward for mobile iris recognition was to detach from hardware adaptations and rely on the imaging power of mobile phones in the visible spectrum. The MICHE-I and II competitions [13–15] pushed forward the research on visible wavelength (VIS) mobile iris detection and recognition. However, the negative effect of VIS imaging on the recognition capability of the iris trait [16, 17] has moved the research in the field to the NIR imaging context again. This renewed focus on iris recognition in mobile devices in the near infra red spectrum was facilitated by the appearance of novel NIR enabled mobile devices, i.e. Samsung S8/9, S8+/9+ and Note 8/9, Microsoft Lumia 950 (XL), and local solutions like Fujitsu NX F-04G in Japan or Iritech and IrisGuard in India. The results observed in literature confirm early studies which noted that NIR iris recognition outperforms its visible counterpart [18] and therefore, several recent works propose methods for NIR imaging in the mobile environment [19–22], also supported by the release of corresponding first public datasets (e.g. CASIA-Iris-M1-S3).

In this work, we tackle the issue of more difficult acquisition conditions in case of mobile capturing and show results on a new NIR mobile iris dataset. In particular, with mobile acquisition iris imagery potentially exhibits a larger extent of rotation and off angle shots due to misalignment of facial plane and sensor plane. Besides requiring more rotation compensation, the increase in rotation might throw off traditional segmentation methods, which often rely on the more or less horizontal alignment of eye-lids. Therefore, we propose to use CNN-based semantic segmentation techniques for more reliable iris segmentation results on such data.

In Section 2, we describe the applied CNN-based iris segmentation and propose a way to use the resulting binary segmentation masks to generate normalised iris texture (applying the rubbersheet transform). Section 3 evaluates the proposed approach with respect to (i) iris segmentation accuracy compared to manual segmentation ground truth and (ii) iris recognition accuracy comparing the approach to traditional, open source segmentation and recognition techniques (USIT) as well as to the OSIRIS toolkit. Additionally, we present results with respect to the required rotation compensation in the matching process. Section 4 concludes this paper and gives outlook to future work.

2 CNN-based Iris Segmentation

Application of convolutional neural networks for iris segmentation has recently received some research attention, and a few CNN-based models were proposed for this task [23]. Historically, the first approach [24] proposed a CNN classifier composed of three blocks of alternative convolution and pooling layers, which takes in a local patch around pixels to extract features, and labels each pixel separately using a final fully connected layer (FC), however, this model lacks speed and efficiency. The authors proposed another model, which includes six blocks of interconnected convolution and pooling layers followed by a single multiplication and a softmax layer. The somewhat restricted empirical assessment makes the results difficult to interpret.

A posterior work [23] used deep convolutional encoder-decoder networks in different variants, i.e. iFCEDN and derivatives, and trained them on manually annotated iris segmenta-

tion masks. The iFCEDN network comprises an encoder part and a decoder part and the 'Basic' and 'Bayesian' networks are just reduced versions of the same network with 'Bayesian' enabling the probabilistic pixel-wise segmentation using Monte-Carlo sampling and drop-out techniques. Segmentation results turned out to be excellent compared to traditional segmentation and in another work [25], the authors used the same 'Basic' network and proposed a domain adaption model for iris images, eliminating the need for generating manually annotated training masks for each iris dataset / sensor separately.

The method proposed in [26] focuses on the segmentation of visible light unconstrained (on-the-move) iris recordings. The CNN used specifically focuses on the segmentation of the iris and requires a prior ROI detection and then adapted a pre-trained VGG network as one-class (iris pixel) classifier.

CNN based iris segmentation algorithms generate a binary mask, separating iris from non-iris pixels. None of the approaches described did go beyond segmentation and none proposed any model for using the iris masks in the actual recognition process. The obtained binary mask is highly accurate on the one hand, but cannot be directly used for extracting and normalizing the iris texture for further use in a biometric system on the other hand. For generating the normalised texture using the rubbersheet transform, a parameterization of the pupillary and limbic boundary are required. Both the CNN-based segmentation and parameterization to utilize the CNN-based mask in a biometric system are described in the following sections.

2.1 CNN-based Mask Generation

Besides using iFCEDN as proposed in [23, 25], we use a more recent semantic segmentation CNN called "RefineNet" [27]. It is a multi-path refinement network, which employs a 4-cascaded architecture with 4 RefineNet units, each of which directly connects to the output of one Residual net [28] block, as well as to the preceding RefineNet block in the cascade. Each RefineNet unit consists of two residual convolution units (RCU), whose outputs are fused into a high-resolution feature map, and then fed into a chained residual pooling block. The final part of each RefineNet block is another residual convolution unit.

The RefineNet architecture was chosen in addition to the improved FCEDN [23] (iFCEDN) for its higher segmentation performance on common public iris datasets, i.e., higher F-measure of generated masks compared to a mask produced by human annotation (ground truth), as shown in Table 3. For information about the ground truth and data sets refer to Section 3.

The binary iris segmentation masks are created by CNNs trained with the manually annotated ground truth mask. There is no overlapping of training and testing sets since a five-fold cross-validation method was used. To do so, initially each database was partitioned into five equally sized subsets. Of the five subsets, a single subset was retained as the testing data, and the remaining four subsets were used as training data. The cross-validation process was then repeated five times, with each of the five subsets used exactly once as the testing data.

Table 1: The training parameters for the network.

Parameter	Batch-size	Epoch	Momentum	Learning rate
Values	8	1000	0.1	0.9

Further relevant parameters for training the RefineNet network are listed in Table 1 while for iFCEDN, identical parameters as in [23] are used.

2.2 Parameterizing the Masks for Normalization

Given binary segmentation masks as produced by CNN semantic segmentation, the outline of the parameterization process for the limbic and pupillary boundaries is as follows: (1) preprocessing by median blur to smooth out the edges, (2) generation of candidate segmentations using a circular Hough transform [29], and (3) selecting the best candidate as a final parameterization. An example for the pupillary parameterization is given in Figure 1, the annuli (see below) belonging to the chosen pupillary parameterization are also shown.

If the core parameters for the Hough transformation are set sanely, i.e. pupillary radius no larger than 1/4th of the smallest image dimension and iris radius greater than pupillary radius, then the resulting circles almost certainly contain the correct parameterization. The selection process then is the key for finding well fitting and stable parameterizations.

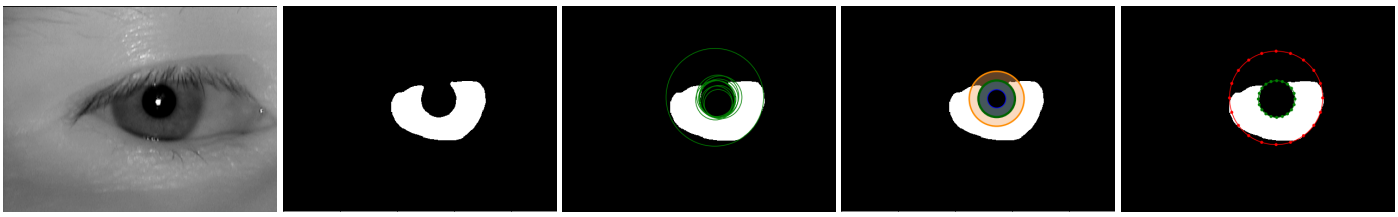
Given that the pupillary boundary is more complete than the iris boundary, because it is less obfuscated by eyelids and eyelashes, we start with that. The goal is to find the best parameterizations for which the parameterized circle fulfills the following properties: (a) surrounded by iris pixels on the outside, and (b) containing no iris pixels inside. We can use annuli adjacent to the circle candidate c, with center at c_x, c_y and radius c_r, and measure the number of iris pixels contained therein. Let the inner annulus be $A_i(c)$ with radii c_r and $0.5{\times}c_r$, and likewise the outer annulus $A_o(c)$ with radii $(1.5 \times c_r, c_r)$. Let $N(a)$ count the number of iris pixels in area a and $\widehat{N(a)}$ be maximum possible number of iris pixels in area a. We can then use this, pupillary, value functions as a, normalized in $[0, 1]$, quality measure for each circle:

$$V^p(c) = \frac{N(A_o(c)) - N(A_i(c))}{\widehat{N(A_o(c))}}. \quad (1)$$

The circle candidate with the highest value will be the circle we choose as parameterization for the pupil c^p. If two candidates have the same value the one with the smaller radius shall be chosen, in order to maximize the available iris texture.

The selection of the limbic circle is done in an analogous fashion, but changing the role of outer and inner annulus, since the limbic boundary should contain all the iris pixels. The largest radius candidate should be used in case of multiple candidates having the same value. There is an additional constraint which can be added to the selector function. The limbic and pupillary parameterizations should be roughly concentric. This leads to the limbic value function

$$V^l(c, c^p) = \max\Bigg(0, \frac{N(A_i(c)) - N(A_o(c))}{\widehat{N(A_i(c))}} - \frac{\sqrt{(c_x - c_x^p)^2 + (c_y - c_y^p)^2}}{c_r}\Bigg). \quad (2)$$



(a) Iris image (b) Mask image (c) c^p Candidates (d) c^p, $A_i(c^p)$ and $A_o(c^p)$ (e) Final parameterization

Figure 1: Outline of the parameterization process.

The lower cap of zero needs to be in place to keep it normalized in a $[0, 1]$ range, since of course $\widehat{N(A_o)} > \widehat{N(A_i)}$, and we further reduce the score by the relative radial offset. The cases where V^i would be less than zero are clearly not good parameterization candidates and as such the loss in differentiation does not impact the result. Like above, we choose the highest value circle c^l as the selected parameterization of the limbic boundary, and normalization is then done by transformation to Faberge coordinates in the regular fashion with $P := c^p$, $L := c^l$ (i.e. the rubbersheet transform).

3 Experiments

3.1 Tools and Datasets

For comparison with tools from literature we utilize the USIT toolkit, described in [30, 31]. Specifically, USIT contains the *CAHT* [32] and *WAHET* [33] segmentation tools, the *lg* and *qsw* feature extraction tools, and `hd` and `gen_stats_np.py` for biometric comparison and the evaluation of the results. Further, it contains the `manuseg` tool which can utilize parameterised pupillary and limbic boundaries, as generated by the algorithm described in Section 2.2 or provided by the various ground truth segmentations.

The databases used for comparison are the well known IIT Delhi Iris Database version 1.0[1] (*iitd*), and the interval subset of the CASIA Iris Image Database version 4.0[2] (*casia4i*). In addition we used a subset of CASIA Iris Subject Ageing Version 1.0 Database (*casiaA*)[3]. The ground truth for the segmentation is available for all used databases: For *iitd* and *casia4i* refer to [34][4]; and for *casiaA* refer to [35][5].

The central experimental dataset used in this work is the iris subset of the second part of the *PROTECT Multimodal DB* [6] [36], which has been acquired in a context focused on border control.

While *iitd*, *casia4i* and *casiaA* are well known databases the PROTECT Multimodal DB is relatively new and so we will briefly describe it here.

The *PROTECT Multimodal DB* aims at being representative of the universe of travelers that cross the borders thus including a wide range of variety in age, gender, ethnicity and skin/eye colour types. The data was collected during a single week in two different environments at the same location. These different environments were created to simulate the Air and Sea Border (ASB) and the Land Border (LB) use cases. In an indoors location, a "biometric corridor" acquisition was created to collect 2D face and periocular, 3D face and anthropometrics. In an outdoors location, a "vehicle collection" environment was created to simulate the situation when a vehicle approaches the land border and the passengers need to have the biometric data captured whilst inside the car. The samples collected in the vehicle setup were iris, 2D face, 3D face, thermal face, and finger & hand veins. The iris subset (***protMI***) contains data of the two eye patterns from 28 subjects (56 different irises) and was acquired inside a vehicle using a mobile device by the passenger. Each volunteer provided 4 samples in the enrolment session and a variable number of samples in one or two additional collections. Therefore, the number of samples per subject is variable starting at a minimum of 10 samples (4 samples in the enrolment plus 6 samples in the first session). In the end a total of 1008 images from 56 iris were recorded, the number of images per eye finally used are given in Table 2 (few images are discarded in experimentation, see below).

The device used for the acquisition of iris samples was the MK212OU from the *IriShield*™series by IriTech which includes ultra-compact, auto-capture iris scanners, complete with onboard iris recognition and a PKI-based security infrastructure that ensures end-to-end data security. This sensor features iris image quality assessment algorithms to provide good quality images. The *IriShield* MK212OU device works at an optimal distance of 5 cm (2 inches) from a single eye and has a focal depth of 6 mm (0.2 inch). The image format is compliant to the ISO Standard 19794-6 (2005 & 2011) with 640 x 480 pixels and 8-bit grayscale. Furthermore, it offers a NIR LED illumination and it can be used outdoors or indoors. The device is depicted in Figure 2.

Despite the good quality of the iris images in principle, the images reflect some of the limitations observed when a self acquisition is performed using a handheld device plus the fact that this was done inside a vehicle in an outdoors setup. Some samples are depicted in Figure 3.

The environmental conditions, the mobile capturing device and the fact that acquisition simulated untrained users (non-habituated acquisition) lead to an increase in rotation in the data. This fact is at the core of our evaluations, and the other databases used in experimentation represent a less rotated base for comparison. The ground truth iris segmentations for *protMI*, based on one operator with ellipse+polynomial con-

[1] http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm
[2] http://biometrics.idealtest.org/dbDetailForUser.do?id=4
[3] http://biometrics.idealtest.org/dbDetailForUser.do?id=14
[4] http://www.wavelab.at/sources/Hofbauer14b
[5] http://www.wavelab.at/sources/Hofbauer16d
[6] http://www.projectprotect.eu/dataset/

Figure 2: The *Irishield* device connected to a mobile phone (Samsung S8).

tours like in [34], will be published jointly with this paper and will be made available at the Wavelab homepage[7].

It should be noted that we excluded 21 images from the *protMI* ground truth database, and consequently from our experiments. The reason is the usage of the parameterization types of the `manuseg` tool, which uses polynomials to delimit upper and lower eyelid. In case an image is rotated so far that the upper or lower eyelid can no longer be described by this function it has to be skipped. An example of an image removed from the database is given in Figure 4. The number of (remaining) images per eye is given in Table 2.

3.2 Experimental Results

3.2.1 CNN-based Segmentation

The results provided by the proposed segmentation method are depicted in Table 3 based on type 1 (E1) and type 2 (E2) errors (as used in the noisy iris challenge evaluation NICE - I [37], for which lower is better) and the F1-measure (for which higher is better). The binary segmentation masks as generated by the CNN semantic segmentation techniques are compared to the manually annotated ground truth.

Results reveal an interesting tendency. While the RefineNet architecture is able to outperform the already excellent iFCEDN results [23] for all datasets acquired in non-mobile acquisition, the *protMI* data obviously can be segmented better using the iFCEDN architecture. These results underpin the importance of carefully selecting the employed CNN architecture for a specific iris target dataset. The existence of an optimal architecture independent of iris dataset properties seems to be highly questionable.

[7] `http://www.wavelab.at/sources/Hofbauer18b`

Table 2: The makeup of the *protMI* database which contains a total of 1008 images from 28 users. The user ID, number of images from the left and right eye are given.

UID	left	right	UID	left	right
01	11	12	18	25	24
02	23	18	20	16	16
03	23	22	21	16	16
05	23	20	23	23	23
06	25	24	24	16	16
07	23	25	25	16	16
09	16	16	27	16	16
10	16	16	28	16	16
11	24	24	29	33	29
12	24	25	30	7	8
13	24	24	31	10	11
15	16	16	37	10	14
16	8	8	38	13	13
17	16	16	39	16	19

Table 3: Average segmentation scores per dataset.

CNN	Database	E1	E2	F1
RefineNet	*casia4i*	0.009	0.011	0.984
	casiaA	0.005	0.012	0.972
	iitd	0.015	0.018	0.974
	protMI	0.044	0.151	0.746
iFCEDN [23]	*casia4i*	0.021	0.028	0.962
	casiaA	0.007	0.020	0.966
	iitd	0.018	0.022	0.970
	protMI	0.008	0.024	0.952

3.2.2 Rotation Compensation

The eye can rotate in the eyesocket, torsional movement is induced by sup/inf rectus and sup/inf olique muscles. This torsional movement is generally limited to ranges of $\pm 10°$, Sparks [38] and Young and Sheena [39].

The USIT toolkit can deal with rotation, and the usual setting is $\pm$16bits. USIT uses textures of width, that is transformed circumference of the iris, of 512 bits, leading to a single bit roughly equaling 0.7° of rotation. This compensation of rotation is costly in terms of biometric recognition accuracy as it searches for better matches of imposter scores, narrowing the gap between genuine and imposter distributions. Furthermore, it is time consuming, a multiplication of a single match by the number of rotations. Both of these disadvantages of classical rotation compensation are known and there are works in literature which aim at reducing the impact of rotation on the recognition, e.g. [40, 41].

The following utilizes the rotation compensation of the USIT in its baseline configuration, because the goal is to look at the capabilities of the segmentation tools to perform their task on rotation-prone / off-angle images.

The first step we conduct is to ascertain the degree of rotation that is actually present in the *protMI* database, to this end we utilize the segmentation ground truth with `manuseg` and *lg*

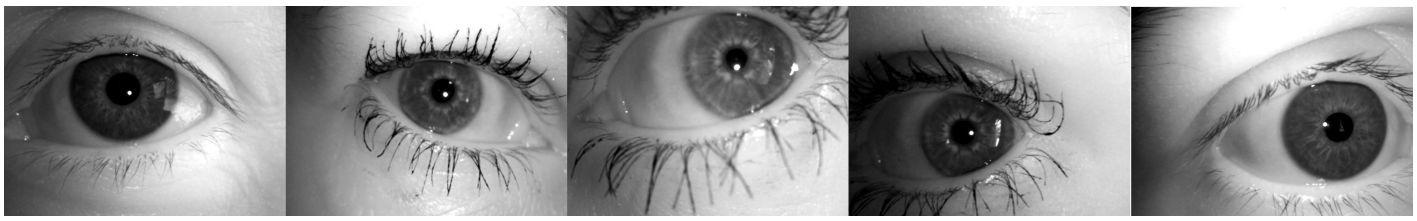


Figure 3: Iris samples from the *PROTECT Multimodal DB* database.

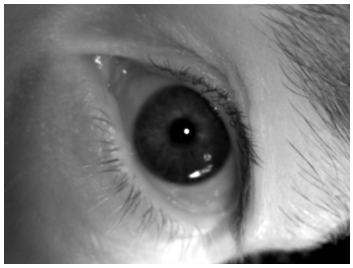


Figure 4: An example of a removed image since the polynomials used for eyelid delimitation can not deal with the strong rotation.

Table 4: Experiment to find the rotation in the *protMI* Database, based on ground truth, `manuseg` and *lg*.

rotation [±bit]	EER [%]	$OP_{0.01}$[%]
16	8.74	50.48
24	6.49	49.56
32	5.12	49.07
40	4.53	48.74
48	4.10	48.43
56	3.96	48.26
64	3.96	48.26
72	3.96	48.26

as feature. The biometric recognition rate (as obtained when comparing all samples excluding symmetric comparisons), or rather the equal error rate (EER) and the operating point of $FNMR_{@FMR=0.01\%}$(as $OP_{0.01}$in the table), will serve as the error function which is to be minimized. An exact pinpointing of the rotation is not necessary since in a practical application rotation will differ from this set. Consequently, we will start at the suggested minimum ±16bit, and increase the rotation in roughly 5° steps, or ±8bit, until EER and $OP_{0.01}$are minimized.

The results of this experiment are given in Table 4. The error is minimized somewhere between ±48 and ±56 bits, so we will stick with ±56 bit for further experiments. The equivalent of ±56bit are ±39.375°.

Since the experiment was performed with the ground truth, the EER and $OP_{0.01}$ also serve as a baseline for the further automated / algorithmic segmentations.

To get a better view of the actual rotations used during rotation compensation we used the ability of the `hd` comparison tool of the USIT to log the exact bitshift used for the final match. We only used genuine comparisons since imposter to genuine comparisons would result in a random shift. To see the difference to a traditional system, that is using stationary acquisition, we plot the frequency for the *casia4i* database as well. Figure 5 gives the result, the y-axis contains the frequency, that is the percentage of all rotations, per bit shift. The total overlap between the frequencies is 73.6%, or put differently, 26.4% of all comparisons are shifted from the center to the extremes. Meaning roughly a quarter of the genuine comparisons would result in a worse comparison score, and likely false rejection, when using traditional rotation compensation as suggested by USIT default settings.

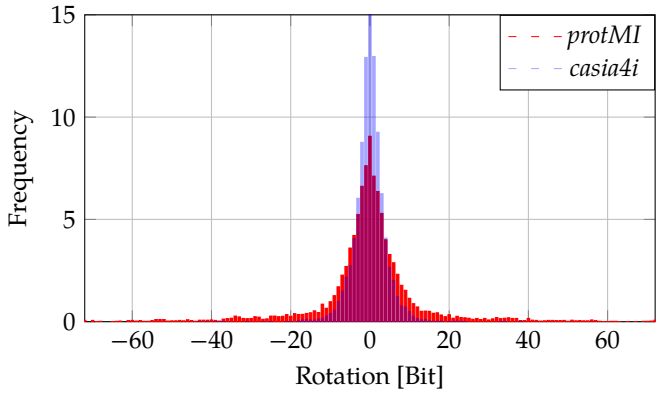


Figure 5: Frequency of rotation on the *protMI* and *casia4i* databases.

3.3 Recognition Performance Evaluation

Many of the traditional segmentation methods, e.g. *CAHT*, assume a cooperative user, a high quality stationary recording device and consequently a high quality frontal image with relatively little obfuscation, i.e. a wide open eye. In actual practice this might not be the case and traditional methods often have problems when it comes to less constraint iris databases, e.g. iris acquisition on the move or using mobile sensors. The utilization of CNNs in the biometric toolchain could solve these problems. Note that this can only compensate segmentation issues not issues arising from the non-linear transformation of the iris texture, see Karakaya [42].

To evaluate the performance of CNNs, and their use in the biometric system (after generating normalised texture as discussed), we will utilize and compare to different segmentation techniques, i.e. *CAHT*, *WAHET* and the manual ground truth, as well as features extraction methods, i.e. *lg* and *qsw*. The rotation compensation will be set to ±56bit as per prior findings. As a representative for CNNs we use the iFCEDN [23]

Table 5: Comparison of different feature extraction and segmentation methods under a rotation of ±56bit. The baseline resulting from ground truth segmentation is also given as *Groundtruth*.

feat.	seg.	EER [%]	$OP_{0.01}$[%]	ME
lg	*Groundtruth*	3.96	48.26	2015
lg	*CNNHT*	7.41	50.02	0
lg	*WAHET*	11.33	62.08	0
lg	*CAHT*	22.64	63.44	441500
qsw	*Groundtruth*	3.87	49.60	2015
qsw	*CNNHT*	7.36	52.03	0
qsw	*WAHET*	11.39	58.84	0
qsw	*CAHT*	23.15	49.43	441500
— OSIRIS —		15.05	99.74	0

(*CNNHT*), due to it's performance on the *protMI* database as per the findings in Section 3.2.1.

Further, to compare with a well known method from literature we included the OSIRIS [43] biometric toolchain. OSIRIS has a relatively good performance but can not be set to a higher rotation compensation.

The results achieved on the *protMI* dataset are given in Table 5. In addition to the EER and $OP_{0.01}$, we also give the number of masking errors (ME). Masking errors happen when the comparison of two iris textures have no unmasked information in common. This happens when masked out bits, eyelids, hair or other obfuscations, take up the majority of the image. In this case no comparison score can be calculated.

What can be seen is that traditional segmentation methods like *CAHT*, which usually outperforms *WAHET* on data acquired in constraint conditions, has obviously problems with rotation. This leads to a high number of masking errors as well as a high EER and $OP_{0.01}$. Segmentation methods which are build to handle more complicated matters are likely to perform better. This is reflected by our choice of *WAHET*, which uses elliptical boundaries to handle off-angle recordings better. A stark decrease in masking errors when compared to *CAHT* is seen and likewise the EER is decreased. The CNN (*CNNHT*) is even more adaptable than *WAHET* and, even though the parameterization also only uses circular boundary fittings, outperforms *WAHET* by quite a margin.

The same basic behaviour can be seen if the *qsw* is used instead of the *lg* feature extraction method, hinting that this performance is actually hinged on the segmentation and not on the choice of feature extraction.

It should also be noted that even the *CNNHT* nearly doubles the EER reached by the ground truth, so while the *CNNHT* clearly outperforms both *CAHT* and *WAHET*, there is ample room for future improvement.

OSIRIS recognition results turn out not to be competitive, but surprisingly still better (at least in terms of EER) as the *CAHT* based results with the extended rotation compensation.

4 Conclusion

Iris data captured by mobile sensors reflect the effect of a more unconstrained acquisition by non-habituated users. Factors such as the higher degree of rotation and off-angle imagery pose problems to classical (segmentation) algorithms used in the iris recognition toolchain. This work has shown that CNN-based semantic segmentation together with a higher degree of explicit rotation compensation during matching significantly improves iris recognition performance for such datasets.

Future work will address the high cost of explicit rotation compensation (by applying e.g. [40, 41] to such data) and we will seek to improve the CNN-based segmentation results by applying elliptical curve models in the iris boundary parametrisation process applied to the binary segmentation masks.

Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 700259.

References

[1] D. Kunda and M. Chishimba, "A survey of android mobile phone authentication schemes," *Mobile Networks and Applications*, pp. 1–9, 2018 (cit. on p. 2).

[2] T. Van Nguyen, N. Sae-Bae, and N. Memon, "Draw-a-pin: Authentication using finger-drawn pin on touch devices," *Computers & security*, vol. 66, pp. 115–128, 2017 (cit. on p. 2).

[3] N. Malkin, M. Harbach, A. De Luca, and S. Egelman, "The anatomy of smartphone unlocking: Why and how android users around the world lock their phones," *GetMobile: Mobile Computing and Communications*, vol. 20, no. 3, pp. 42–46, 2017 (cit. on p. 2).

[4] K. Bowyer and M. Burge, Eds., *Handbook of Iris Recognition*. Springer-Verlag, 2016. DOI: https://doi.org/10.1007/978-1-4471-6784-6 (cit. on p. 2).

[5] C. Rathgeb, A. Uhl, and P. Wild, *Iris Recognition: From Segmentation to Template Security*, ser. Advances in Information Security. Springer Verlag, 2013, vol. 59 (cit. on p. 2).

[6] D. Kim, Y. Jung, K.-A. Toh, B. Son, and J. Kim, "An empirical study on iris recognition in a mobile phone," *Expert Systems with Applications*, vol. 54, pp. 328–339, 2016 (cit. on p. 2).

[7] D. H. Cho, K. R. Park, and D. W. Rhee, "Real-time iris localization for iris recognition in cellular phone," in *6th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing and First ACIS International Workshop on Self-Assembling Wireless Networks*, IEEE, 2005, pp. 254–259 (cit. on p. 2).

[8] D. S. Jeong, H.-A. Park, K. R. Park, and J. Kim, "Iris recognition in mobile phone based on adaptive gabor filter," in *International Conference on Biometrics*, Springer, 2006, pp. 457–463 (cit. on p. 2).

[9] K. R. Park, H. Park, B. J. Kang, E. C. Lee, and D. S. Jeong, "A study on iris localization and recognition on mobile phones," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, p. 20, 2008 (cit. on p. 2).

[10] S. Kurkovsky, T. Carpenter, and C. MacDonald, "Experiments with simple iris recognition for mobile phones," in *Seventh International Conference on Information Technology: New Generations (ITNG'10)*, IEEE, 2010, pp. 1293–1294 (cit. on p. 2).

[11] J.-S. Kang, "Mobile iris recognition systems: An emerging biometric technology," *Procedia Computer Science*, vol. 1, no. 1, pp. 475–484, 2010 (cit. on p. 2).

[12] H. Lu, C. R. Chatwin, and R. C. Young, "Iris recognition on low computational power mobile devices," in *Biometrics-Unique and Diverse Applications in Nature, Science, and Technology*, InTech, 2011 (cit. on p. 2).

[13] M. De Marsico, M. Nappi, D. Riccio, and H. Wechsler, "Mobile iris challenge evaluation (MICHE)-I, biometric iris dataset and protocols," *Pattern Recognition Letters*, vol. 57, no. C, pp. 17–23, May 2015, ISSN: 0167-8655. DOI: 10.1016/j.patrec.2015.02.009. [Online]. Available: http://dx.doi.org/10.1016/j.patrec.2015.02.009 (cit. on p. 2).

[14] M. D. Marsico, M. Nappi, F. Narducci, and H. Proença, "Insights into the results of MICHE I - Mobile Iris CHallenge Evaluation," *Pattern Recognition*, vol. 74, pp. 286–304, 2018, ISSN: 0031-3203. DOI: https://doi.org/10.1016/j.patcog.2017.08.028. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320317303412 (cit. on p. 2).

[15] M. D. Marsico, M. Nappi, and H. Proença, "Results from MICHE II – Mobile Iris CHallenge Evaluation II," *Pattern Recognition Letters*, vol. 91, pp. 3–10, 2017, ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2016.12.013. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865516303683 (cit. on p. 2).

[16] V. Haaf, M. Neukamp, J. Hämmerle-Uhl, and A. Uhl, "Real-world non-nir illumination and wavelength-specific acquisition variants in iris recognition (best paper award)," in *Proceedings of the International Conference on Vision, Image and Signal Processing (ICVISP 2017)*, Osaka, Japan, 2017 (cit. on p. 2).

[17] P. Wild, P. Radu, and J. Ferryman, "On fusion for multispectral iris recognition," in *Proceedings of the IAPR International Conference on Biometrics (ICB'15)*, May 2015, pp. 1–6 (cit. on p. 2).

[18] C. Boyce, A. Ross, M. Monaco, L. Hornak, and X. Li, "Multispectral iris analysis: A preliminary study," in *Proceedings of Computer Vision and Pattern Recognition Workshop on Biometrics (CVPRW'06)*, IEEE, 2006, pp. 1–9 (cit. on p. 2).

[19] H. Li, Q. Zhang, and Z. Sun, "Iris recognition on mobile devices using near-infrared images," *Human Recognition in Unconstrained Environments: Using Computer Vision, Pattern Recognition and Machine Learning Methods for Biometrics*, p. 103, 2017 (cit. on p. 2).

[20] A. V. G. Niamh M. Fitzgerald Christopher Dainty, *Two iris imaging over an extended depth of field with a mobile phone camera*, 2018. DOI: 10.1117/12.2312147. [Online]. Available: https://doi.org/10.1117/12.2312147 (cit. on p. 2).

[21] Q. Wang, X. Su, Z. Cai, and X. Zhang, "Mobile iris recognition via fusing different kinds of features," in *Chinese Conference on Biometric Recognition*, Springer, 2017, pp. 401–410 (cit. on p. 2).

[22] Q. Zhang, H. Li, Z. Sun, and T. Tan, "Deep feature fusion for iris and periocular biometrics on mobile devices," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2897–2912, 2018 (cit. on p. 2).

[23] E. Jalilian and A. Uhl, "Iris segmentation using fully convolutional encoder–decoder networks," in *Deep Learning for Biometrics*, B. Bhanu and A. Kumar, Eds., (ZG) Switzerland: Springer, 2017, ch. 6, pp. 133–155, ISBN: 978-3-319-61656-8. DOI: https://link.springer.com/chapter/10.1007/978-3-319-61657-5_6 (cit. on pp. 2, 3, 5, 6).

[24] 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 *2016 International Conference on Biometrics (ICB)*, 2016, p. 8 (cit. on p. 2).

[25] E. Jalilian, A. Uhl, and R. Kwitt, "Domain adaptation for cnn based iris segmentation," in *Proceedings of the 16th International Conference of the Biometrics Special Interest Group (BIOSIG'17)*, Darmstadt, Germany, 2017, pp. 51–60 (cit. on p. 3).

[26] M. Arsalan, H. G. 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*, vol. 9, no. 11, 2017. DOI: 10.3390/sym9110263. [Online]. Available: http://www.mdpi.com/2073-8994/9/11/263 (cit. on p. 3).

[27] G. Lin, M. Anton, S. Chunhua, and I. Reid, "Refinenet: Multi-path refinement networks for high-resolution semantic segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017 (cit. on p. 3).

[28] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2016 (cit. on p. 3).

[29] H. K. Yuen, J. Princen, J. Illingworth, and J. Kittler, "A comparative study of hough transform methods for circle finding," *Image Vision Computing*, vol. 8, no. 1, pp. 71–77, 1990 (cit. on p. 3).

[30] C. Rathgeb, A. Uhl, P. Wild, and H. Hofbauer, "Design decisions for an iris recognition sdk," in *Handbook of Iris Recognition*, ser. Advances in Computer Vision and Pattern Recognition, K. Bowyer and M. J. Burge, Eds., second edition, Springer, 2016. DOI: 10.1007/978-1-4471-6784-6 (cit. on p. 4).

[31] University of Salzburg, *USIT – University of Salzburg iris toolkit*, http://www.wavelab.at/sources/USIT, version 2.x, University of Salzburg, 2017 (cit. on p. 4).

[32] L. Ma, T. Tan, Y. Wang, and D. Zhang, "Efficient iris recognition by characterizing key local variations," *IEEE Transactions on Image Processing*, vol. 13, no. 6, pp. 739–750, Jun. 2004, ISSN: 1057-7149. DOI: 10.1109/TIP.2004.827237 (cit. on p. 4).

[33] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, University of Western Australia, 2003 (cit. on p. 4).

[34] H. Hofbauer, F. Alonso-Fernandez, P. Wild, J. Bigun, and A. Uhl, "A ground truth for iris segmentation," in *Proceedings of the 22th International Conference on Pattern Recognition (ICPR'14)*, Stockholm, Sweden, 2014, 6pp. (Cit. on pp. 4, 5).

[35] H. Hofbauer, I. Tomeo-Reyes, and A. Uhl, "Isolating iris template ageing in a semi-controlled environment," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'16)*, Darmstadt, Germany, 2016, p. 8 (cit. on p. 4).

[36] U. of Reading, *PROTECT Multimodal DB Dataset*, Available by request at projectprotect.eu/dataset., Jun. 2017 (cit. on p. 4).

[37] H. Proenca and L. A. Alexandre, "The nice.i: noisy iris challenge evaluation - part i," in *2007 First IEEE International Conference on Biometrics: Theory, Applications, and Systems (BTAS'07)*, Sep. 2007, pp. 1–4. DOI: 10.1109/BTAS.2007.4401910 (cit. on p. 5).

[38] D. L. Sparks, "The brainstem control of saccadic eye movements," *Nature Reviews Neuroscience*, vol. 3, no. 12, pp. 952–964, 2002. DOI: 10.1038/nrn986 (cit. on p. 5).

[39] L. R. Young and D. Sheena, "Survey of eye movement recording methods," *Behavior Research Methods & Instrumentation*, vol. 7, no. 5, pp. 397–429, 1975, ISSN: 1554-3528. DOI: 10.3758/BF03201553. [Online]. Available: http://dx.doi.org/10.3758/BF03201553 (cit. on p. 5).

[40] P. Drozdowski, C. Rathgeb, H. Hofbauer, J. Wagner, A. Uhl, and C. Busch, "Towards pre-alignment of near-infrared iris images," in *Proceedings of the IAPR/IEEE International Joint Conference on Biometrics (IJCB'17)*, 2017, p. 8 (cit. on pp. 5, 7).

[41] C. Rathgeb, H. Hofbauer, A. Uhl, and C. Busch, "Triplea: Accelerated accuracy-preserving alignment for iris-codes," in *Proceedings of the 9th IAPR/IEEE International Conference on Biometrics (ICB'16)*, 2016, pp. 1–8 (cit. on pp. 5, 7).

[42] M. Karakaya, "A study of how gaze angle affects the performance of iris recognition," *Pattern Recognition Letters*, vol. 82, pp. 132–143, 2016, ISSN: 0167-8655. DOI: https://doi.org/10.1016/j.patrec.2015.11.001. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167865515003852 (cit. on p. 6).

[43] D. Petrovska and A. Mayoue, "Description and documentation of the biosecure software library," Project No IST-2002-507634 - BioSecure, Deliverable, 2007 (cit. on p. 7).