Received: 3 December 2020 | Revised: 16 May 2021 | Accepted: 22 June 2021

IET Biometrics

DOI: 10.1049/bme2.12052

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

CNN-based off-angle iris segmentation and recognition

Ehsaneddin Jalilian[1] | **Mahmut Karakaya**[2] | **Andreas Uhl**[1]

[1]Department of Computer Science, University of Salzburg, Salzburg, Austria

[2]Department of Computer Science, Kennesaw State University, Marietta, Georgia, USA

Correspondence

Ehsaneddin Jalilian, Department of Computer Science, University of Salzburg, Jakob-Haringer-Strasse 2, 5020, Salzburg, Austria.
Email: ejalilian@cs.sbg.ac.at

Funding information

Austrian Science Fund for Advanced Methods and Applications for Fingervein Recognition, Grant/Award Number: P 32201-NBL; Österreichische Forschungsförderungsgesellschaft FFG KIRAS project AUTFingerATM, Grant/Award Number: 864785; Secure and Trustworthy Cyberspace Secure and Trustworthy Cyberspace (SaTC) program of NSF, Grant/Award Number: CNS-2100483

Abstract

Accurate segmentation and parameterisation of the iris in eye images still remain a significant challenge for achieving robust iris recognition, especially in off-angle images captured in less constrained environments. While deep learning techniques (i.e. segmentation-based convolutional neural networks (CNNs)) are increasingly being used to address this problem, there is a significant lack of information about the mechanism of the related distortions affecting the performance of these networks and no comprehensive recognition framework is dedicated, in particular, to off-angle iris recognition using such modules. In this work, the general effect of different gaze angles on ocular biometrics is discussed, and the findings are then related to the CNN-based off-angle iris segmentation results and the subsequent recognition performance. An improvement scheme is also introduced to compensate for some segmentation degradations caused by the off-angle distortions, and a new gaze-angle estimation and parameterisation module is further proposed to estimate and re-project (correct) the off-angle iris images back to frontal view. Taking benefit of these, several approaches (pipelines) are formulated to configure an end-to-end framework for the CNN-based off-angle iris segmentation and recognition. Within the framework of these approaches, a series of experiments is carried out to determine whether (i) improving the segmentation outputs and/or correcting the output iris images before or after the segmentation can compensate for some off-angle distortions, (ii) a CNN trained on frontal eye images is capable of detecting and extracting the learnt features on the corrected images, or (iii) the generalisation capability of the network can be improved by training it on iris images of different gaze angles. Finally, the recognition performance of the selected approach is compared against some state-of-the-art off-angle iris recognition algorithms.

1 | INTRODUCTION

Iris recognition is one of the most reliable and accurate techniques in biometrics used for human identification. The iris is the only internal organ in humans visible to the outside world. Therefore, its pattern is well protected, virtually stable and very difficult to be forged. Iris acquisition is performed in a contactless and non-invasive manner and under ideal acquisition conditions. Iris data maintains an accurate recognition with a very low false accept rate [1]. Localisation and segmentation of the iris in eye images is the key initial step in iris recognition, which plays a vital role in the accuracy of subsequent feature extraction and recognition processes. Ever since the first iris recognition system proposed by John Daugman [2], a wide variety of techniques have been proposed to perform segmentation in eye images captured typically in a frontal view, under a controlled or constrained environment. In practice, however, many of the users or operators of these systems are inexperienced and often capture images where the subjects are looking in the wrong direction, due to inadvertent eye movement. Meanwhile, the emerging stand-off iris biometric systems and the recent trend towards 'on-the-move-acquisition' are transforming iris biometric systems from being operated in a well-controlled setup to being smart standoff modalities. The iris images captured under such conditions are more likely to be off-angle images, and additional off-angle-related distortions may be incorporated. Segmentation tasks in such images become quite challenging as the iris boundaries are dilated, of

elliptical shape, or even missing in the extreme off-angle images. Most classical segmentation approaches, which are mainly based on the integrodifferential, circular Hough Transform and edge detection techniques that rely on the visibility of clear iris contours, fail to perform segmentation in such images. Also, most feature comparison algorithms, operating under the assumption that the iris texture lies on a flat frontal plane and possesses a circular geometric property, fail to perform the comparison task properly as well [3]. To address such challenges, off-angle iris recognition has recently become an important research topic within the biometrics community.

With recent advancements in deep learning techniques, some convolutional neural networks (CNN) were proposed for the challenging task of iris segmentation (e.g. [4–6]). While there are already several studies addressing the off-angle distortions in the classical iris segmentation literature, there exists no detailed research that investigates and quantifies the affecting mechanism of the off-angle eye-structure-related distortions on the segmentation capability of the CNNs, and no comprehensive framework is dedicated specifically for off-angle iris recognition. The parameterisation and normalisation tools proposed for the obtained iris segmentations are limited to frontal iris images only [7], and most algorithms that proposed to estimate the iris gaze angles and/or re-project them back to frontal view rely on extra supplementary data, such as head position [8], intrinsic parameters, locations and orientations of cameras, lights and monitors, cornea curvature, angular offset between optical and visual axis etc. [9, 10].

In this work, we extend our previous studies [11, 12] within the frame of the proposed end-to-end recognition pipelines to enable the usage of the CNN-based segmentations for the final task of recognition. In particular, in the 'improved-homogeneous' approach, we estimate the gaze angle of the iris images and classify them into different classes according to this angle, so that we can train and apply a dedicated CNN on (homogeneous in terms of gaze angle) the iris images of each distinct gaze angle. The segmentation outputs are then improved, as will be explained in Section 6. In the 'improved-heterogeneous' approach, we train a network with iris images exhibiting different gaze angles, aiming to improve the generalisation capability of the network in a way that can help obtain (hopefully) better results than what we obtained using the angle-specific configuration, as outlined above (where the training and testing data have identical gaze angles). Also, we utilise our off-angle iris parameterisation and correction module along with a couple of other correction algorithms to geometrically re-project the corresponding off-angle iris images back to frontal view, before unwrapping and normalising the extracted iris features. These techniques are termed 'corrected-homogeneous' and 'corrected-heterogeneous', respectively. Using these approaches, we aim to clarify if correcting the off-angle iris texture can compensate for the degradations imposed by the off-angle distortions. Simultaneously, we investigate the effect of the applied interpolation and the possible imperfections of the correction algorithms on the subsequent recognition performance of these experiments. Furthermore, in the 'corrected-frontal' approach (as proposed newly in this work), we apply to the corrected images a network trained only on the frontal eye images, and likewise, evaluate the subsequent recognition performance using each correction algorithm. Doing so, we opt to determine if a network trained on frontal eye images is able to detect and extract the learnt features on the corrected images, eliminating the need for training with off-angle images.

The main new contributions of the current work compared to previous studies [11, 12] lie in the following:

- Proposing a gaze-angle estimation algorithm (free of need for any auxiliary data or instrument) to enable the following: Gaze-angle estimation in the off-angle images, re-projection of the off-angle images back to the frontal view in a realistic manner, as well as angle-specific training when using the segmentation-based CNNs.
- Including other state-of-the-art gaze-angle estimation and correction algorithms in our experiments, aiming to disentangle and investigate the agonising effect of the interpolation applied during the correction procedure and the possible imperfections of the correction algorithms in the frame of several end-to-end recognition pipelines.
- Introducing a new approach (recognition pipeline) termed 'corrected-frontal', in which a network trained only on frontal eye images is applied to off-angle images, which have been geometrically re-projected back to frontal view. This is done utilising the proposed off-angle iris parameterisation and correction module along with the other improvement algorithms.
- Evaluation of the recognition performance of the best-performing approach (pipeline) against some well-known off-angle iris recognition algorithms, after specifying the proper configuration.
- Extending the experimental validation to cover all subjects in the dataset [13] used (doubling the scope of the experiments).
- Extending the recognition experiments to include a further deep-learning-based segmentation architecture for a selected approach.

2 | RELATED WORK

Existing classical iris segmentation methods can be broadly categorised into three types. The first and most popular type is feature-based methods, which aim to locate the inner and outer boundaries of the iris in the iris image. The Hough transform finds the circularity by edge-map voting within the given range of the radius, which is known as the Wildes approach [14]. Daugman's integrodifferential operator is another scheme that finds the boundaries using an integral derivative [15]. Many advancements have been made in these algorithms ever since the introduction of these two algorithms [16, 17]. The methods of the second type use texture-based discriminating features to differentiate between iris and non-iris pixels [18, 19]. The third type of segmentation methods employs active contour models [20, 21]. There exist many current and ongoing research

studies on iris segmentation using the classical approaches specified. But due to the limitation of space, and in order to keep the focus on off-angle iris segmentation, here we introduced only a selection of techniques in each category. For a general overview, please refer to for example [22, 23].

Several different techniques, which stem from classical methods, have been proposed to address the off-angle iris segmentation and recognition problem. Daugman proposed to detect the inner and outer iris boundaries using an active contour method, based on the discrete Fourier series expansion of the contour data [1]. Shah and Ross combined snake segmentation with geometric active contours [20]. Generally, active contour methods require parameters that are distribution dependent and hard to generalise. Zuo et al. [24] used the intensity, shape, and localisation features from the iris and pupil to automatically segment non-ideal iris images. Their method demonstrated performance improvement on challenging iris images, up to 30°gaze angle. Kennell et al. [25] proposed to segment images containing non-circular irises with morphological operations. Gangwar et al. proposed the IrisSeg algorithm [26], which adopts a coarse-to-fine strategy to localise different boundaries. The pupil is coarsely detected using an iterative search method and by exploiting dynamic thresholding and multiple local cues. The limbic boundary is first approximated in the polar space using adaptive filters and then refined in the Cartesian space. Uhl and Wild proposed the WAHET (Weighted adaptive Hough and ellipsopolar transforms) algorithm [27], which uses an adaptive Hough transform at multiple resolutions to estimate the approximate position of the centre of the iris. A subsequent polar transform detects the first elliptic limbic or pupillary boundary, and an ellipsopolar transform finds the second boundary based on the outcome of the first. Few other works such as [28, 29] focussed on refining the features of the iris after extraction to mitigate the off-angle-related distortions.

Some other works tried to compensate for the off-angle iris distortions. Schuckers et al. [30] used a method in which the angle estimates were searched to pick those that maximise the value of the integrodifferential operator on iris images. The algorithm utilises a brute-force technique to re-project the input image to all possible gaze angles. In [31], the authors utilised the boundary segmentation of the iris and pupil and compared the boundaries with a look-up table generated by using a biologically inspired biometric eye model and finding the closest feature point in the look-up table to estimate the gaze angles. Bolme et al. [32] presented four methods for correcting off-angle iris images so that they appear frontal. They tested their affine transformation method on segmentations which are manually corrected. The displacement and generic algorithms they used require a corresponding frontal image to perform the correction for each off-angle image (which is not realistic). Also, the refractive method they used is based on a synthesised eye model and requires supplementary information, such as the focal length of the camera, distance to the iris etc., to perform the correction on the iris images. Price et al. [33] developed a generalised eye model to correct for perspective and refractive distortion of the iris pattern using ray tracing techniques. They reported a median reduction of Hamming distance for synthetic eyes, with gaze up to 60°. A main drawback in the majority of recent approaches is that they rely on classical boundary detection methods, which operate on the input iris images and generally tend to localise false iris boundaries (specially in off-angle images). Furthermore, in many cases, they used synthesised eye models or manually adjusted data to evaluate their algorithms, which does not reflect real world scenarios.

To address the drawbacks of classical segmentation methods and to reduce the complexity of intensive pre- and post-processing, a fourth category of segmentation methods evolved recently, which is based on data-driven learning methods. Within this category, deep learning techniques and, in particular, convolutional neural networks are the most ideal and popular schemes due to their accuracy and performance. Liu et al. [5] located the iris region in non-cooperative environments using convolutional neural networks. In their study, a hierarchical CNN (HCNNs) and a multi-scale FCN (MFCNs) were used to locate the iris region automatically. Jalilian and Uhl [6] proposed three types of fully convolutional encoder–decoder networks for iris segmentation and evaluated their performance on off-angle iris images available in the UBIRIS.v2 database[1]. Their results showed the superior capability of CNNs to deal with off-angle iris data compared to some classical methods. The method presented by Arsalan et al. [34] roughly estimates the iris region using an edge detection algorithm and then classifies the pixels into two classes (iris and non-iris) by using a CNN. They performed a fine tuning of a VGG CNN, and tested their model on iris images captured in non-ideal environments. Rot et al. [35] presented a deep multi-class eye segmentation model built around a semantic segmentation architecture. They also examined the sensitivity of the network to the change of view for four directions (left, right, up and straight), generally. In order to overcome the requirement of large quantities of labelled data in the approaches mentioned above, Jalilian et al. [36] proposed a domain adaption technique for CNN-based iris segmentation. Bazrafkan et al. [37] introduced a CNN to perform iris segmentation on lower quality iris images (including off-angle images). They further investigated the effect of network tuning on the results of the segmentation. The work presented in [4] proposed a deep network called IrisDenseNet, which is based on VGG-16, to deal with low-quality iris images, such as side views, glasses, off-angle eye images and rotated eyes. Roig et al. [38] proposed to segment the iris region using a multi-class approach, which differentiates additional classes, for example the pupil or sclera, aiming to improve the iris segmentation in non-cooperative environments using a CNN. Nevertheless, none of the above works provided a systematic analysis on the effect of the different gaze angles on ocular biometrics and the resulting iris segmentations and recognition using CNNs. There exist other studies that employed deep learning models for feature extraction or comparison stages of

[1] http://iris.di.ubi.pt/ubiris1.html.

an iris recognition system capable of dealing with off-angle images. Karakaya et al. [39] proposed a deep-learning-based iris recognition model, in which they adopted the AlexNet [40] classifier to simply classify off-angle and frontal iris images belonging to 52 subjects (as available in their database) to 52 corresponding classes.

3 | EYE STRUCTURE-RELATED DISTORTIONS AND THEIR EFFECTS ON THE CNN-BASED IRIS SEGMENTATION

In addition to the known degradation factors affecting the constrained (frontal) iris imaging (e.g. focus, motion blur, specular reflections, and illumination variations), off-angle iris imaging introduces further challenging eye structure-related distortions to the iris images. In the following sections, we investigate the interaction mechanism of light rays within the anterior and posterior eye structure elements, such as the cornea, limbus, sclera, anterior chamber (aqueous humour), iris, and lens (as illustrated in Figures 1 and 2), and analyse the distortions they may introduce to the actual iris image, depending on the image acquisition angle and the way they can affect the segmentation capability of the network.

3.1 | Three-dimensional structure of iris

The structure of the iris consists of several types of dilator muscles, forming a three-dimensional texture on the iris plane. As the gaze angle changes, the 2D image of the captured iris texture changes, amending the key content and features learnt by the CNN networks during the training process (see Figure 3i, which shows the difference between the normalised images of an iris captured frontally (Figure 3g), and from −50° gaze angle (Figure 3h) in red, where constant parts are depicted in several shades of blue). Those pixels located on (the side closer to the camera) the border region of the iris inner boundary may get occluded in steeper view angles (see the light rays unseen (blue) and seen (green) by the camera in different angles in Figure 2). There will be considerable changes in the distribution of iris features when the gaze angle of training and testing iris images differs, and the network may not be able to spot the corresponding features (as learnt in the training session) in the test images, failing to segment the iris region accurately.

3.2 | Limbus occlusion

The limbus is the semitransparent organ that joins the sclera and the cornea texture, where the fully transparent cornea cannot reach to the bottom of the anterior chamber and ends at a higher level than the iris plane. Due to the distance between the ending points of the cornea and the iris plane, the diameter of the cornea-limbus border is slightly smaller than the anterior chamber width. Therefore, the limbus consistently occludes the boundary region of the iris texture (especially) in extreme off-angle view. The extent of occlusion of the iris texture on the side closer to the camera increases as the gaze angle increases (see the red-dotted reflections in Figure 2). Thus, the off-angle images do not exhibit certain outer iris boundary information, as present in the frontal training images. This causes a CNN trained on the frontal images to fail to accurately detect the iris region (especially its outer boundary) in the (off-angle) test images.

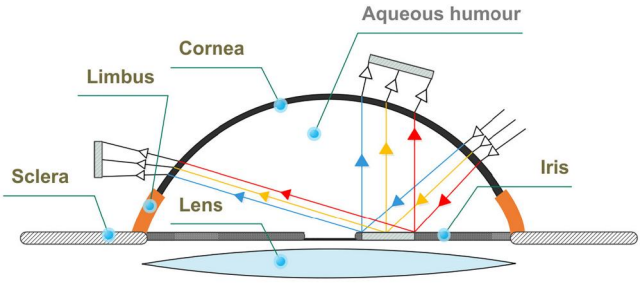

FIGURE 1 Posterior eye structure and the perspective and refraction distortions affecting the iris texture geometry

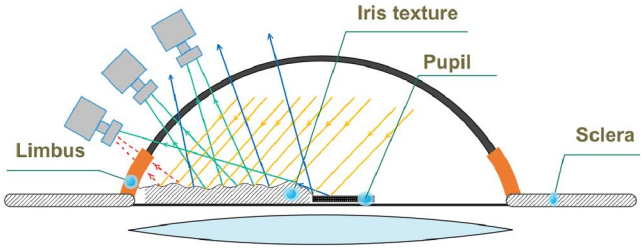

FIGURE 2 Three-dimensional structure of iris and Limbus occlusion distortions

3.3 | Perspective and refraction distortion

The geometric properties of an object's image on the camera sensor change if the coordinates of the camera change with respect to the object. This phenomenon is generally referred to as 'perspective distortion'. The cornea is the transparent structure of the eye located at the outermost layer of the eye. The aqueous humour is the transparent watery fluid that is located between the cornea and the iris and fills the anterior chamber. Therefore, incoming and outgoing light rays are first refracted at the cornea and then refracted at the aqueous humour due to the refraction index differences between air, the cornea, and the aqueous humour. When capturing iris images at steeper angles, light rays refract more at the cornea, causing the geometric property of the reflected iris features to transform (e.g. get scaled, dilated or eroded), as shown in Figure 1. Correspondingly, we can see that the circular shape of the iris image captured in a frontal manner in Figure 3d is transformed to an ellipse in Figure 3e, when captured from a −50° angle, mainly due to perspective distortion. We can also observe the effect of this distortion along with the refraction distortion on

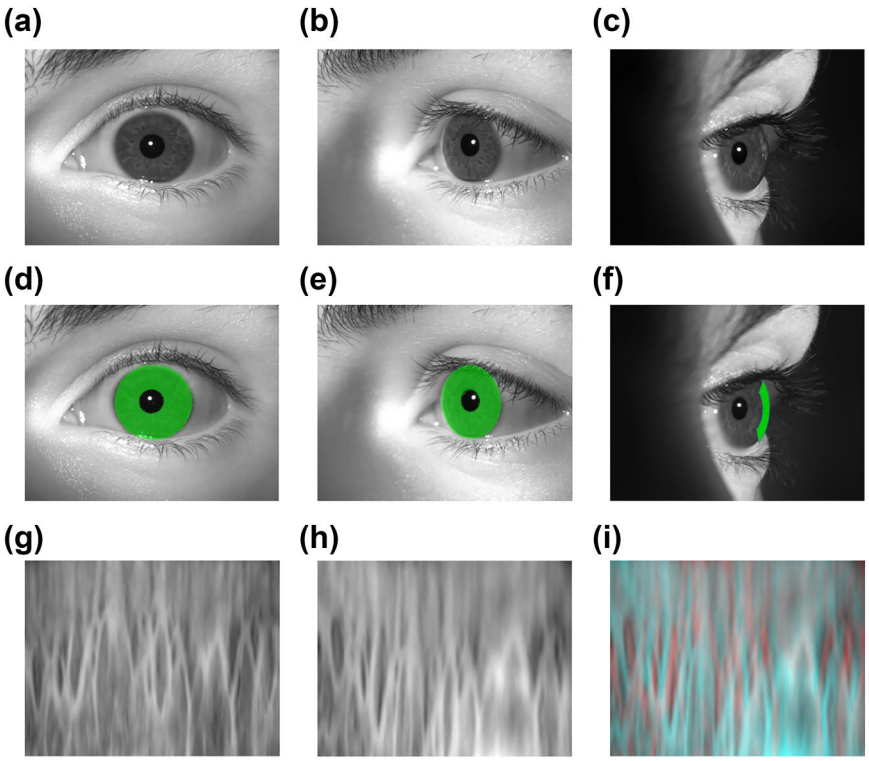

FIGURE 3 Examples of the iris images in the dataset and the corresponding off-angle distortions

the geometric properties of the corresponding normalised iris textures (Figure 3g, and Figure 3h, respectively), as presented in Figure 3i. Basically, CNNs learn scale-dependent patterns in a specific combination of image size and network architecture, and thus, they face serious difficulties in spotting the learnt patterns in the testing data if their geometric properties (such as boundaries and texture information) are changed with respect to the training data.

3.4 | Iris missing boundary in extreme angles

The sclera is the outer layer of the eye with bright white colour which strongly contrasts with the coloured iris' texture, forming a clear boundary between these two tissues. In frontal imaging, this boundary is clearly visible. But as the gaze angle gets steeper (especially towards the right-most gaze angle as we consider the left eye (i.e. +50°)), the boundary erodes and finally disappears (see the green curve, showing the missing iris boundary in Figure 3f). The learning process in CNNs starts with convolving filters that can be thought of as feature identifiers, which convolve over the input, looking first for low-level features, such as edges and boundaries, and then building up to more abstract concepts through further filtering layers. Thus, low-level features, such as edges (boundaries), play a scaffolding role in encoding the feature representations of target regions. Therefore, if these features (learnt during the training process) are not presented in the testing data (which might have a steeper gaze angle than the training data), the network will not be able to retrieve the accurate boundary pixels (which are missing) or may wrongly spot false boundaries (false-positives) in the image.

3.5 | Posterior eye effect

The space behind the iris and in front of the vitreous body is referred to as the posterior eye chamber, which includes the ciliary body muscle and the lens complex. The reflected light rays (as received by the camera during the iris acquisition) do not directly interact with the posterior eye structures. However, during accommodation of the eye, the variation of the lens thickness can change the curvature of the iris surface by pushing it forward and backward. To this extent, the effects of the posterior eye on the iris is indirect, and thus, should be studied within the frame of anterior eye structures' effects, which have already been discussed in the previous subsections.

4 | OFF-ANGLE IRIS GAZE ESTIMATION

Gaze estimation is an important prerequisite for analysing, in case of a need, correct off-angle iris images. There is extensive literature on the video-based gaze-angle estimation and eye tracking applications that is [41–43]. In these applications, various intrusive and non-intrusive techniques and instruments are used, depending upon the required level of accuracy, to estimate the gaze angles of the eyes. Generally, non-intrusive methods are cheaper, more comfortable, more practical and less risky, but less accurate than intrusive methods. These

methods can be roughly classified into four main categories: corneal reflection-based methods [9], mapping functions-based methods [44], model-based methods [31], and appearance-based methods [10]. To achieve accurate gaze-angle estimation with these non-intrusive methods, typically, a set of parameters (e.g. intrinsic parameters, locations and orientations of cameras, lights and monitors, cornea curvature, head position [45], angular offset between optical and visual axis etc.) needs to be set initially by conducting well-controlled calibration and/or training steps. We have already reviewed the algorithms that propose to estimate the gaze angle without the use of such supplementary data in eye images and discussed their main drawbacks in Section 2. Additional works proposed within this category, for example [46], are mainly theoretical solutions tested only on synthesised data, which is generated by (controlled) mathematical setups and generally does not apply to realistic applications.

Addressing this need, we developed a (non-intrusive) method, which requires no additional instrument or supplementary information to estimate the iris gaze angles in the eye images, with very high accuracy. The principle of the algorithm is based on measuring the relative distance of certain off-angle iris and pupil feature points in the segmentation output masks generated by the CNNs. The information is then fed to a fine K-Nearest-Neighbourhood (KNN) classifier to classify the iris images to their corresponding gaze angles. Based on our experiments, this model is able to successfully classify iris images into their corresponding gaze-angle classes with about 97% accuracy. To achieve this, we first decouple the iris and pupil regions in each segmentation mask. To improve the masks and remove false-negative pixels, we compute and superimpose the convex hull of each region, where the iris convex hull boundary information is used also to decouple the iris and pupil region in the masks where there is no clear boundary between these regions (see Figure 4 for an example). Next, we find the right and left-most extreme iris and pupil pixels on the horizontal axis passing through the centre points of the corresponding regions. Using these 5 points we develop10 relative measures, which are expected to be unique for each gaze angle (see Figure 5a showing the corresponding measures).

The rationale behind this is the fact that as the eye gaze angle moves (towards the left or right) on the horizontal axis, the geometric features of the iris (e.g. horizontal diameter, radius etc.) change as well. The proportion of such changes is assumed to be unique for each gaze angle and ought to provide an estimation of the eye gaze angle. Based on the same logic, we support the obtained data by adding further measures calculated using the segmentations' extrema points (8 points each). These measures are, in particular, obtained by subtracting the left side points (Top-left (TL), Left-top (LT), Left-bottom (LB), and Bottom-left (BL)), from the right side points (Top-right (TR), Right-top (RT), Right-bottom (RB), and Bottom-right (BR)), as illustrated in Figures 5b and 5c.

Out of the 42 measures obtained, 11 were found (experimentally) to be unique for each gaze angle. So, we associated each iris image with a 21-elements (10 + 11) vector (as illustrated in Table 1), and trained a fine KNN classifier, which makes finely detailed distinctions between classes with the number of neighbours set to 1, with this information. Figure 6 shows the corresponding algorithm performance on the different classes in the form of a confusion plot. As the results show, the model is able to determine the iris gaze angles with high accuracy (≈97%), with very low prediction deviations (possessing standard deviation of ≈0.48), in a way that, in most cases, the falsely predicted gaze angles are adjacent to the target ones. This, in fact, leads to less deviations when performing angle-specific training as well as the iris correction tasks. We used a 20-fold training scheme for training the KNN, where the training and testing data were fully separated with no overlap. It was proven that the primary 10 measures (by experiment) contribute the most to the algorithm's accuracy, as we were able to obtain about 80% accuracy by including only these measure. Further, by adding the secondary 11 measures we could improve this result up to 97% accuracy.

To get better insight into the performance of the model, we compared the model's performance with another setup applied to the same data. For this purpose, we first implemented the core algorithm proposed by Schuckers et al. in [30] (as a generic angle-correction algorithm) in which the estimates are searched to pick those that maximise the value of the integrodifferential operator on iris images possessing different gaze angles, terming it as the 'Integrodifferential-based algorithm'. We updated the code to enable gaze-angle estimation within the range of the angles available in our dataset (−50° to +50°) and then applied the algorithm to our dataset and compared the results with those obtained using the KNN-based algorithm. Figure 7 shows the results obtained in this experiment in the form of a confusion plot. As can be seen, the performance of the KNN-based algorithm is far better than this algorithm in terms of accuracy. Also, the run time of 5 min per image (using an Intel-i5-6500-3.20 GHz CPU) makes the algorithm computationally very inefficient (compared to the KNN-based algorithm with a run time of less than 1 s) as well. It is also interesting to note that, in the majority of the cases, the algorithm finds the maximum value of the integrodifferential operator in images which encountered no re-projection (the horizontal middle row in the table), regardless of the actual iris gaze angles in the images. Technically, this can reflect the negative effect of interpolation (applied when re-projecting the irises to different gaze angles) on the key iris boundary feature, which causes the integrodifferential operator to fail to find the maximum value on the corrected circular irises.

5 | ELLIPSE-BASED OFF-ANGLE IRIS PARAMETERISATION AND CORRECTION

As already mentioned in Section 1, in this work, we propose an off-angle parameterisation method to outline the elliptic boundary of the iris in the segmentation outputs of off-angle images and use this information to re-project the segmentation masks along with their corresponding off-angle iris images back to frontal view. Currently, there is a lack of suitable algorithms, which enable a true parameterisation of the off-angle iris region

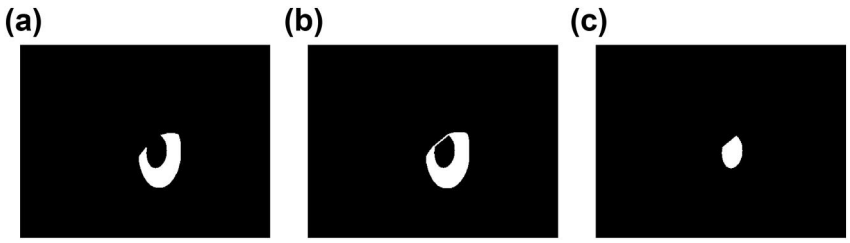

FIGURE 4 Example of a segmentation output mask (4a), where the iris convex hull boundary is superimposed (4b) to decouple the pupil region (4c)

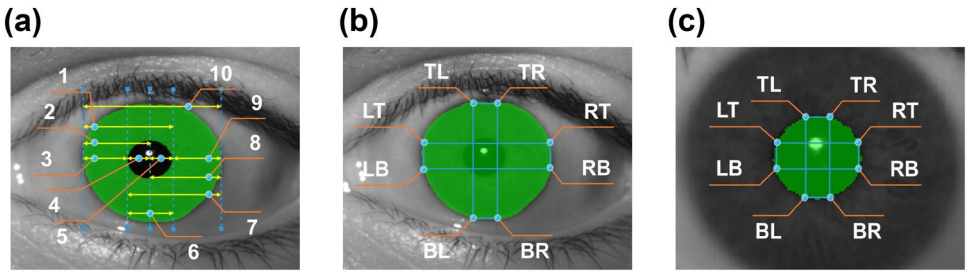

FIGURE 5 Examples of the (zoomed) iris (5a, 5b) and pupil (5c) segmentation outputs (green areas) and the corresponding measures considered

TABLE 1 21-elements vector associated with each image

pTL-pBL	pTL-pBR	pTR-pBL	iTR-iBL	1	2	3
4	5	6	7	8	9	10
iLB-iRB	iLT-iRT	iTL-iRT	iTR-iLT	iRT-iBL	iTL-pBL	iLT-pRT

in CNN-based segmentation outputs. The available parameterisation algorithms are limited to the frontal segmentation outputs, where the circular Hough transform is used to parameterise the iris region. In fact, off-angle iris images normally possess an elliptic shape and, to this extent, circular-based algorithms fail to perform well on such images. The main obstacle to applying an elliptic parameterisation to the actual segmentation outputs is the tendency of such models to overly oblong or obround, due to occlusion of the iris by the eyelid or eyelashes. To resolve this issue, we propose to search only for the vertical edges in the segmentation outputs. It is ensured that the edges detected here correspond to the actual iris boundaries, as the iris region is already specified by the CNN. The information on the horizontal edge is further used to filter out the horizontal noise. Thus, the resulting edge points secure the proper fitting of an ellipse to the actual iris region. A least-squares criterion is then used for estimation of the best ellipse fit to the given set of points (see Figure 8 for an example). In the next step, we extract the horizontal and vertical axes information of the ellipse. This information is used for re-projecting (correcting) the segmentation outputs and their corresponding off-angle iris images back to frontal view as follows: We assume that our ellipse is in the following parametric form:

$$x = x_0 + Q \times \begin{bmatrix} a \times cos(\theta) \\ b \times sin(\theta) \end{bmatrix}, \tag{1}$$

where x and x_0 are 2-dimensional vectors and $a > b > 0$ corresponds to the horizontal and vertical axes of the ellipse,

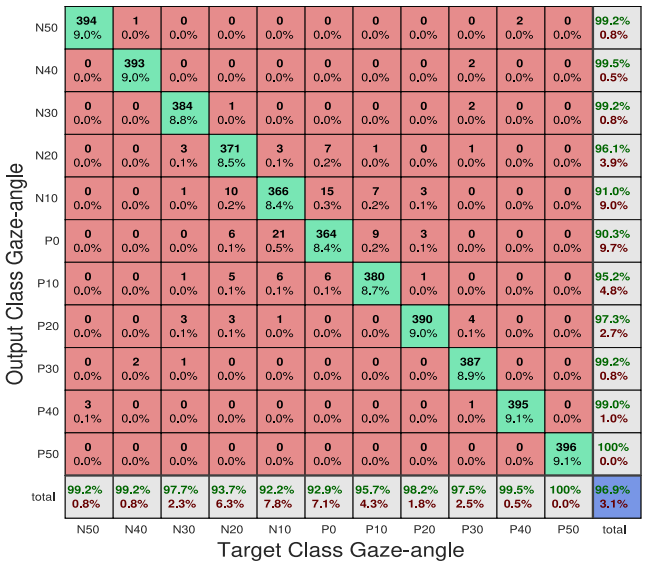

FIGURE 6 Gaze-angle estimation results on the different gaze-angle groups using the K-nearest-neighbourhood-based algorithm

respectively. Q is the rotation matrix and θ represents the rotation angle. We assume a vertical ellipse, as our rotation angles are to the left and right only. Thus,

$$Q = \begin{bmatrix} cos(90) & -sin(90) \\ sin(90) & cos(90) \end{bmatrix}. \tag{2}$$

We want our transformation to produce y in the shifted, rotated coordinates as follows:

$$y = \begin{bmatrix} 1 & 0 \\ 0 & a/b \end{bmatrix} \cdot \begin{bmatrix} a \times cos(\theta) \\ b \times sin(\theta) \end{bmatrix}, \tag{3}$$

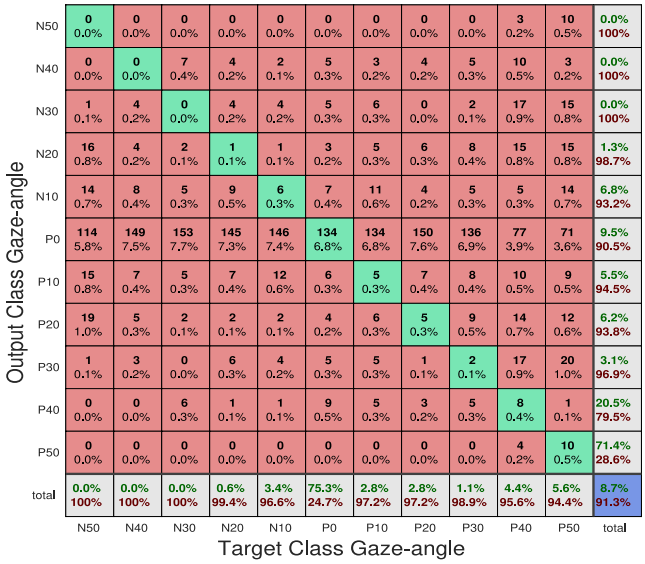

FIGURE 7 Gaze-angle estimation results on the different gaze-angle groups using the integrodifferential-based algorithm

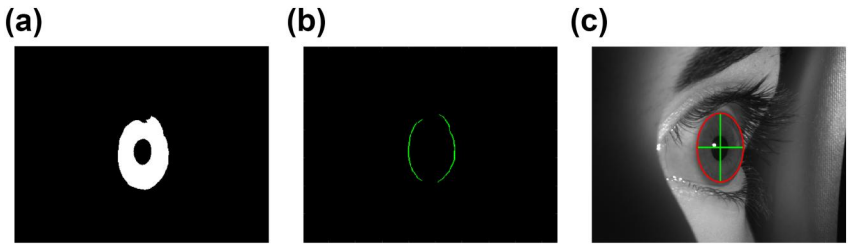

FIGURE 8 Sample iris segmentation output (8a), the extracted vertical boundaries (8b), and the corresponding fitted ellipse (8c)

where x represents the original coordinates. By submitting to equation (1), we can infer the affine transformation matrix we need to re-project the parameterised ellipse back to frontal view, so that it possess a circular shape.

$$x = \left[Q\begin{bmatrix}1 & 0\\ 0 & a/b\end{bmatrix}Q'\right]x + \left[\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - Q\begin{bmatrix}1 & 0\\ 0 & a/b\end{bmatrix}Q'\right]x_0 \tag{4}$$

6 | SEGMENTATION IMPROVEMENT

From the findings in our previous work [11], we already know that the network tends to produce some false-positive detection, in particular, along the segmentation output masks' borders. Therefore, we improved the segmentation outputs by applying some specific morphological operation. So, we first defined a marginal area (A) along each border of the segmentation output masks (with a width (in pixel) equal to 1/5 of the length of the same border) and then performed an opening operation with a big (disk-shape) structuring element (B) as follows:

$$A \circ B = (A \ominus B) \oplus B, \tag{5}$$

where $\ominus$ and $\oplus$ denote erosion and dilation, respectively. We further performed another opening operation on the overall segmentation output masks using a small (disk-shape) structuring element to remove small false-positive detections outside the iris region. Figure 9 shows a sample segmentation output mask and its corresponding improved version.

7 | EXPERIMENTAL FRAMEWORK

Dataset: For our experiments, we used a subset (containing 4400 left eye images captured from 40 subjects) of an off-angle iris database [13]. The iris images in this dataset are captured by two near-infrared sensitive IDS-UI-3240ML-NIR cameras. Images at the 0°gaze angle were captured by a frontal fixed camera, and off-angle images were captured by a frontal moving camera rotating horizontally from −50°(N50) to +50° (P50) at an angle with a 10°step-size. Each camera captured 10 iris images per stop, giving 10 frontal and 100 off-angle iris images captured from each subject, to comprise about 400 images per angle (examples of images in the dataset are already presented in Figure 3). The database is accessible on request (from the authors), and further details about it can be found in [13]. We developed the ground-truth labels (required for training the network) for all the images available in the dataset using the iris, pupil, upper and lower eyelid parameters specified manually (the data will be made available upon acceptance of the manuscript at http://wavelab.at). We used a

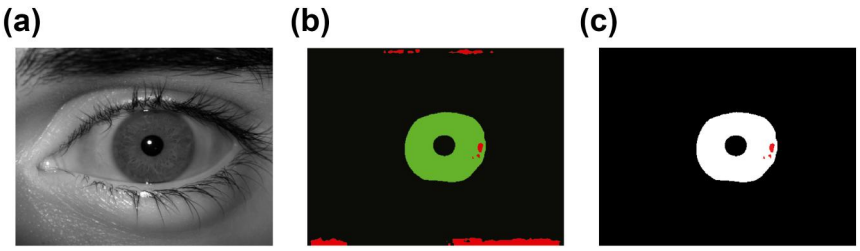

FIGURE 9 Sample iris image with P0 gaze angle (9a) and its corresponding segmentation (green area) and error mask (red area) before (9b), and after correction (9c), using the network trained on P0 images

2-fold training scheme to carry out experiments on segmentation. For this, we divided the whole dataset into two equal parts (with no overlap according to the subjects). In the first fold of the training, we used one part as our testing data and the other one as our training data, and in the second fold of the training, the roles of the datasets where switched.

Fully convolutional neural network (FCN): We selected the RefineNet [47] to perform the iris segmentations in our experiments. The network is already proven to enable high-resolution prediction, and at the same time, preserve the boundary information (which is crucial to our parameterisation mechanism). It is a multi-resolution refinement network, which employs a 4-cascaded architecture with 4 refining units, each of which directly connects to the output of one residual net [48] block as well as to the preceding refining block in the cascade. Each refining unit consists of two residual convolution units (RCU), which include two alternative ReLU and 3 × 3 convolutional layers. The output of the RCU units are processed by 3 × 3 convolution and up-sampling layers incorporated in multi-resolution fusion blocks. A chain of multiple pooling blocks, each consisting a 5 × 5 max-pooling layer and a 3 × 3 convolution layer, then operates on the feature maps, such that one pooling block takes the output of the previous pooling block as its input. Therefore, the current pooling block is able to re-use the result from the previous pooling operation and thus access the features from a large region without using a large pooling window. Finally, the outputs of all the pooling blocks are fused together with the input feature maps through summation of the residual connections. We used an Adam optimiser with a learning rate of 0.0001, executing 40,000 iterations to train the network.

Recognition pipeline: The segmentation outputs (after applying correction or improvement), are parameterised using the technique introduced in [7]. The extracted iris patterns are normalised by unwrapping the circular region into a rectangular block of constant dimensions. Each isolated iris pattern is then demodulated to extract its phase information (feature) using quadrature 1-D Gabor wavelets. To compare the unique extracted features with one another, the Hamming distance, with rotation correction, was calculated in the comparison phase. We used the University of Salzburg implementation of these algorithms, as provided in the Iris Toolkit (USIT)[2]. Figure 10 illustrates the overall recognition pipeline, along with the proposed parameterisation and correction module.

Segmentation evaluation and measures: In order to facilitate proper quantification of the accuracy of the segmentations in each experiment, we considered the *nice*1 iris segmentation error rate, which is based on the NICE1 protocol[3] and used in several iris segmentation challenges. Accordingly, the segmentation error rate (*nice*1) for each segmentation output mask I_i is given by the proportion of corresponding disagreeing pixels (through the logical exclusive or operator) with the ground-truth mask, over all the output mask as follows:

$$nice1 = \frac{1}{c \times r} \sum_{c'} \sum_{r'} O(c', r') \otimes C(c', r'), \qquad (6)$$

where c and r are the dimensions of the segmentation, and $O(c', r')$ and $C(c', r')$ are the pixels of the segmentation and the ground-truth mask, respectively. The value of (*nice*1) is in the [0, 1] interval, where 1 and 0 are the worst and the best scores, respectively.

8 | EXPERIMENTS AND ANALYSIS

In the first step, we investigated the effect of different gaze angles on the CNN-based off-angle iris segmentations, aiming to address the primary research question: If a gaze-angle specific training is required for high segmentation accuracy. We extended the experiments and analysis to the subsequent recognition as well. So, initially, we used the KNN-based module to estimate the gaze angles of the iris images in our dataset. Then, we trained the network following the improved-homogeneous approach, that is the training data consists of iris images with identical gaze angles (200 images per gaze angle for each fold, as in our training data). For each available gaze angle, a dedicated network was trained, and then we conducted segmentation on all the testing data, differentiating and grouping results into the different gaze angles available (starting from −50°(N50) to +50°(P50)). An analysis of the network segmentation outputs (examples displayed in Figure 11) shows that the extensive use of the pooling filters and residual information, as utilised in the RefineNet architecture, makes the network vulnerable to texture-related distortions, such as perspective, refraction and also 3-D iris structure distortions, leading to many false-positive detections

[2] http://www.wavelab.at/sources/USIT.

[3] http://nice1.di.ubi.pt/.

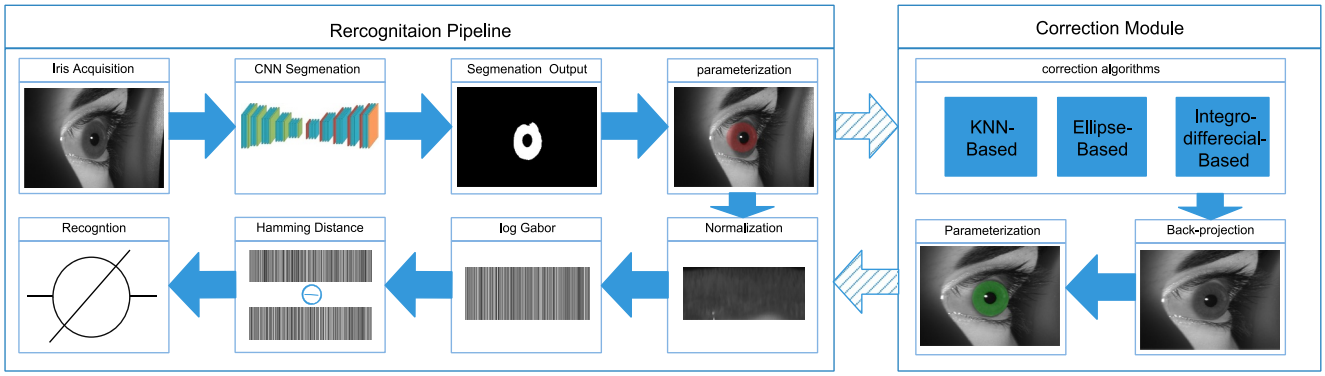

FIGURE 10 Recognition pipeline and the correction module

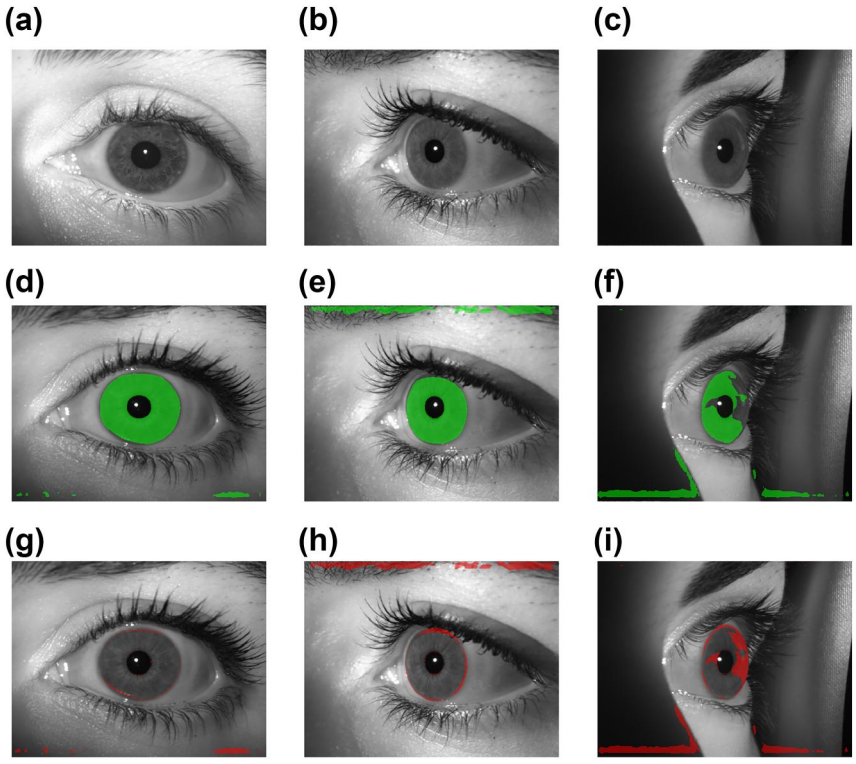

FIGURE 11 Sample iris images of P0 (11a), N50 (11b), and P50 (11c) gaze angles in the dataset, and their corresponding segmentation masks (11days, 11e, 11f) and error masks (11g, 11h, 11i), using the network trained on the frontal iris data

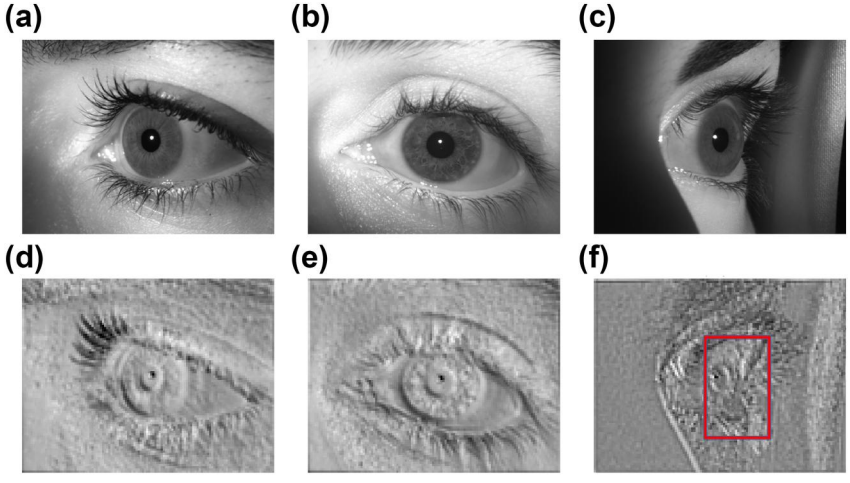

FIGURE 12 Sample iris images with N50 (12a), P0 (12b), and P50 (12c) gaze angles, and their corresponding outputs (12d, 12e, and 12f) extracted from the intermediate convolutional neural network layers

(specially along the outputs' borders) as well as some undetected iris pixels (false-negatives), especially on the main texture of the iris, as reflected in the segmentation outputs and their corresponding error masks in Figure. 11g, 11h, and 11i, respectively.

The effect of the missing iris boundary is also visible in Figure 11i (as missed iris boundary pixels), while the effect of limbus occlusion seems to be negligible. Figure 12 visualises the outputs of some intermediate CNN layers for the configurations where the network is trained on the frontal eye images and tested on the N50, P0, and P50 data. In particular, we may notice the network failure to retrieve the accurate iris boundary pixels (inside the red marked area) in the case of P50 iris images due to the missing boundary features, as demonstrated in Figure 12f.

The rule of background variation on the distribution of false-positive detections for different gaze angles (as in Figure 11h, where the eyebrow region is falsely detected as the iris, or in Figure 11f, where the periocular region is falsely detected) is also worth mentioning. While an analysis of such a variation can certainly contribute to improving the network performance in these cases, we believe that it is beyond the scope of this study and needs to be addressed in a separate research study.

To address these issues, we opt to improve the segmentation outputs using the improvement scheme already explained in Section 6. Figure 13 shows the corresponding segmentation results, as the average *nice*1 error, after improving the segmentation outputs. As the results show, the improvement resulted in considerable enhancements in almost all the segmentation results (especially for the right off-angle (P) images) compared to the segmentation results obtained in [11], as the average segmentation error decreased (about 2 times) from 0.030 to 0.013. Also, in accordance with what we found using the identical training scheme (Homogeneous) and the network (RefineNet) already in [11], we can see the direct relation of the network's performance to the similarity of the gaze angles of the training and testing images, hereafter the morphological improvement as well. Yet, the key new finding is that the performance gradually improves as the gaze angles of the training and testing data converge in terms of the angle but may also diverge in terms of the direction. The corresponding gradual improvement of the results, as they get distanced from the middle vertical column of the table towards the sides (except for the extreme right gaze angles, which are most affected by the missing boundary distortion), clearly reflects this fact. Technically, this shows the dominance of the gross content features (i.e. iris boundaries) learnt by the network over the fine eye image context features. To be more precise, the network is able to detect the elliptical features of the symmetric iris in the images captured from the same angle (with respect to frontal view) but in opposite direction. The applied improvement, which, in fact, compensated for some false-positive detections (caused by the off-angle distortions), allowed us to figure out this capability of the network.

In the next step, we fed the improved segmentations along with their corresponding images to the recognition pipeline to

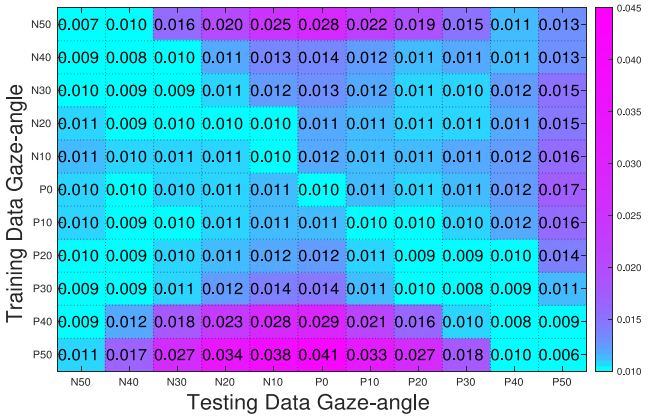

FIGURE 13 Segmentation performance on the different gaze-angle groups under the improved-homogeneous approach as average *nice*1 error

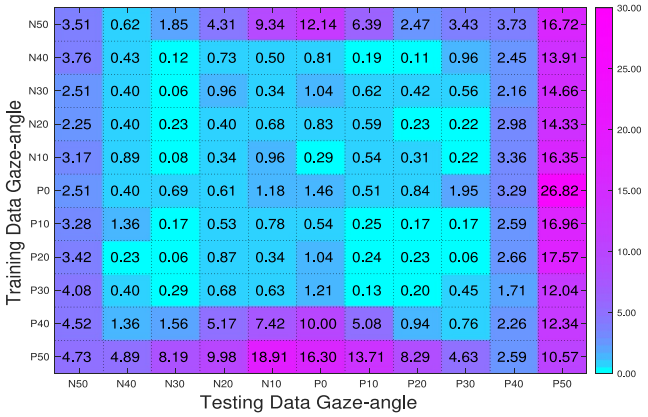

FIGURE 14 Recognition performance on the different gaze-angle groups under the improved-homogeneous approach in terms of equal error rate error

investigate the recognition performance in terms of equal error rate under this approach. Figure 14 shows the results of this experiment. As expected, we observe that the segmentation results are more or less translated into the recognition scores, following the same trends discussed in the segmentation experiments. The only visible difference here is the lower recognition performance of the extreme gaze-angle images (i.e. N50 and specially P50), which, as already mentioned, is mainly due to the extreme 3-D and perspective distortions on the extracted iris' textures.

In the improved-heterogeneous approach, we considered investigating if the generalisability of the network can be improved by switching to a heterogeneous training setting, where we include iris images with different gaze angles in the training data. For this purpose, first we trained the network with all the iris images that have different gaze angles in our training data and then tested it on all the iris images in our testing data. Likewise, the segmentation outputs were improved by morphological opening and the performance was evaluated, differentiating and grouping the results into the different gaze angles available. While the heterogeneous configuration was expected to deliver good results (compared to the angle-specific

training configuration), based on the findings in [11], here we (i) evaluated the extent to which the improvement applied can enhance the segmentation performance and (ii) verified if the improved segmentations can eventually improve the recognition performance beyond the improved angle-specific training configuration.

Figure 15 demonstrates the segmentation results for this experiment in the form of a boxplot for each gaze-angle group (after the improvement). As the results show, by applying the improvement, we obtained considerable enhancement in almost all segmentation results (especially for the right off-angle (P) images) that we had already obtained in the identical heterogeneous configuration without improvement in [11], as the average segmentation error decreased from 0.023 to 0.008. Also, as the results show, we obtained almost the same segmentation performance as obtained using the improved-homogeneous approach for the angle-specific configurations.

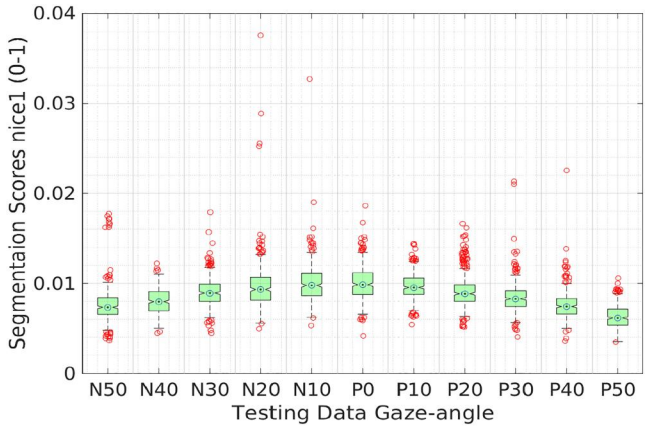

FIGURE 15 Segmentation performance on the different gaze-angle groups under the improved-heterogeneous approach as average *nice*1 error

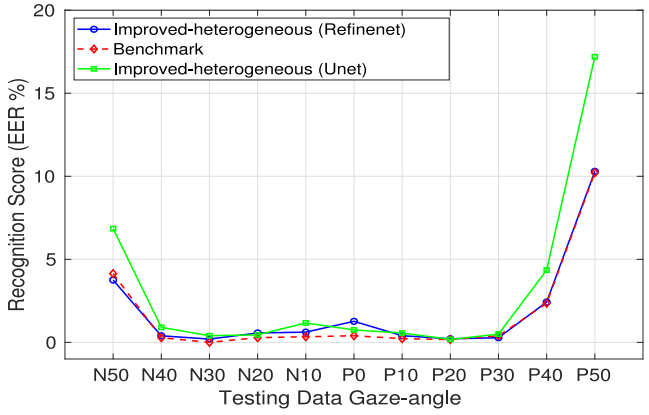

FIGURE 16 Recognition performance benchmark and the corresponding experimental results on the different gaze-angle groups under the improved-heterogeneous approach using different convolutional neural networks, in terms of *EER* error. EER, equal error rate

Figure 16 shows the subsequent recognition results obtained using the corresponding images. As can be seen in the figure, in spite of minor degradation in some results (i.e. N50, P40 and P50 gaze-angle groups' results), we could maintain, on average, the same recognition performance (≈1.85) as that obtained under the improved-homogeneous approach for the angle-specific configurations (≈1.87). Of course, this is a very positive result, as it enables us to refrain from the angle-specific training strategy, and even better, there is no need to determine the iris images gaze angles or carry out the correction procedure before deploying the recognition system.

To provide a proper evaluation of the actual performance of the model, we developed a performance benchmark as well. For this, we segmented the iris region on the eye images using the manual ground truths (resembling the perfect condition) and then fed the segmented iris regions into the recognition pipeline. We also considered another well-known CNN ('Unet' [49]) to be trained (using the manual ground truths) and ran on the same data. As can be seen in Figure 17, the segmentation outputs generated by Unet also contain corresponding errors (false-positive and false-negative pixels), especially in the boundary areas, to a certain extent. So, we applied the same improvements (as proposed in Section 6) on the segmentation outputs and proceeded to the recognition experiment. As the corresponding results in Figure 16 show, the RefineNet segmentation outputs maintained better recognition performance (which is close to the benchmark results in the majority of cases) than those of the Unet.

In the corrected-homogeneous approach, we aim to address if re-projecting the off-angle iris images back to frontal view and correcting the off-angle iris texture can compensate for the degradations imposed by the off-angle distortions, and we eventually improve the system recognition performance. To address this, we re-projected the improved segmentation outputs along with their corresponding iris images back to frontal view using our KNN-based algorithm. To disentangle and investigate the agonising effect of the interpolation applied during the correction procedure, and the possible imperfections of the correction algorithm, we further included the integrodifferential-based and the ellipse-based algorithms in our experiment.

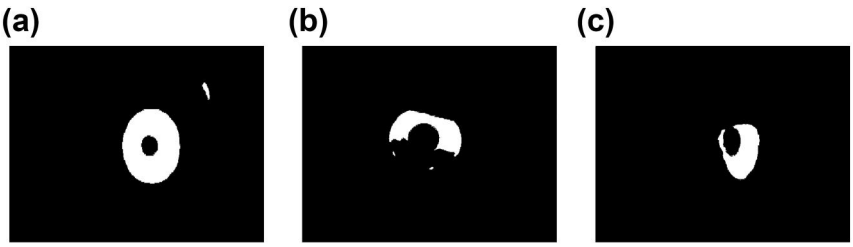

FIGURE 17 Sample segmentation outputs for N50 (17a), P0 (17b), and P50 (17c), using Unet network

Re-projecting the off-angle images back to frontal view (correcting them) concerns with rotating the images back, along the rotation direction. The off-angle iris images in our database are captured by rotating the camera around the vertical axis by $\theta°$. Thus, in order to correct the (off-angle) images, theoretically, all we need to do is to develop the corresponding rotation matrix, substitute the rotation value (e.g. as obtained using our gaze-angle estimation anti-logarithms), and apply it to the (off-angle) images as follows:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & cos(\theta) & -sin(\theta) \\ 0 & sin(\theta) & cos(\theta) \end{bmatrix} . \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (7)$$

Applying such a transformation to the images practically results in certain changes in the image dimensions, depending on the value of the rotation applied (see Figure 18 for an example). To compensate for this, we first calculated a ratio (R) for the corresponding change as follows:

$$R = \frac{newheight \times previouswidth}{previousheight} \quad (8)$$

A new image width was then calculated using this ratio for each image, and the images were cropped, centring the middle of the new calculated widths on the images' centre pixels. The resulting images then got rescaled to their original sizes, as demonstrated in Figure 18.

After correcting the iris images in our testing data using each algorithm, we fed them into the recognition pipeline to evaluate the recognition performance. Figure 19 shows the recognition results using the KNN-based algorithm. When comparing the results to those obtained using the improved-homogeneous approach (see Figure 13), a notable improvement in the majority of results belonging to right gaze-angle configurations is observable. The improvement gradually tends to increase as the testing data gaze angles diverge more from the frontal view (i.e. P20, P30 etc.), so that we can observe the maximum improvement in the results of the extreme right gaze-angle groups (i.e. P40 and P50). However, this is not true in the case of the left gaze-angle images, and the results seem to be coherent with those obtained under the improved-homogeneous approach, with a slight degradation of all the results (see Figure 14 for a comparison). This seems to be as expected, as the gaze-angle prediction results obtained by this algorithm for the right gaze-angle groups were better than those of the left-angle groups (see Figure 6).

Figure 20 shows the corresponding results using the integrodifferential-based algorithm. As expected, we can observe an overall degradation in the results (specifically in those of the frontal view images) compared to the results obtained using the improved-homogeneous approach and the KNN-based approach, which is clearly due to the poor per-

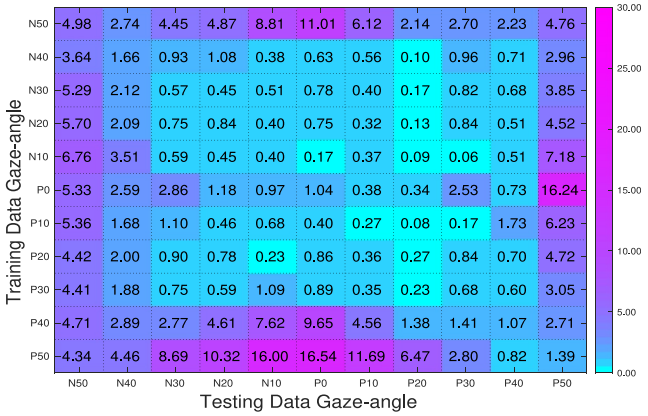

FIGURE 19 Recognition performance on the different gaze-angle groups under the corrected-homogeneous approach, using the K-nearest-neighbourhood-based algorithm, in terms of equal error rate error

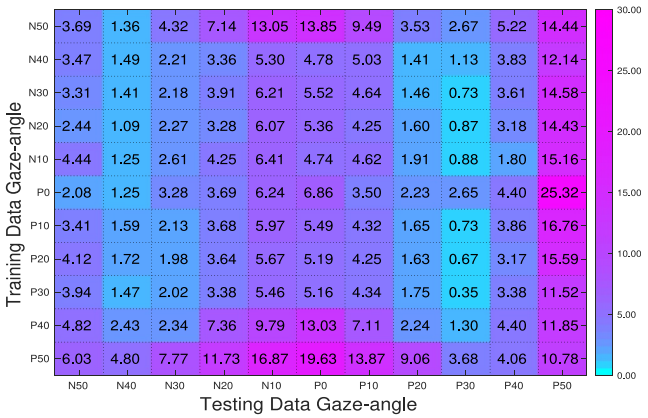

FIGURE 20 Recognition performance on the different gaze-angle groups under the corrected-homogeneous approach, using the Integrodifferential-based algorithm, in terms of equal error rate error

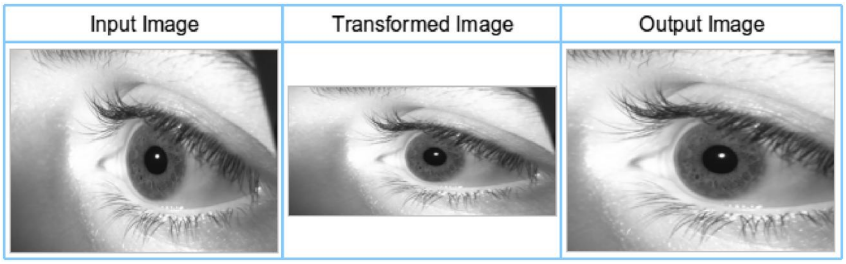

FIGURE 18 Sample off-angle iris image (left) and its corresponding transformed image (middle) and the resulting output after cropping and rescaling

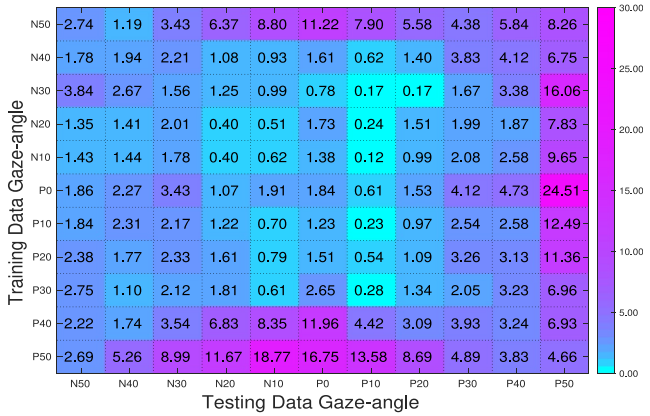

FIGURE 21 Recognition performance on the different gaze-angle groups under the corrected-homogeneous approach, using the Ellipse-based algorithm, in terms of equal error rate error

formance of the algorithm when detecting the correct iris gaze angles. The results obtained using the ellipse-based algorithm are partially different (see Figure 21), and except for a slight improvement in those of the extreme right and left gaze-angles groups (i.e. N50 and P50), all other results encountered a medium degradation compared to those obtained using the improved-homogeneous approach. Nonetheless, these results are generally better than those obtained using the integrodifferential-based algorithm, and except for some improvements in the results of the extreme left gaze-angles images (i.e. N40 and N50), do not show a notable improvement compared to those obtained using the KNN-based algorithm.

Considering the fact that the KNN-based algorithm was very highly accurate in predicting the iris gaze angles, we inspected the corresponding outputs of the N40 and N50 image groups (as demonstrated in Figure 22c) to get better understanding of the slight improvement obtained on these gaze-angles groups, using the ellipse-based algorithm. The inspections revealed that the ellipse-based algorithm tends to underestimate the iris gaze angles (and thus, applies less interpolation) for the extreme right off-angle images (i.e. N40 and N50), as the value of interpolation in this algorithm is estimated based on the actual circularity of the iris shape in the image rather than the capture angle (as in the case of the KNN-base algorithm) recorded. All in all, regardless of the type of correction algorithm applied, the overall degradation of the results, as the gaze angles of the test data diverge more from the frontal view (moving from the central horizontal column in the corresponding tables to the sides), is more than the corresponding degradations observable in the segmentation results obtained under the improved-homogeneous approach. So, the experiments here decoupled and clearly distinguished the negative effect of the interpolation (in distorting the key iris features) applied during the correction process from the possible defects of the correction algorithms.

We further considered the corrected-heterogeneous approach in which we investigated if correcting the off-angle iris texture can compensate for the degradations imposed by the off-angle distortions within a heterogeneous training configuration. So here, after training the network on iris images with different gaze angles and testing it on the images of each gaze angle separately, the segmentation outputs were morphologically improved, parameterised and re-projected back to frontal view, using all the three correction algorithms already used in the previous experiment. Subsequently, the recognition performance was evaluated. Figure 23 demonstrates the results of this experiment for each gaze angle and algorithm, along with the results we obtained using the improved-heterogeneous approach. As can be seen, the results obtained using the KNN-based and ellipse-based algorithms show slight degradation for the left gaze-angle images (N50, N40,\enleadertwodots), while showing considerable improvement for the P50 gaze-angle images compared to the results obtained under the improved-heterogeneous approach. The results for the rest of the gaze-angle groups are more or less the same. The integrodifferential-based algorithm generally does not deliver any promising results under this approach either. Thus, we can see that depending on the correction algorithm used, the correction setup can only improve the performance on some of the most extreme gaze-angle images (i.e. N50, P50), while it generally does not improve the results in other cases. Of course, the type and the scale of the changes in the performance in each case are subject to the influence of the two factors already explained.

The corrected-frontal approach experiments are meant to address if we can enable segmentation in the iris images of different gaze angles using a network trained on the frontal eye images, by correcting their gaze angles and bringing them back to frontal view. For this reason, we first corrected all off-angle iris images available in our testing data, using their corresponding ground-truth masks, with the help of the correction algorithms. Then, we trained the network only with the iris images belonging to frontal view in our training data and performed segmentation in the corrected images in our testing data (400 samples per gaze angle). The segmentation outputs were then morphologically improved, and the recognition performance was evaluated subsequently.

Figure 24 shows the results obtained in this experiment along with the corresponding results obtained under the improved-heterogeneous approach (as the selected approach), the results obtained by applying the same network to the corrected off-angle images as obtained by the KNN-based algorithm (sixth row of the table in Figure 19), and also those of the angle-specific training obtained under the improved-heterogeneous approach. As the results show, almost in all the cases, either the results do not show any specific improvement or are degraded to a certain extent. While the KNN-based algorithm delivered the best correction results, the corresponding segmentation results for the extreme left gaze-angle images (i.e. N50, N40) obtained by this algorithm degraded very intensely compared to those of the other correction algorithms used. Figure 25 demonstrates the sample segmentation outputs obtained using each correction algorithm along with their corresponding error masks in these experiments. As can be observed, there are many false negative detections (red pixels in Figure 25g) in

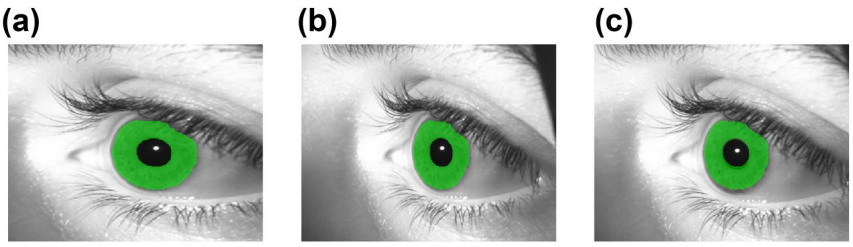

FIGURE 22 Sample corrected iris images of N50 corrected by KNN-based (22a), integrodifferential-based (22b), and ellipse-based (22c) algorithms

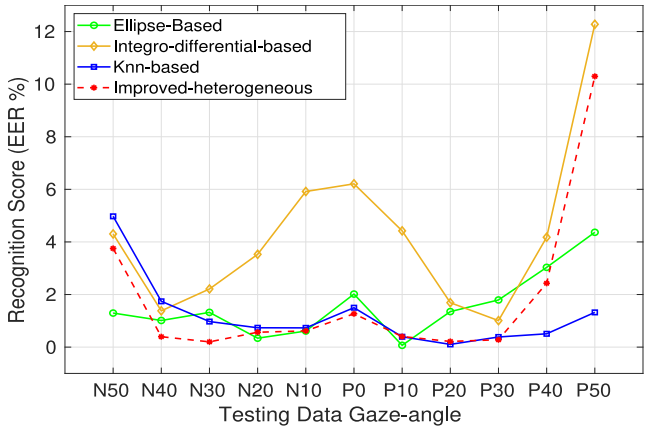

FIGURE 23 Recognition performance on the different gaze-angle groups under the corrected-heterogeneous approach, using the different correction algorithms, in terms of *EER* error. EER, equal error rate

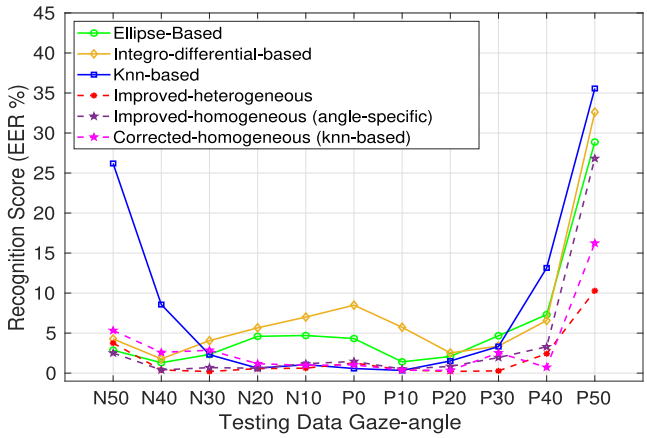

FIGURE 24 Recognition performance on the different gaze-angle groups under the corrected-frontal approach, using the different correction algorithms, in terms of *EER* error. EER, equal error rate

the corresponding segmentation outputs of this algorithm. These again show the negative role of the interpolation applied during the correction process, and the fact that, as it increases, it can alter the unique texture patterns of the iris learnt by the network and cause the network to fail to detect the true iris pixels in the test images.

After investigating the different approaches to figure out a solid model that can best deal with the iris images captured from different angles using the CNNs, we further proceeded with evaluating the model's performance within the frame of other off-angle iris recognition algorithms. For this purpose, we implemented some well-known off-angle iris recognition algorithms, which are based on deep learning, as well as classical techniques. In particular, we considered the IrisSeg [26] algorithm and the WAHET [27] algorithm as the classical approaches and the algorithm proposed in [39] as our deep-learning-based setup. The technical details of these algorithms are already explained in Section 2. We applied each algorithm to our testing data and prepared the results for each gaze-angle group separately, as presented in Figure 26. As it can be seen in the plot, the selected approach (improved-heterogeneous) is superior to all other classical algorithms. Yet, the deep learning algorithm has the optimal performance. Of course, the main drawback of deep-learning-based classifiers of this type is their non-scalability. To be more precise, these networks have to be trained each and every time a new subject is introduced to the system. Furthermore, the performance of such classifiers gradually declines as the number of classes increases [50]. Thus, they may show good performance on a dataset containing a limited number of subjects, but their performance may deteriorate as the size of the dataset (the number of classes) increases.

9 | CONCLUSION

We investigated the effect of different gaze angles on ocular biometrics, and analysed their distorting mechanism on off-angle iris segmentation using CNNs, as well as their subsequent recognition performance. The effect of perspective, refraction, and 3-D iris structure distortions on the network mainly appeared as missing (undetected) iris texture in the network segmentation outputs. These distortions also caused considerable changes in the unique iris features, especially in the extreme (right) gaze-angle images, undermining the corresponding recognition results for these gaze-angle groups as well. The missing and dilated iris boundary and the limbus distortions caused the network to fail to accurately extract the iris boundary pixels (false-negatives) in the corresponding segmented areas of the images with steeper gaze angles. In this case, the recognition results also degraded due to the missing boundary information, which is required for seamless recognition. In fact, in all the cases, the effect of the distortions was more severe on the images captured from the right angles than those captured from the left angle, due to the more severe off-angle distortions they include when captured from the right angles, as we consider left eyes only.

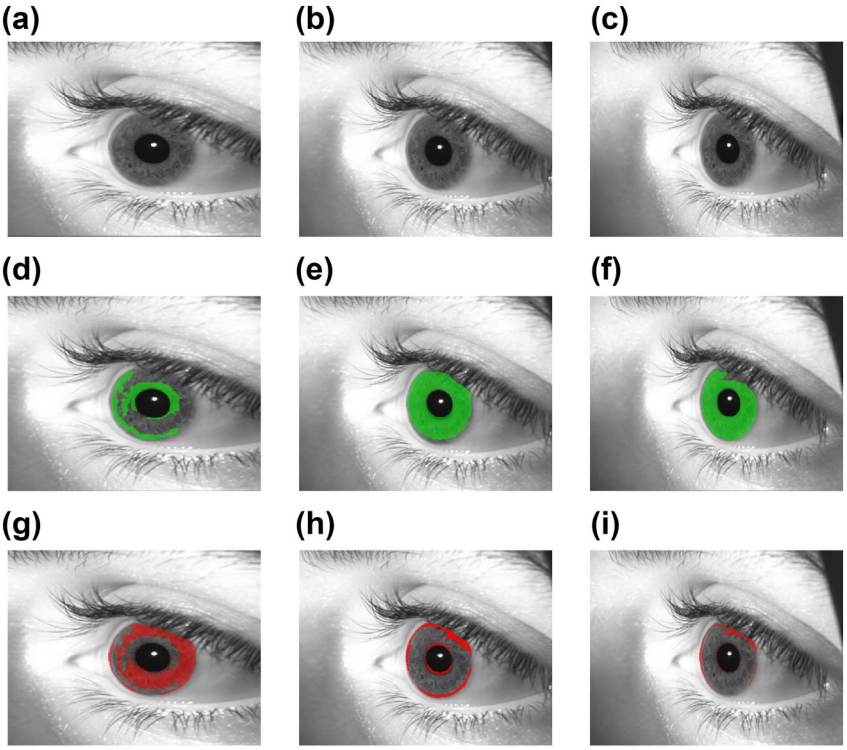

FIGURE 25 Sample N50 iris images, corrected using the knn-based (25a), the ellipse-based (25b), and the integrodifferential-based (25c) algorithms, and their corresponding segmentations outputs (second row) and the error masks (third row)

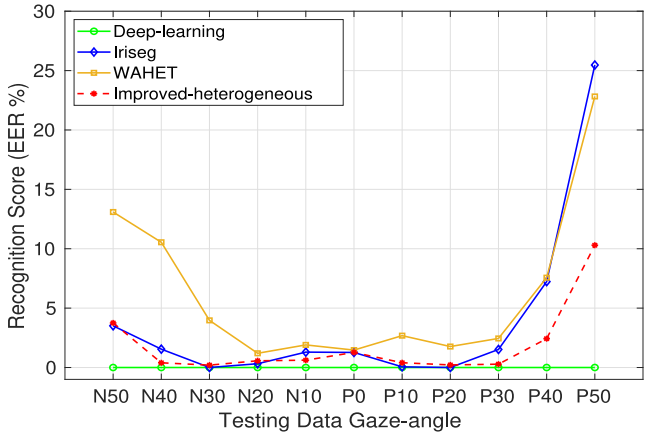

FIGURE 26 Recognition performance on the different gaze-angle groups using the different recognition algorithms, in terms of *EER* error. EER, equal error rate

The results obtained under the improved-homogeneous approach showed that the performance of the network has a direct relation to the correspondence of the gaze angles of the training and testing images, and it declines as the gaze angles diverge. We further found that the network performance gradually improves as the gaze angle of the training and testing data converges in terms of the angle but diverges in terms of direction. This showed the capability of the network to detect the symmetric iris contents in the images captured from the same angle, but in the opposite direction, which was figured out as the result of the segmentation improvement. Also, the morphological improvement technique proved to compensate for some false-positive segmentation errors and enhanced the segmentation results beyond the results obtained in [11], in identical configurations. Furthermore, the proposed KNN-based algorithm for estimating the iris gaze angles in the eye images successfully classified the iris images into their corresponding gaze-angle classes with about 97% accuracy, without the need for any supplementary information (i.e. head position, camera parameters etc.).

The experiments carried out to investigate the effect of correcting the off-angle iris images on the recognition performance showed that the interpolation applied during the correction procedure and the imperfections of the correction algorithms can diversely influence the unique iris features in the eye images. In some cases, the recognition performance is improved, while in other settings it is reduced. While the better correction performance of the KNN-based algorithm (compared to the other two algorithms) resulted in the general improvement of the recognition results, the lower extent of interpolation applied by the ellipse-based algorithm, which tended to underestimate the angle in the right gaze-angle images, improved the recognition performance beyond the values obtained by the KNN-based algorithm on these gaze-angle groups. The experiments carried out under the corrected-heterogeneous approach also showed similar outcomes. And except in some extreme gaze angles (e.g. P40, P50) or when certain correction algorithms (the KNN-based algorithm) were used, the correction process did not improve the recognition performance considerably. So, based on these results we may conclude the following: (i) provided that a suitable correction algorithm is applied, the correction setups are recommended only for the extreme off-angle images (i.e. gaze angles greater than 30°) and (ii) the type and scale of the improvement

obtained in this case is defined as a function of the influence of the value of interpolation applied during the correction procedure and the possible imperfections of the correction algorithm.

The experiments carried out under the corrected-frontal approach, in which we applied the network trained only on iris images belonging to frontal view to the corrected images from different gaze angles, showed that the network is not able to detect and extract the learnt features on the corrected images in a way that can help us obtain better results than those obtained when applying the same network to the 'uncorrected' off-angle images, using the same configuration.

Also, our experiments actually showed that we can maintain almost the same segmentation and recognition performance, as that obtained in the angle-specific configurations, using the improved-heterogeneous approach. Further experiments showed the superiority of this approach (improved-heterogeneous) over all other classical off-angle recognition algorithms. These are, in fact, very promising results, as the selected approach enabled us to refrain from (i) the angle-specific training strategy, which requires determination of the iris images gaze angles in advance, and (ii) even more important, carrying out the correction procedure (which proved to be complicated, based on our findings) before deploying the recognition system. Also, in spite of the fact that the deep-learning-based recognition algorithms performed better (i.e. [39])- than the classical algorithms, the improved-heterogeneous model was preferred, as it addresses some key drawbacks of these algorithms (e.g. poor scalability).

ORCID

Ehsaneddin Jalilian https://orcid.org/0000-0001-5925-5735

REFERENCES

1. Daugman, J.: Probing the uniqueness and randomness of iriscodes: results from 200 billion iris pair comparisons. Proc. IEEE. 94(11), 1927–1935 (2006). https://doi.org/10.1109/jproc.2006.884092
2. Daugman, J.: How iris recognition works. In: The Essential Guide to Image Processing, pp. 715–739. Elsevier (2009). https://doi.org/10.1016/b978-0-12-374457-9.00025-1
3. Jinyu Zuo, J., Schmid, N.A.: On a methodology for robust segmentation of nonideal iris images. IEEE Trans. Syst., Man, Cybern. B. 40(3), 703–718 (2010). https://doi.org/10.1109/tsmcb.2009.2015426
4. Arsalan, M., et al.: IrisDenseNet: robust Iris segmentation using densely connected fully convolutional networks in the images by visible light and near-infrared light camera sensors. Sensors. 18(5), 1501 (2018). https://doi.org/10.3390/s18051501
5. Liu, N., et al.: Accurate iris segmentation in non-cooperative environments using fully convolutional networks. In: 2016 International Conference on Biometrics (ICB), pp. 1–8. IEEE (2016). https://doi.org/10.1109/icb.2016.7550055
6. Jalilian, E., Uhl, A.: Iris segmentation using fully convolutional encoder-decoder networks. In: Bir Bhanu, A.K. (ed.) Deep Learning for Biometrics. (ZG) SwitzerlandSpringer, 6. pp. 133–155. (2017). https://doi.org/10.1007/978-3-319-61657-5_6
7. Hofbauer, H., Jalilian, E., Uhl, A.: Exploiting superior cnn-based iris segmentation for better recognition accuracy. Pattern. Recogn. Lett. 120, 17–23 (2019). https://doi.org/10.1016/j.patrec.2018.12.021
8. En-Wong, T., et al.: Gaze estimation using residual neural network. In: 2019 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), pp. 411–414. IEEE (2019). https://doi.org/10.1109/percomw.2019.8730846
9. Guestrin, E.D., Eizenman, M.: General theory of remote gaze estimation using the pupil centre and corneal reflections. IEEE Trans. Biomed. Eng. 53(6), 1124–1133 (2006). https://doi.org/10.1109/tbme.2005.863952
10. Xu, L., Machin, D., Sheppard, P.: A novel approach to real-time non-intrusive gaze finding. In: Carter, J.N., Nixon, M.S. (eds.) Proceedings of the British Machine Vision Conference, pp. 1–10. British Machine Vision Association (1998). https://doi.org/10.5244/c.12.43
11. Jalilian, E., Uhl, A., Karakaya, M.: Gaze-angle impact on iris segmentation using cnns. In: Proceedings of the IEEE 10th International Conference on Biometrics: Theory, Applications and Systems, pp. 1–8. (2019). https://doi.org/10.1109/btas46853.2019.9185970
12. Jalilian, E., Karakaya, M., Uhl, A.: End-to-end off-angle iris recognition using cnn based iris segmentation. In: Proceedings of the IEEE 19th International Conference of the Biometrics Special Interest Group, pp. 1–12. (2020)
13. Karakaya, M., et al.: Limbus impact on off-angle iris degradation. In: International Conference on Biometrics (ICB), pp. 1–6. (2013). https://doi.org/10.1109/icb.2013.6612971
14. Wildes, R.P.: Iris recognition: an emerging biometric technology. Proc. IEEE. 85(9), 1348–1363 (1997). https://doi.org/10.1109/5.628669
15. Ann-Roy, D., S-Soni, U.: Iris segmentation using daughman's Daughman's method. In: International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 2668–2676. (2016)
16. Jeong, D.S., et al.: A new iris segmentation method for non-ideal iris images. Image Vis. Comput. 28(2), 254–260 (2010). https://doi.org/10.1016/j.imavis.2009.04.001
17. Tan, T., He, Z., Sun, Z.: Efficient and robust segmentation of noisy iris images for non-cooperative iris recognition. Image Vis. Comput. 28(2), 223–230 (2010). https://doi.org/10.1016/j.imavis.2009.05.008
18. Khan, T.M., et al.: Automatic localization of pupil using eccentricity and iris using gradient based method. Opt. Laser. Eng. 49(2), 177–187 (2011). https://doi.org/10.1016/j.optlaseng.2010.08.020
19. Parikh, Y., Chaskar, U., Khakole, H.: Effective approach for iris localization in nonideal imaging conditions, pp. 239–246. IEEE Students Technology Symposium (2014)
20. Shah, S., Ross, A.: Iris segmentation using geodesic active contours. IEEE Trans. Inf. Forensics Secur. 4(4), 824–836 (2009). https://doi.org/10.1109/tifs.2009.2033225
21. Yuen-Chai, T., et al.: Local chan-vese segmentation for non-ideal visible wavelength iris images. In: Conference on Technologies and Applications of Artificial Intelligence (TAAI), pp. 506–511. (2015). https://doi.org/10.1109/taai.2015.7407059
22. Jan, F.: Segmentation and localization schemes for non-ideal iris biometric systems. Signal Process. vol. 133, pp. 192–212. (2017). https://doi.org/10.1016/j.sigpro.2016.11.007
23. De Marsico, M., et al.: Insights into the results of MICHE I - mobile Iris CHallenge evaluation. Pattern Recogn. 74, 286–304 (2018). https://doi.org/10.1016/j.patcog.2017.08.028
24. Zuo, J., Schmid, N.: On a methodology for robust segmentation of nonideal iris images. IEEE Trans Syst Man Cybern B Cybern. 40(3), 703–18 (2009). https://doi.org/10.1109/TSMCB.2009.2015426
25. R-Kennell, L., W-Ives, R., M-Gaunt, R.: Binary morphology and local statistics applied to iris segmentation for recognition. In: International Conference on image processing, pp. 293–296. IEEE (2006). https://doi.org/10.1109/icip.2006.313183
26. Gangwar, A., et al.: A fast and robust iris segmentation framework for non-ideal iris images. In: 2016 International Conference on Biometrics (ICB), pp. 1–8. (2016).
27. Uhl, A., Wild, P.: Weighted adaptive hough and ellipsopolar transforms for real-time iris segmentation. 5th IAPR International Conference on

Biometrics (ICB). pp. 283–290. IEEE (2012). https://doi.org/10.1109/icb.2012.6199821.
28. Thornton, J., Savvides, M., Vijayakumar, B.V.K.: Robust iris recognition using advanced correlation techniques. In: International Conference Image Analysis and Recognition, pp. 1098–1105. Springer (2005). https://doi.org/10.1007/11559573_133
29. Vatsa, M., Singh, R., Noore, A.: Improving iris recognition performance using segmentation, quality enhancement, match score fusion, and indexing. IEEE Trans. Syst., Man, Cybern. B. 38(4), 1021–1035 (2008). https://doi.org/10.1109/tsmcb.2008.922059
30. Schuckers, S.A.C., et al.: On techniques for angle compensation in nonideal iris recognition. IEEE Trans. Syst., Man, Cybern. B. 37(5), 1176–1190 (2007). https://doi.org/10.1109/tsmcb.2007.904831
31. Karakaya, M., et al.: Gaze estimation for off-angle iris recognition based on the biometric eye model. In: Proceedings of SPIE - the International Society for optical Engineering, vol. 8712, pp. 87120F (2013)
32. Bolme, D.S., et al.: Off-angle Iris correction methods, pp. 497–518. LondonSpringer London (2016). https://doi.org/10.1007/978-1-4471-6784-6_21
33. R-Price, J., et al.: On the efficacy of correcting for refractive effects in iris recognition. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–6. IEEE (2007). https://doi.org/10.1109/cvpr.2007.383380
34. Arsalan, M., et al.: Deep learning-based iris segmentation for iris recognition in visible light environment. Symmetry. 9(11), 263 (2017). https://doi.org/10.3390/sym9110263
35. Rot, P., et al.: Deep multi-class eye segmentation for ocular biometrics. In: International work Conference on Bioinspired Intelligence, pp. 1–8. (2018). https://doi.org/10.1109/iwobi.2018.8464133
36. Jalilian, E., Uhl, A., Kwitt, R.: Domain adaptation for cnn based iris segmentation. In: Proceedings of the IEEE 16th International Conference of the Biometrics Special Interest Group (BIOSIG'17), pp. 51–60. (2017). https://doi.org/10.23919/biosig.2017.8053502
37. Bazrafkan, S., Thavalengal, S., Corcoran, P.: An end to end deep neural network for iris segmentation in unconstrained scenarios. Neural Network. 106, 79–95 (2018). https://doi.org/10.1016/j.neunet.2018.06.011
38. Osorio-Roig, D., et al.: Visible wavelength iris segmentation: a multi-class approach using fully convolutional neuronal networks. In: International Conference of the Biometrics Special Interest Group, pp. 1–5. (2018). https://doi.org/10.23919/biosig.2018.8553162
39. Karakaya, M.: Deep learning frameworks for off-angle iris recognition. In: 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems, pp. 1–8. IEEE (2018)
40. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. Adv. Neural Inf. Process Syst. 60, 1097–1105 (2012)
41. Hyoki, K., et al.: Quantitative electro-oculography and electroencephalography as indices of alertness. Electroencephalogr. Clin. Neurophysiol. 106(3), 213–219 (1998). https://doi.org/10.1016/s0013-4694(97)00128-4
42. Babcock, J.S., Pelz, J.B.: Building a lightweight eyetracking headgear. In: Proceedings of the 2004 Symposium on Eye Tracking research & applications, pp. 109–114. (2004). https://doi.org/10.1145/968363.968386
43. Duchowski, A.: Eye tracking methodology: Theory and practice. 01. (2007)
44. Nguyen, B.L., Chahir, Y., Jouen, F.: Free eye gaze tracking using Gaussian processes. Proceedings of the 2009 International Conference on Image Processing, Computer Vision vol. 2, CSREA, USA (2009)
45. Park, S., Spurr, A., Hilliges, O.: Deep pictorial gaze estimation. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 721–738. (2018). https://doi.org/10.1007/978-3-030-01261-8_44
46. Yang, T., et al.: Gaze angle estimate and correction in iris recognition. In: 2014 IEEE symposium on computational Intelligence in biometrics and identity Management (CIBIM). IEEE, pp. 132–138. (2014). https://doi.org/10.1109/cibim.2014.7015454
47. Lin, G., et al.: Multi-path refinement networks for high-resolution semantic segmentation. CoRR. abs/1611, 06612 (2016).
48. He, K., et al.: Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), abs/151203385 vol. 1, (2015)
49. Ronneberger, O., Fischer, P., Brox, T.: U-net: convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention, pp. 234–241. Springer (2015). https://doi.org/10.1007/978-3-319-24574-4_28
50. Jung, M., Chi, S.: Human activity classification based on sound recognition and residual convolutional neural network. Autom. ConStruct. 114, 103177 (2020). https://doi.org/10.1016/j.autcon.2020.103177

How to cite this article: Jalilian, E., Karakaya, M., Uhl, A.: CNN-based off-angle iris segmentation and recognition. IET Biome. 1–18 (2021). https://doi.org/10.1049/bme2.12052