Human Identification in Metaverse Using Egocentric Iris Recognition

Kuo Wang, Ajay Kumar
The Hong Kong Polytechnic University, Hong Kong

Abstract—In recent years, electronic glasses, including augmented reality (AR), virtual reality (VR), and mixed reality (MR) devices that connect the natural world and virtual world seamlessly, have significantly developed. Ocular images are inherently acquired during the immersion experiences from such devices, and can enable the verification of privileged identities during a live broadcast or meetings in virtual spaces. Lack of any such public database, and any specialized framework, is one of the key challenges in advancing iris recognition capability in metaverse or such virtual spaces. We introduce first or a new public iris images database, from 384 different subjects, to advance iris recognition using a generalized AR/VR device. Conventional iris recognition methods can only offer limited performance on such challenging iris images. This paper introduces an accurate and generalizable framework for iris recognition using AR/VR devices. The proposed framework is based on a convolutional network that uses a specifically designed shifted and extended quadlet loss function, enabling the network to accurately learn the discriminant iris features preserved in close-range and off-angle iris images. The framework introduced in this work can also adaptively consolidate the spatially corresponding features and abstract features from the other ocular details for more accurate matching. Thorough experimental results presented in this paper, using several classical and state-of-art iris recognition methods, are consistently outperforming and validate the effectiveness of the proposed approach with improvement of 96.30%, 30.58% and 27.23% for true accept rate (at false accept rate =0.0001), and 85.65%, 49.91% and 76.56% for equal error rate respectively.

Index Terms—Metaverse security, egocentric vision, iris recognition, periocular recognition, person authentication in virtual spaces, biometrics.

1 Introduction

Accurate personal identification using egocentric biometric images is one of the most critical and challenging tasks to meet the growing demand for stringent security. Iris recognition has emerged as one of the most accurate, low-cost, convenient methods for personal identification that also enables better hygiene due to contactless imaging. Iris patterns are unique among different subjects, even between identical twins, and they can be easily acquired using low-cost cameras. Iris images from billions of citizens have been acquired worldwide and incorporated into national ID programs to benefit citizens and effective e-governance. In recent years, electronic glasses, including augmented reality (AR), virtual reality (VR), and mixed reality (MR) devices that connect the natural world and virtual world seamlessly, have significantly developed and offer tremendous potential to augment human capabilities. Those wearable devices are a trend of electronic devices evolution, and they are potential substitutes for our smartphones if they are more portable [1]. However, significant algorithmic advances are required to achieve highly accurate user identification capabilities from such egocentric vision [2]. Foveated rendering [3] is one of the core technologies that enable a high-quality immersive experience with low computational cost and acquire egocentric iris images for gaze estimation. Eye interaction play vital important role in metaverse, and iris recognition is therefore the most feasible and accurate biometric for human identification in various metaverse applications, as shown in Figure. 1. This paper proposes to advance currently available iris recognition capabilities to meet ever-growing demand of stringent security during immersion experience from such AR/VR devices, with a multitude of applications in sectors such as online education, e-business, healthcare, and entertainment.

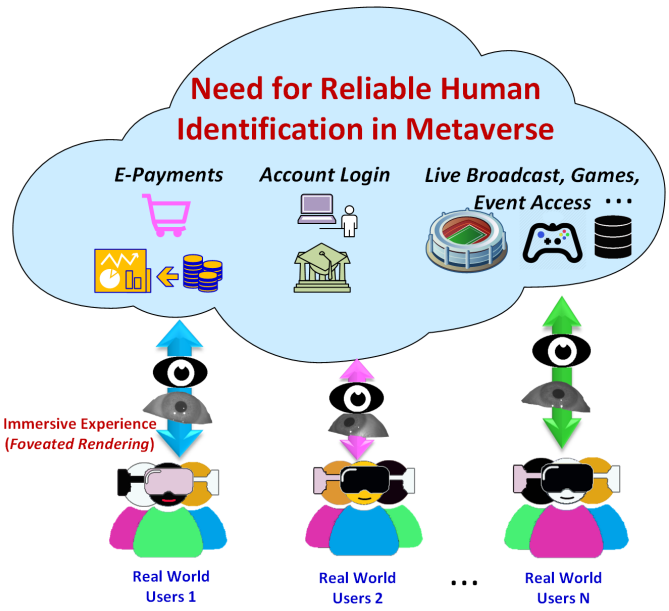


Fig. 1: Iris recognition for reliable human identification in various metaverse applications.

1.1 Related Work

Classical iris recognition methods use rubber-sheet model to normalize the conventional iris images and 2D Gabor filter-based phase encoding to generate the (feature) templates

TABLE 1: Summary of comparative performance from our work and state-of-the-art algorithms on three databases.

Reference	Egocentric Iris Recognition	Performance Evaluation		Comparative Results		
		Recognition	Open-Set Evaluation	Databases	TAR (FAR=0.0001)	EER
IrisCode(2D) [4]	No	No	No	(a)	22.16%	9.45%
				(b)	40.01%	12.66%
				(c)	22.86%	17.12%
IrisCode(1D) [5]	No	No	No	(a)	18.90%	12.57%
				(b)	39.52%	17.04%
				(c)	28.92%	12.58%
Ordinal Filter [6]	No	No	No	(a)	22.21%	8.80%
				(b)	39.80%	10.98%
				(c)	35.90%	12.65%
UniNet.v2 [7], [8]	No	No	No	(a)	43.78%	8.40%
				(b)	48.25%	10.04%
				(c)	43.01%	11.93%
DRFNet [9]	No	No	No	(a)	36.55%	6.27%
				(b)	53.80%	5.75%
				(c)	40.15%	11.05%
Maxout CNN [10]	No	No	No	(a)	31.87%	7.50%
				(b)	46.31%	10.56%
				(c)	17.19%	25.05%
Ours	Yes	Yes	Yes	(a)	85.94%	0.90%
				(b)	70.25%	2.88%
				(c)	54.72%	2.59%

(a) PolyU AR/VR Iris Images database (b) QFIRE database (c) CASIA-Iris-Degradation-V1 database.

[11]. These binarized templates are referred to as IrisCodes and efficiently matched using the Hamming distance. Variations of such an approach have used 1D log-Gabor filters to generate IrisCodes [5]. There are several other feature extraction approaches in the literature, e.g., using ordinal filters [6], sparse representation [12], phase correlation [13], etc., that have also shown exciting results to accurately match iris images that are acquired using conventional iris sensors. The stop-and-stare mode of the conventional iris recognition systems [14] limits its usage for other real-world applications. Therefore the literature has also reported notable efforts to accurately segment and match iris images that are acquired under-less constrained environment [15]. Daugman [16] introduced elliptical segmentation of iris images using active contours while reference [17] incorporated perspective transformation to recover frontal view iris image. Some recent efforts [18] using a trained network to match off-angle iris images are primitive, i.e., using manual or ground truth masks for segmentation, while such [19] learned features fail to preserve spatial correspondences.

More recent and promising efforts in the literature use deep neural networks and present exciting results in advancing the matching accuracy for iris recognition. Ganwar et al. [20] proposed an approach called DeepIrisNet using deep convolutional neural network (DCNN) technique for general iris recognition. This work is a direct application of classic CNN on iris recognition problem without any iris-specific optimization. Nguyen et al. [21] investigated the off-the-shelf iris features extracted from pretrained open-source CNNs, and the final decision is made using support vector machine (SVM). Reference [8] introduced a UniNet that included two branches of fully convolutional network (FCN) [22] to generate spatial corresponding IrisCode-like features and iris masks, to compute Hamming distances. Later in [7], this work employed the Mask R-CNN [23] to optimize the iris mask generation and further enhance the performance of proposed work. Another dense iris feature learning framework using dilated residual kernel is investigated in [9]. In [24], Boutros et al. investigated the iris recognition on an eye-gazing database [25] with state-of-the-art pretrained CNN models including ResNet [26], DenseNet [27] and MobileNetV3 [28].

Periocular information is inherently available in the raw iris images and therefore a range of methods have been incorporated for such periocular recognition. Reference [29] provides related summary of such references and methods in the literature. Attention-based deep networks in [30] can adaptively consider the importance of different ocular features and offer highly accurate results on multiple databases. Therefore such an approach can serve as a promising baseline in developing more effective periocular image matching techniques. Simultaneous use of iris and periocular features have advanced iris recognition accuracy and has been extensively studied in the literature [31], [32]. More recently, Zhang et al. [10] provided a promising framework from the maxout CNNs to enhance mobile-based iris recognition performance.

Despite exciting results and from the less-constrained iris recognition, there has been a lack of any attention for close-range iris recognition using off-angle iris images often acquired in wearable AR or VR devices. Some eye gaze tracking databases developed in the past, such as the OpenEDS dataset [25], are no longer in the public domain, do not provide user identity information for recognition task, were acquired from a small number of subjects and therefore not useful for iris recognition research.

1.2 Our Work and Contributions

This paper focuses on addressing critical limitations of currently available technologies for off-angle iris recognition from closer distances using head mounted devices, by developing a framework for accurately matching discriminative iris features from such images. We summarize the

main differences between our work and other related work in Table. 1 in the context of generalized iris recognition.

Foveated rendering [3] is a core technology that enables a high-quality immersion experience, with low computational cost, in head-mounted AR/VR sensors using gaze estimation. Images acquired for such gaze interaction provide a tremendous potential of seamless iris recognition, and we introduce the first such public database. We develop and incorporate a specific protocol that enables the acquisition of off-angle iris images from *nine different viewpoints* in such an AR device to ensure significant intra-class variability in this database (more details in Section 5). Therefore, a new database from this work is more realistic and provides challenging images for egocentric iris recognition.

Key contributions can be summarized as follows:

(a) Conventional methods for the iris recognition are not adequate to accommodate significant intra-class variations that are frequently observed in the iris images acquired from closer distances using head mounted AR/VR devices, and therefore can only offer limited performance. This paper introduces a new framework to accurately match such off-angle iris image. Our comparative experimental results detailed in Section 6 of this paper indicate outperforming results, over the state-of-the-art baselines, e.g.,improvement of 96.30% for true accept rate (at FAR=0.0001) and 85.65% in EER for egocentric iris recognition. We also present such comparative performance on other databases with off-angle images, e.g., improvement of 30.58% and 27.23% in TAR (at FAR=0.0001), 49.91% and 76.56% in EER respectively, to validate the effectiveness of our approach.

(b) Lack of any publicly available database for iris recognition using AR/VR devices is one of the critical limitations for much needed further research in this area. Therefore, this paper develops a two-session egocentric iris images database that includes images from 384 different subjects. To the best of our knowledge, this is the first two-session off-angle egocentric iris images database from AR/VR devices in the public domain. Each of the subjects contributed to 360 samples acquired from *nine different viewpoints* and is a large database for further research on egocentric iris recognition.

(c) Egocentric iris images that are acquired during the immersion experience using AR/VR sensor often present varying regions of effective iris pixels according to the gaze changes. Our framework adaptively consolidates such abstract features and spatially corresponding features for more accurate matching (more details in Section 4). Our comparative experimental results presented in Section 6 of this paper, using challenging protocols and other state-of-the-art methods, validate merit of the proposed approach.

In addition, we also develop and incorporate a specialized iris image quality detector and a more effective elliptical segmentation and normalization approach to address such egocentric iris recognition challenges. The rest of this paper is organized as follows. Section 2 presents our elliptical iris segmentation and normalization. Section 3 details our iris recognition network and proposed loss function. Section 4 provides the details of our periocular-assisted iris recognition framework. We introduce our new collected database in Section 5 and present the experimental results and discussion in Section 6. Finally, Section 7 presents a summary of key conclusions and future work in this area.

2 Iris Segmentation and Normalization

Accurate recognition of human identities using head mounted AR/VR sensors requires specialized preprocessing steps and they are introduced Figure. 2. These key steps in our iris recognition framework, to recover the normalized iris texture from such close distance and off-angle images, are briefly described in the following.

2.1 Pupil and Iris Boundary Detection

Our first step is the iris detection which involves pixel-level identification of iris and non-iris regions, i.e. excluding sclera, eyelash and source reflections to locate region of interest (ROI) and predict the binary iris mask. We use SOLOv2 [33] for such instance segmentation which is fine-tuned using 150 images manually segmented egocentric iris images [34]. Earlier references [16], [35] have shown that the boundary of off-angle iris images cannot be fitted using circles that are widely used in conventional methods [6], [8], [10], [36]. We therefore choose a non-conventional approach using an arc-support ellipse detector [37] to detect the pupil boundary and iris boundary on the ROI from the instance segmentation. In order to suppress the adverse effect from the eyelashes, we use bilateral filter on the images for the pupil ellipse detection. One can directly use the arc-support ellipse detector to locate the pupil boundaries without any parameter finetuning if the image quality is sufficient, i.e., adequate contrast, occlusion, etc. However, the iris boundary may not appear as a sharp edge due to the limbus effect [38]; therefore, detecting such boundaries in the ROI is difficult. To address such challenges, we use the mask segmentation results generated from the instance segmentation to assist in the accurate detection of iris boundaries. We use vertical Sobel operators with three different thresholds on the binary mask images to generate iris boundary candidates. Only the vertical one is deployed in our experiments to eliminate the eyelid's effect. The final ellipse for the iris boundary is selected from the three candidates using the following heuristics: (a) The pupil should be totally inside the predicted iris ellipse; (b) The predicted iris ellipse should include 90% of predicted iris pixels; (c) If two or more candidates satisfy both (a) and (b) mentioned above, we select the ellipse more similar to the pupil ellipse as the iris or limbus boundaries. The similarity is computed by their respective semi-major axis length over the semi-minor axis length.

The boundary information is represented with a ten elements vector $[x_p, y_p, l_p, s_p, \theta_p, x_i, y_i, l_i, s_i, \theta_i,]$, where (x_p, y_p) and (x_i, y_i) are pupil ellipse center and iris ellipse center; l_p and l_i are the length of the corresponding semi-major axis; s_p and s_i are the length of the corresponding semi-minor axis: θ_p and θ_i are orientation angle of the corresponding semi-major axis. The pupil radius r_p and iris radius r_i changes with the rotation angle $\phi \quad (0 < \phi \leq 2\pi)$. Such radius of detected pupil and iris ellipse can be computed as follows:

$$\begin{aligned} r_p^2 =& (l_p * cos(\theta_p) * cos(\phi) - s_p * sin(\theta_p) * sin(\phi))^2 \\ &+ (l_p * cos(\theta_p) * sin(\phi) + s_p * sin(\theta_p) * cos(\phi))^2 \\ r_i^2 =& (l_i * cos(\theta_i) * cos(\phi) - s_i * sin(\theta_i) * sin(\phi))^2 \\ &+ (l_i * cos(\theta_i) * sin(\phi) + s_i * sin(\theta_i) * cos(\phi))^2 \end{aligned} \tag{1}$$

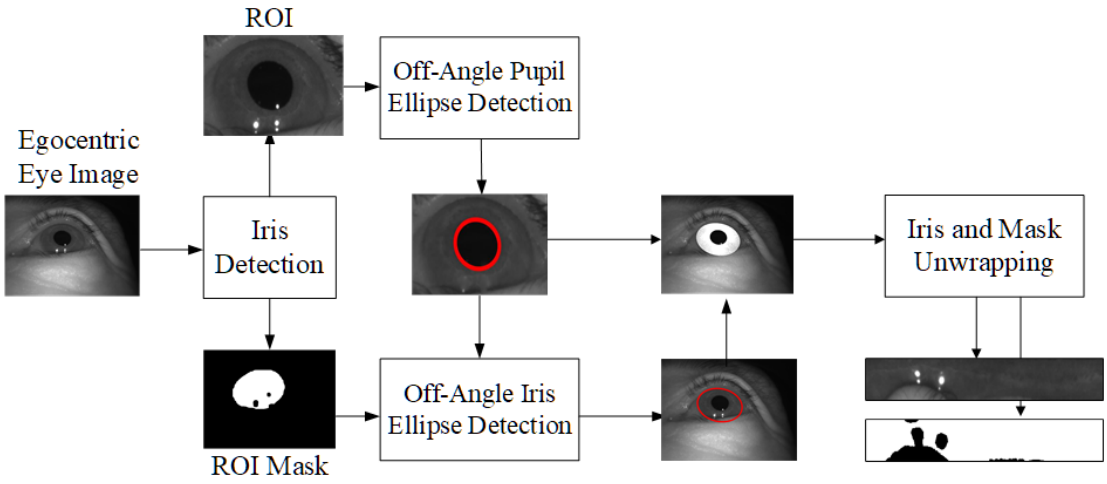


Fig. 2: Our key steps for egocentric iris segmentation and normalization.

Figure. 3 presents our database images with detected ellipses.

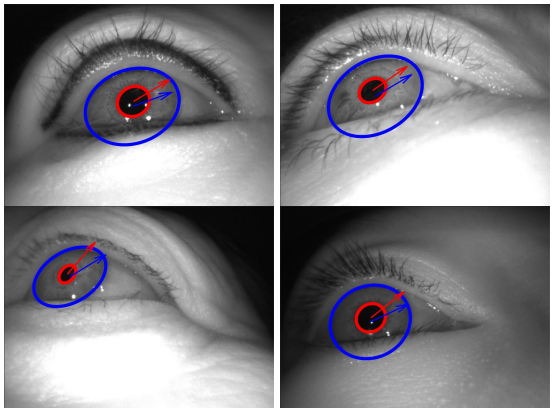


Fig. 3: Sample images with the orientations of two detected ellipses.

2.2 Iris Normalization

Similar to the conventional normalization [5], we achieve such normalization of iris images by replacing the circles using respective ellipses detected from pupil and iris boundaries shown in Figure. 4. Different from the circles, the radius of such ROI region changes with the angle in the ellipse. The center of the detected pupil ellipse is the reference point to sample the pixels along the radial axis with a specific angle. Each of the iris region pixels, from the intersection point $(x_p(\theta), y_p(\theta))$ of the pupil boundary to the intersection point $(x_i(\theta), y_i(\theta))$ of the iris boundary, are scanned along the radial axis which is oriented at angle θ, and then in the anti-clockwise direction for all other angles. The r is the radial distance between the pupil boundary and iris boundary. The locations of intersection points are computed using the respective slope from the pupil center. The mapping of every pixel in the off-angle iris image from the polar coordinate (r, θ) to the Cartesian coordinate (x, y) in the unwrapped iris image can be summarized as follows.

$$\begin{aligned} x(r,\theta) &= (1-r)x_p(\theta) + r x_i(\theta) \\ y(r,\theta) &= (1-r)y_p(\theta) + r y_i(\theta) \end{aligned} \tag{2}$$

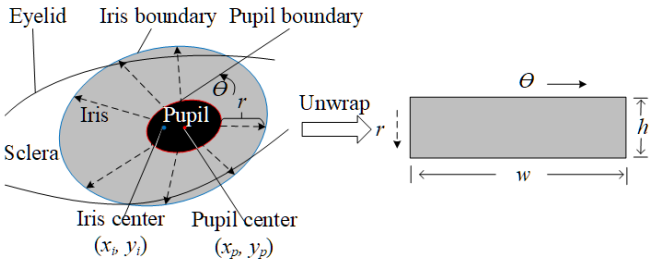


Fig. 4: Elliptical unwrapping model for off-angle iris images.

3 Egocentric Iris Recognition Network

Each segmented and normalized iris image is matched using a deep neural network that also uses respective iris mask. We consider a FCN [22] architecture to preserve spatial feature correspondences in the generated templates. Such a FCN-based architecture UniNet.v2 [7] can aggregate features from three different scales and be also employed in this work. The genuine or within-class distances in close

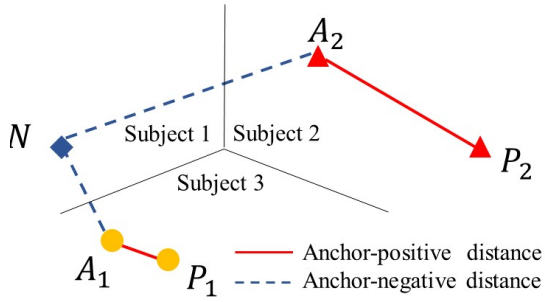


Fig. 5: Limitations of conventional triplet architecture for egocentric iris recognition.

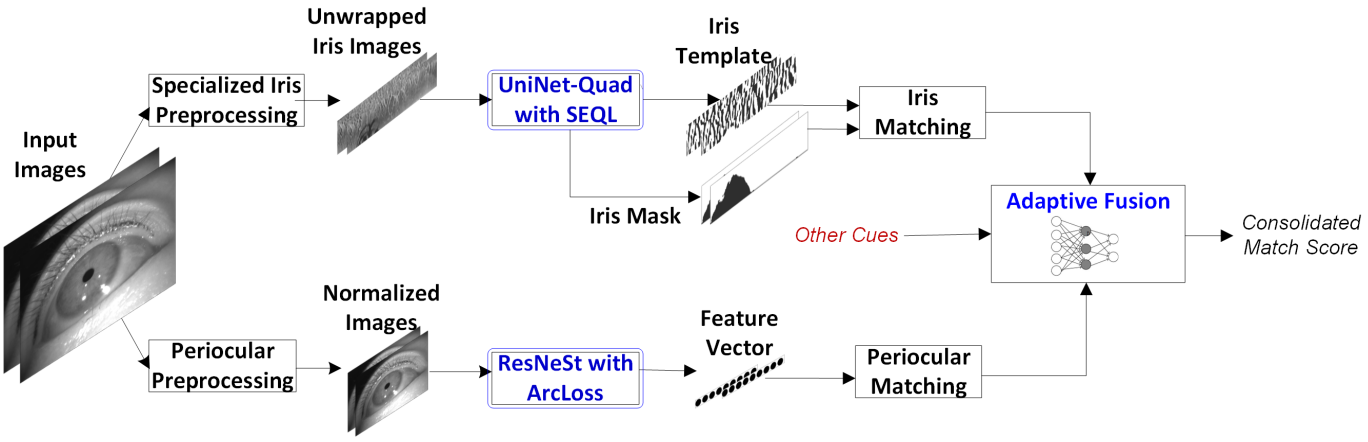


Fig. 6: Adaptive iris recognition framework for egocentric images acquired using AR/VR devices.

distance off-angle iris images are considerably higher than those for the conventional iris recognition [19], [39], [40], and can significantly degrade the matching accuracy. Therefore, we use a quad-based network architecture and a newly introduced loss function to accurately match iris images.

The motivation for using a quad architecture, instead of triplet architecture, is that the triplet architecture only considers the gap or distance between the genuine and impostor scores for the *same* anchor images. This problem can be observed from the conventional triplet loss L_{trip} i.e.

$$L_{trip} = \frac{1}{M}\sum_{i=1}^{M}[D_{A,P} - D_{A,N} + \alpha]_+ \tag{3}$$

where M is the batch size, $D_{A,P}$ is the distance between the anchor image and genuine image, $D_{A,N}$ is the distance between the anchor and imposter images, α is a hyperparameter controlling the margin between the anchor-positive and anchor-negative distances, $[.]_+$ means truncation to zero if the value is negative.

It should be noted that during the actual application, i.e. verification or identification, a global threshold will be set for all match scores instead of a unique threshold for each of the iris-class or the subject. This can be observed from an example case in Figure. 5 which shows five samples from three different subjects or classes when two triplets have already satisfied Eq. (3). We can however find that the genuine distance between A_2 and P_2 is *larger* than the imposter distance between A_1 and N. If we deploy this system in our application, it will consider that A_1 and N are from the same subject whereas A_2 and P_2 are from different subjects. Therefore, inspired by quadruplet [41], one more constrains is added to mitigate this effect and ensure the distance D_{A_2,P_2} is less than $D_{A_1,N}$. This can be achieved by adding one more branch for the negative sample from a different class regarding both the anchor and negative sample. The new sample N' is from a third subject which is different from other samples, i.e., A, P, N. Therefore, we can compute one more imposter score $D_{N,N'}$ between the negative sample N and N'. The quadlet loss L_{quad} is defined as follows.

$$L_{quad} = \frac{1}{M}\sum_{i=1}^{M}([D_{A,P} - D_{A,N} + \alpha]_+ + [D_{A,P} - D_{N,N'} + \alpha]_+) \tag{4}$$

Similar to the IrisCode Hamming distance calculation, we perform the horizontal and vertical bit shift in our training process. Also, the mask information is utilized to filter out the noisy pixels in our loss calculation. The overall quadlet loss with bits shift and mask is our new proposed shifted and extended quadlet loss.

4 Adaptive Egocentric Iris Recognition

Raw images acquired from the AR/VR sensors are essentially periocular images with discriminative ocular features. Therefore, we can also match such images to generate match scores that can be adaptively consolidated with iris match scores. Such an approach is illustrated in Figure. 6 and discussed in following.

We incorporate a powerful feature extractor (ResNeSt) [42] pre-trained with ImageNet [43], and ArcLoss [44] function to learn ocular features from the input images. Any effective mechanism to simultaneously utilize the varying cues from the iris and periocular matching should carefully consider their image quality. A varying number of valid (ROI) iris pixels, incorporated to generate respective iris match scores, indicates the reliability of such match scores and can serve as an important cue to adaptively consolidate the match score. We define the number of valid pixels as the *mask rate* and is computed as the fraction of occluded iris pixels, using the iris masks, in two matched iris templates. The differences in mask rates can be utilized to adaptively reinforce the periocular information. Therefore, we incorporate a multilayer perceptron network (MLP) to consolidate such multiple pieces of discriminative details and generate a more reliable match score.

Our MLP is composed of three fully connected layers. It will receive a seven-element feature vector as an input, including the iris match score, periocular match score, mask rate, two eye image quality score from our quality checking network, and two eye orientation information from the segmentation steps. The eye orientation is the direction of the detected ellipse during iris segmentation. The network

is trained offline using the genuine and impostor pairs from the respective training dataset. The trained network generates consolidated match scores, from the softmax value in the final output layer, in the 0-1 range.

5 Database Acquisition and Organization

A general-purpose sensor is employed to acquire the iris images for gaze estimation that also enable immersion experience in AR/VR environment. Therefore, the acquired images reflect real-world and dual-use for the subject identification using close-range and off-angle iris patterns. The subjects who provided images in this database were volunteers from different ethnic communities, including Chinese, Indian and European. These volunteers consented to support our research, were not paid any honorariums, and we did not record their personal details. They wore our head-mounted AR device and observed a 3D rectangular pattern with a green point displayed on the glasses. All the volunteers were requested to gaze sequentially at nine different locations as illustrated in Figure. 7, while the camera captured their iris images. Sample iris images acquired from different gaze points from a volunteer are shown in Figure. 8. We acquired 20 different image samples from one eye at each of the gaze points, and therefore a total of 360 image samples were acquired from one subject during one session.

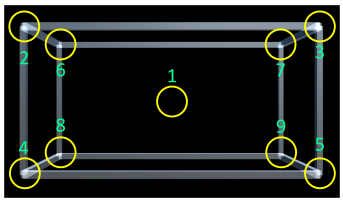


Fig. 7: Gaze points displayed on head-mounted display during the egocentric iris database acquisition.

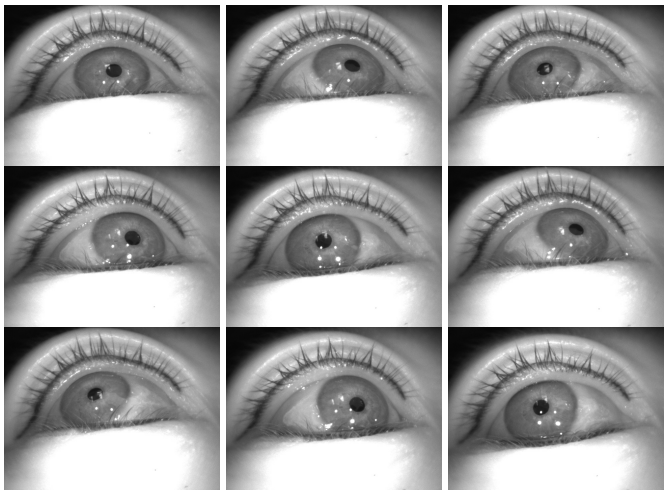


Fig. 8: Sample iris images from a user in the acquired database.

We acquired first session image samples from 384 different subjects while only 114 were available for the second session data acquisition. The *minimum* interval between two image sessions was four months. Therefore, the entire database consists of 138,240 image samples from the first session and 41,040 image samples from the second session. This entire database is publicly available [34] to advance further research in this area. Among potential iris image quality degradation listed in [45], Table. 2 provides list of these problems in the context of AR/VR iris images database acquired during this research. Figure. 9 illustrates the distribution of estimated iris image orientations from the egocentric views in the acquired database. Average number of effective iris pixels and equivalent iris diameter in this database is 18,434.39 and 116.26 pixels respectively.

TABLE 2: Common image quality degradation during egocentric iris acquisition.

Possible Acquisition Problem	Egocentric Iris Database
Occlusion by Finger	No
Motion Blur	Yes
Mislabeled Eyes	No
Eye Rotation	No
Closed or Squinting Eye	Yes
Specular Highlights from Glasses	No
Off-axis Gaze	Yes
Highly Dilated or Constricted Pupil	Yes
Focus Blur	Yes
Iris Absent from Image	No
Poor Illumination and Low Contrast	Yes

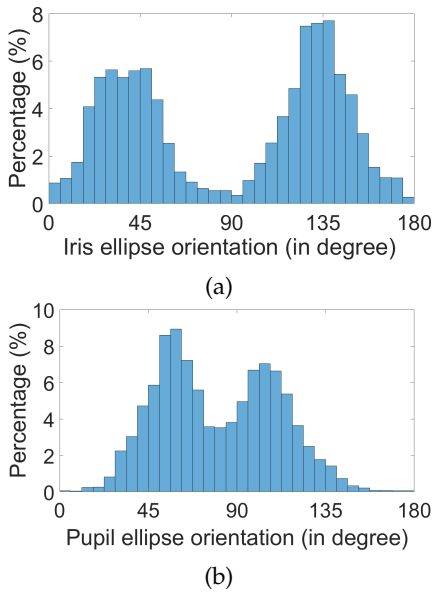


Fig. 9: Estimated distribution of (a) iris and (b) pupil ellipse orientations, from egocentric view, in the database images.

We then train a quality checker using MobileNetV3 [28] to automatically discard very low-quality samples while considering the NIST guidelines except for the off-angle requirement [45]. This quality checker can discard the closed eye, serious off-axis gaze, motion blur and very low contrast, while such samples are shown in Figure. 10. Table. 3 summarizes the statistics of acquired egocentric iris images database and the number in brackets represent images after the quality check.

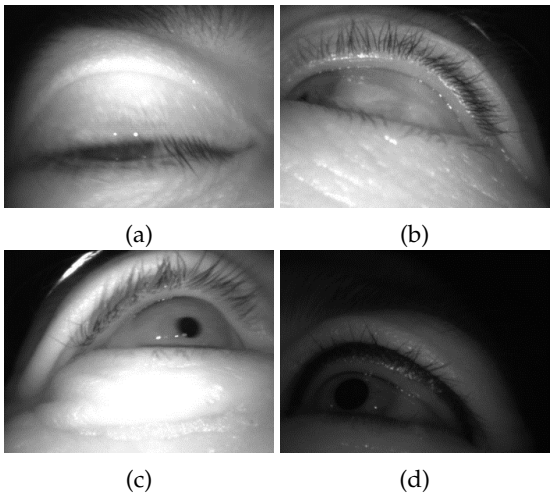


Fig. 10: Sample images from egocentric iris image database with image quality degradation including (a) closed eye (b) off-axis gaze (c) motion blur (d) very low contrast.

TABLE 3: Summary of statistics from images in PolyU AR/VR Iris Images database.

	Session 1		Session2	
	Left	Right	Left	Right
Number of Subjects	384	384	114	114
Number of Samples	69,120 (52,563)	69,120 (53,393)	20,070 (18,107)	20,070 (17,942)
Resolution	640*480	640*480	640*480	640*480

6 Experiments and Results

We performed extensive experiments to evaluate the performance of the proposed approach for AR/VR sensor-based egocentric iris recognition. The details on the exact match protocols and the corresponding reproducible [34] results are provided in the following section. We use the elliptical segmentation approach proposed in Section 2 for the iris segmentation on all the databases used in this paper. The thresholds of the vertical Sobel operator were empirically fixed as 0.35, 0.4, and 0.45 for all our experiments. Our iris normalization generated unwrapped iris images of $512 * 64$ pixels in Figure. 4 for all the experiments and databases used in this paper. In our quadlet training, we set the horizontal shift u as 32, stepsize t as 2, vertical shift v as 5 and margin α as 0.15. The network is trained using SGD with learning rate of 0.001 and momentum of 0.9. The batch size is fixed as 8 for the 80,000 iterations.

6.1 Databases and Protocols

6.1.1 PolyU AR/VR Iris Images Database

We perform experiments using the two-session and one-session (all-to-all) protocol. Two session experiments use 1,967 images from 202 different eyes in the second session to compose a probe set and 1,879 samples in the first session to compose a gallery dataset. Such a two-session match protocol will generate 1,967 genuine and 395,367 imposter match scores. One-session experiments selected samples from 242 subjects that are stated in a text file made available from [34]. We use image samples from the first 60 subjects for the training, whereas ten samples from the rest of 182 subjects are used for the performance evaluation. Test samples are evenly selected from the five different directions while considering the gaze points at the same corner as the same direction. Therefore, there will be 3,640 samples in the performance evaluation test dataset, and 16,380 genuine and 6,606,600 imposter match scores will be generated. The reason for such selection in one-session experiment is to ensure that there are at least two good-quality image samples at each gazing direction from both the eyes of each of the subjects. Similarly, we also attempt to select at least two good-quality images from each subject at each gaze direction for the two-session experiments. We can therefore only use images from 101 different subjects in two-session experiments.

6.1.2 Quality Face and Iris Research Ensemble Database

The Quality in Face and Iris Research Ensemble (Q-FIRE) database [46] is publicly available, providing off-angle iris images acquired using the OKI IRISPASS EQ5016A sensor. Although this database is not acquired from any AR/VR sensor that commonly generates close-range off-angle iris images, the iris images available in this database can help us ascertain the proposed approach's effectiveness for off-angle iris recognition. We select a subset of images acquired from a 5-feet distance under higher illumination. The original image samples illustrate the upper part of the face, and we automatically segment the periocular region with a Fast-RCNN based detector [47]. The samples in our experiments are chosen from the front, up, down, left, right, as illustrated in Figure. 11. We selected iris samples from 44 subjects (88 classes) as our training set, 2,640 iris samples from the last 132 subjects (264 classes) as our testing set. Therefore, this database in our experiments generates 11,880 genuine and 3,471,600 imposter match scores.

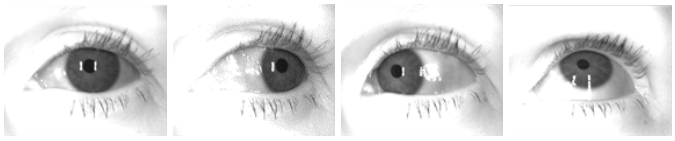


Fig. 11: Sample off-angle iris images from QFIRE database.

6.1.3 CASIA-Iris-Degradation-V1 Database

The CASIA-Iris-Degradation-V1 database [48] consists of 36,539 images from 255 subjects. The challenging samples are divided into 15 categories: illumination, off-angle, occlusion, etc. We use the off-angle subset composed of iris samples from frontal, top-left, down-right, down-left and down-right views, as shown in Figure. 12. We selected iris samples from 60 subjects (120 classes) as our training set, 2,088 iris samples from the last 95 subjects (190 classes) as our testing set. Therefore, this testing set in our experiments generates 13,277 genuine scores and 2,165,551 imposter scores.

6.2 Experimental Results

We performed a range of experiments to validate the effectiveness of the proposed framework. These experimental

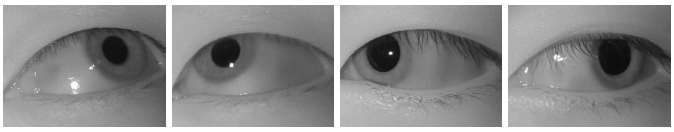
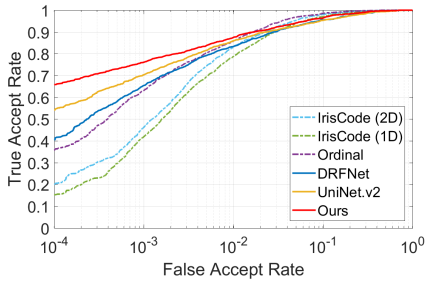
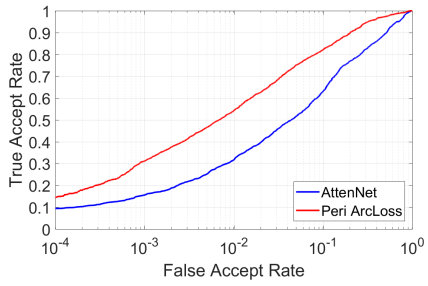


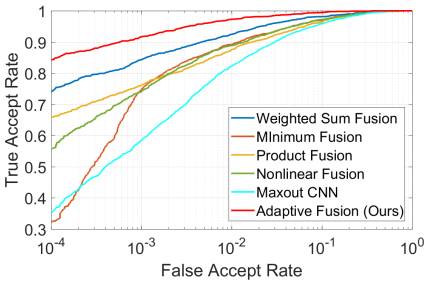
Fig. 12: Sample off-angle iris images from CASIA ID-V1 database.



(a)



(b)



(c)

Fig. 13: Comparative results for two-session verification performance on PolyU AR/VR Iris Images database. ROC plots from (a) iris, (b) periocular and (c) the combined performance.

results are presented using the receiver operating characteristic (ROC) curves to evaluate the performance for the verification. We also evaluate the generalization capability of our model using the cross-database performance evaluation.

6.2.1 Two-Session Matching Performance

Matching AR/VR sensor-based iris images from the second session with those from the first session acquired during the registration can provide most realistic performance evaluation under the popular deployment scenario. Therefore, this performance evaluation is firstly presented using the PolyU AR/VR Iris Images database. We also present comparative performance with many other classical and deep network-based competing benchmark methods. The original IrisCode (2D) introduced by Daugman [11] has been implemented in OSIRIS [4] and is an important baseline for the conventional iris recognition. IrisCode (1D) [5] is also a popular and widely used benchmark for iris recognition performance and uses 1D log-Gabor filter to generate IrisCodes. Another classical algorithm that can provide ordinal measurements using multi-lobe ordinal filters [6] was also considered during our evaluation. In order to ensure fairness, we use same test samples to evaluate all baseline methods under the same protocol as stated in Section 6.1. We also compared with competing deep neural network-based methods that were proposed for the conventional iris recognition. The UniNet.v2 [7] employs the fully convolutional network to generate binarized feature templates while DRFNet [9] generates consolidates features from the residual network using dilated kernels. It is important to note that *only* the iris encoding algorithms from these baseline methods are employed, and we use our segmentation (Section 2) to ensure fairness in such evaluation as the segmentation methods used in those baseline algorithms are not designed for our problem i.e. close-range and off-angle iris recognition. Since the method proposed for matching off-angle periocular images in this work is new or not explored in the literatures, we also provide comparison of our periocular matching model with the state-of-the-art AttenNet [30] method. Our adaptive fusion results are also comparatively evaluated with the maxout CNNs [10] and some other classical fusion algorithms, including weighted-sum, minimum, product, and non-linear fusion [49]. Comparative ROC plots for this evaluation are presented in Figure. 13. Table. 4 presents comparative summary of equal error rates (EER) and genuine or true accept rate (TAR) at low (10^{-4}) false accept rate (FAR).

TABLE 4: Comparative summary from two-session egocentric iris recognition performance evaluation.

Category	Methods	Two Session TAR(%)	Two Session EER(%)
Iris	IrisCode (2D) [4]	20.37	5.65
	IrisCode (1D) [5]	15.53	5.50
	Ordinal [6]	35.54	5.56
	DRFNet [9]	41.38	6.31
	UniNet.v2 [7]	54.75	6.39
	Ours	65.89	5.43
Peri.	AttenNet [30]	9.80	22.04
	Peri ArcLoss (Ours)	14.38	14.05
Fusion	Wighted Sum Fusion	74.50	3.61
	Minimum Fusion	32.28	4.77
	Product Fusion	65.57	4.28
	Nonlinear Fusion [49]	56.33	5.43
	Maxout CNN [10]	36.26	6.07
	Adaptive Fusion (Ours)	84.46	1.88

Additional performance comparisons using off-the-shelf CNN features, a COTS matching, and related compulsory analysis are provided in the attached *Appendix A-C*.

6.2.2 One-Session Matching Performance

Matching one-session iris images can generate large number of match scores [8] and such performance evaluation is also presented in this section. The comparative experimental

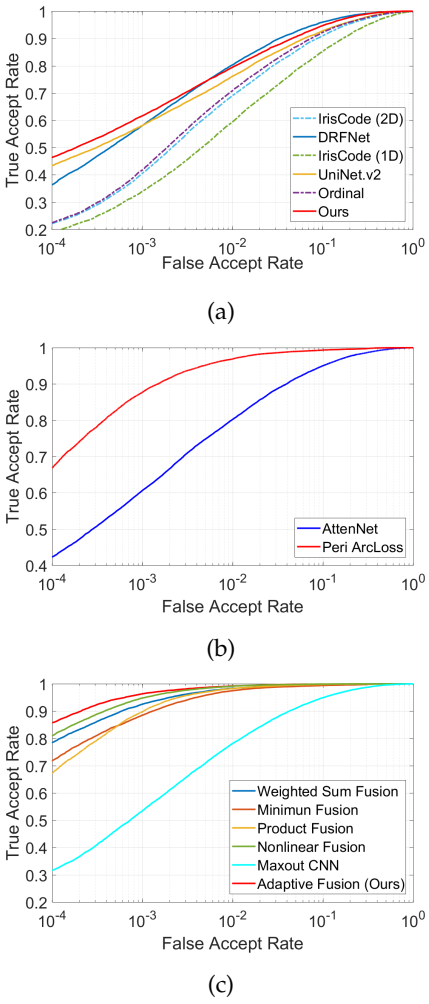


Fig. 14: Comparative results for one-session performance on PolyU AR/VR Iris Images database. ROC curves of (a) iris, (b) periocular, and (c) the combined performance.

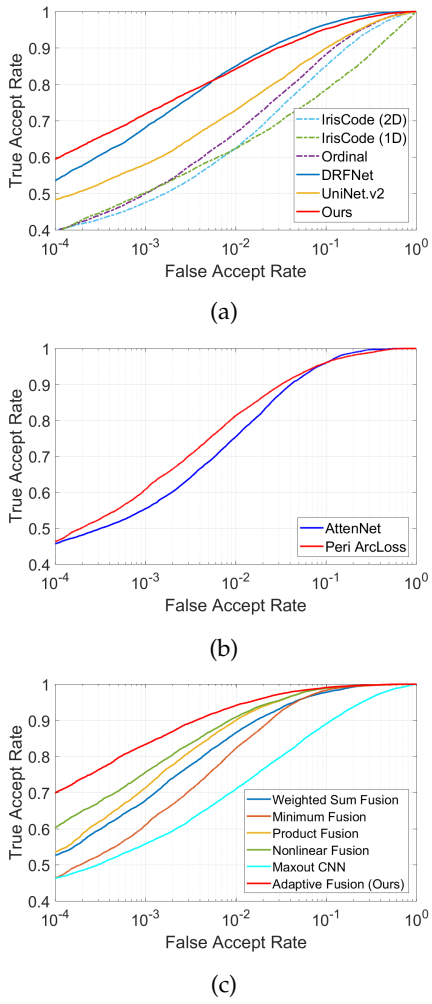


Fig. 15: Comparative results for one-session performance on QFIRE database. ROC curves of (a) iris, (b) periocular, and (c) the combined performance.

results from our approach and the respective benchmark methods are presented in Figure. 14 for PolyU AR/VR Iris Images database. We also present respective ROC results in Figure. 15 using the QFIRE database protocol detailed in Section 6.1.2 and ROC results in Figure. 16 using the CASIA ID-V1 database protocol detailed in Section 6.1.3. Table. 5 presents comparative summary of EER and TAR at low (10^{-4}) FAR constantly improved on the performance.

6.2.3 *Cross-Database Performance*

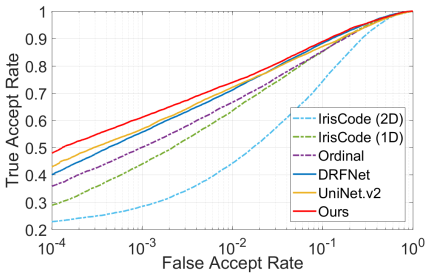
Although there is no other publicly accessible close-range and off-angle egocentric iris database to the best of our knowledge, we still attempted to ascertain cross-database performance using the model trained on PolyU AR/VR Iris Images database and evaluated on the CASIA ID-V1 database. Such cross-database evaluation can help to ascertain the generalization capability from the proposed model. Here we employ the directly trained model using PolyU AR/VR Iris Images database and use it to ascertain the matching performance for the CASIA ID-V1 database (Section 6.1) *without* any finetuning. We compare the performance with the other competing deep network-based algorithms. The number of test images is the same as respective database in the previous experiments. The comparative performance from the cross-databases experiment is shown in Figure. 17 while Table. 6 summarizes respective EER and TAR at low FAR (10^{-4}).

6.3 Discussion

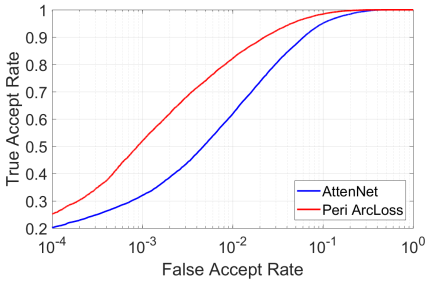
In order to ascertain the effectiveness of the proposed loss function for the iris recognition using AR/VR sensors, we also performed comparisons with popular CNN models using other loss functions. We performed comparisons using the quadlet loss with the same UniNetv2 architecture and additional comparison using the triplet loss. We also compare another popular baseline method, DeepIrisNet [20], which has shown promising results for conventional iris recognition. The hyperparameters for the effective training of respective network architectures were carefully investigated to achieve the best possible performance. All these comparisons were performed PolyU AR/VR Iris Images database in this work (Section 5) and the resulting performance is shown from the ROCs in Figure. 18. The poor performance from the quadlet loss function strongly suggests

TABLE 5: Comparative summary from one-session matching performance.

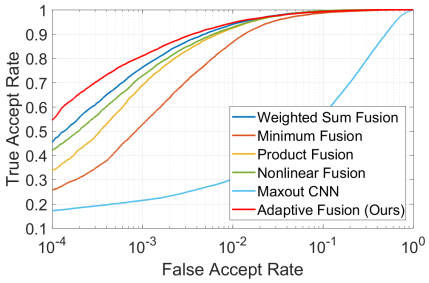
Category	Methods	Egocentric Iris Database		QFIRE Database		CASIA ID-V1 Database	
		TAR(%)	EER(%)	TAR(%)	EER(%)	TAR(%)	EER(%)
Iris	IrisCode (2D) [4]	22.16	9.45	40.01	12.66	22.86	17.12
	IrisCode (1D) [5]	18.90	12.57	39.52	17.04	28.92	12.58
	Ordinal [6]	22.21	8.80	39.80	10.98	35.90	12.65
	DRFNet [9]	36.55	6.27	53.80	5.75	40.15	11.05
	UniNet.v2 [7]	43.78	8.40	48.25	10.04	43.01	11.93
	Ours	46.53	7.19	59.96	6.49	48.14	10.69
Periocular	AttenNet [30]	42.48	6.79	47.17	6.25	20.17	7.53
	Peri ArcLoss (Ours)	66.89	1.87	46.22	6.18	25.31	4.75
Fusion	Weighted Sum Fusion	78.90	1.37	53.08	4.66	45.72	2.62
	Minimum Fusion	72.25	1.44	54.11	4.75	33.93	3.68
	Product Fusion	67.11	1.39	46.39	3.74	25.83	2.77
	Nonlinear Fusion [49]	81.33	0.95	60.41	3.70	42.22	2.80
	Maxout CNN [10]	31.87	7.50	46.31	10.56	17.19	25.05
	Adaptive Fusion (Ours)	85.94	0.90	70.25	2.88	54.72	2.59



(a)



(b)



(c)

Fig. 16: Comparative results for one-session performance on CASIA ID-V1 database. ROC curves of (a) iris, (b) periocular, and (c) the combined performance.

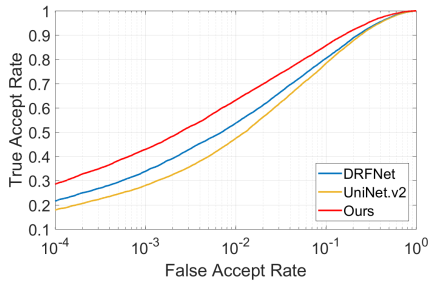


Fig. 17: Comparative ROC from the cross-database performance evaluation.

TABLE 6: Performance summary from cross-database evaluation.

Methods	CASIA ID-V1 Database	
	TAR(%)	EER(%)
DRFNet [9]	21.65	14.91
UniNet.v2 [7]	17.90	15.31
Ours	28.58	12.26

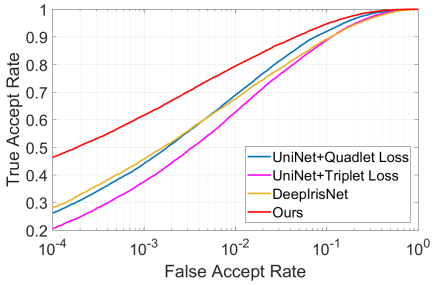


Fig. 18: Comparative ROC from different loss functions.

that it is necessary to account for the spatial bit translation to accurately learn the spatially corresponding features from the close-range and off-angle iris images.

Although our results from the proposed framework to adaptively match iris images have shown promising results, significant work is required to improve the performance for large-scale matching and deployment. Egocentric iris images from wearable AR/VR sensors can exhibit significantly high intra-class variations, and accurate localization of iris boundaries in many such degraded quality images is quite challenging.

7 Conclusions and Further Work

This paper has developed a new framework to accurately match off-angle iris images acquired from closer distances using head-mounted AR/VR devices. We also designed a shifted and extended quadlet loss function to provide effective supervision in learning discriminative features for the convolutional neural network considered in this work. Our extensive and reproducible experimental results presented in Section 6 achieve outperforming results and validate the merit of the proposed framework for iris recognition from the popular AR/VR devices that inherently acquire ocular images under near-infrared illumination for gaze estimation. We also developed and introduce a new database in the public domain [34] to advance further research and development in this area. The focus of this work is on egocentric iris recognition, instead of iris segmentation [50], which would require generation of a number of manually annotated iris and pupil regions in egocentric iris images. Therefore, further extension of this work should focus on more accurate localization of iris boundaries, or more robust iris features from egocentric iris images with significantly high intra-class variations.

Appendix A
Comparison with off-the-shelf CNN features

We also perform two-session experiments using the same matching protocol to compare with the off-the-shelf features approaches [21]. There are five different feature extractors pretrained with ImageNet [43] employed to provide off-the-shelf features using normalized iris images, including AlexNet [51], VGGNet [52], InceptionNetV3 [53], ResNet152 [26] and DenseNet201 [27]. The feature vectors fed into the support vector machine (SVM) are selected from the peak layers for performance, including layer 7 for AlexNet, layer 9 for VGGNet, layer 10 for InceptionNetV3, layer 12 for ResNet152 and layer 5 for DenseNet201 respectively. According to the description in origin paper, all the CNN models are not finetuned and only one-over-all multiclass SVMs are trained with corresponding training data. We use the predicted probability as the matching score, and plot the ROC in Figure. 19 while EER results are shown in Table. 7. We can find that our approaches provide superior performance than the off-the-shelf approaches. The off-the-shelf features cannot effectively address the iris recognition problem under such challenging scenario without any iris-specific design, and one-over-all learning strategy cannot minimize the intra-class difference.

Appendix B
The Model Complexity Analysis

We also evaluate the complexity of the iris recognition framework to ascertain its feasibility for the deployment. Table. 8 presents comparative summary of the computational time for the feature extraction and storage requirements. We use a Ubuntu 18.04 machine with i9-7900X CPU, 32GB RAM, and 11GB 1080ti GPU. These results indicate that the space and time complexity of our trained model is not large and quite suitable for the online iris recognition. The run time requirement for the iris image quality checker, iris and mask detection, ellipse detection with iris segmentation and normalization, are respectively 11.3, 62.4 and 223.7 milliseconds (ms), which can be further reduced by the code optimization and choice of more simplified model.

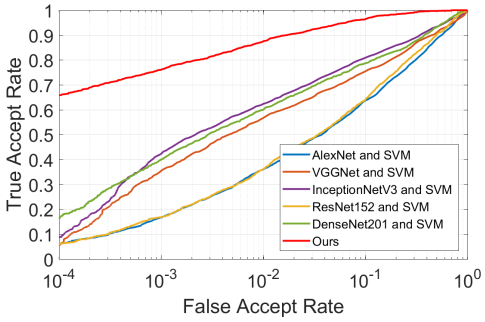


Fig. 19: Comparative ROC results from off-the-shelf features and our approach using two-session matching protocol.

TABLE 7: Comparative EER results from off-the-shelf features and our approach using two-session matching protocol.

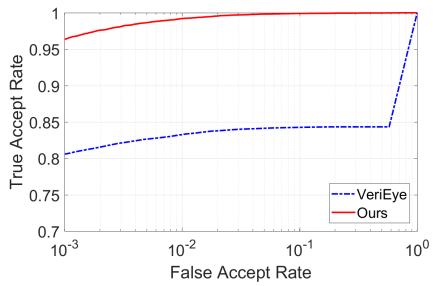
Algorithm	EER
AlexNet and SVM	23.91%
VGGNet and SVM	19.30%
InceptionNetV3 and SVM	15.78%
ResNet152 and SVM	22.45%
DenseNet201 and SVM	17.27%
Ours	5.43%

TABLE 8: Comparative summary of complexity analysis for baseline models.

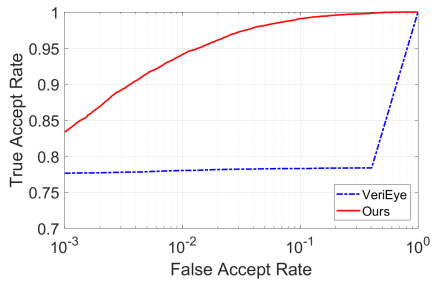
	Parameters	Extraction Time (ms)
DRFNet [9]	125K	8.1
Maxout CNN [10]	4,095K	10.5
DeepIrisNet [20], [53]	55,420K	13.8
DenseNet201 [27]	20,242K	24.6
Ours	129K	8.9

Appendix C
Comparison with COTS

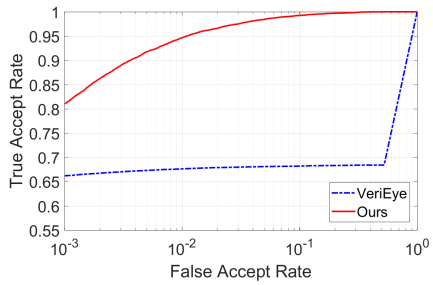
Although we have already provided reproducible comparisons with widely cited baseline methods that have shown competitive performance in the literature, it may also be interesting to ascertain comparisons with the commercial-off-the-shelf (COTS) which are believed to be widely used and optimized for the real-world deployment. VeriEye is one such popular iris recognition SDK developed by Neurotechnology [54]. Its core functions are not open source so we do not know its processing details. This SDK accepts entire eye images as the input and can generate one match score for each pair of inputs, just like for our one session all-to-all matching. Therefore, we provide comparisons from our method on PolyU AR/VR Iris Images dataset in this



(a)



(b)



(c)

Fig. 20: Comparative ROC plots from the VeriEye SDK on (a) PolyU dataset (b) QFIRE dataset and (c) CASIA dataset.

work and QFIRE dataset using the same one-session protocol. The ROC curves from such comparisons are shown in Figure. 20.

The VeriEye SDK has built-in quality assessment function that does not generate match scores for low-quality images and this can explain the nature of the above performance curve. The ROC plots in Figure. 20 indicate outperforming results on both datasets, except for the low FAR side on QFIRE dataset. This COTS software is not an open-source and therefore we can only present an indicative performance as cannot ascertain its segmentation process.

REFERENCES

[1] R. T. Azuma, "A survey of augmented reality," *Presence: Teleoperators & Virtual Environments*, vol. 6, no. 4, pp. 355–385, 1997.

[2] A. Betancourt, P. Morerio, C. S. Regazzoni, and M. Rauterberg, "The Evolution of First Person Vision Methods: A Survey," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 744–760, 9 2014.

[3] A. Patney, M. Salvi, J. Kim, A. Kaplanyan, C. Wyman, N. Benty, D. Luebke, and A. Lefohn, "Towards foveated rendering for gaze-tracked virtual reality," *ACM Transactions on Graphics*, vol. 35, no. 6, pp. 1–12, 2016.

[4] N. Othman, B. Dorizzi, and S. Garcia-Salicetti, "OSIRIS: An open source iris recognition software," *Pattern Recognition Letters*, vol. 82, pp. 124–131, 10 2016.

[5] L. Masek, "Recognition of human iris patterns for biometric identification," Master's thesis, University of Western Australia, 2003.

[6] Z. Sun and T. Tan, "Ordinal Measures for Iris Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 12, pp. 2211–2226, 2008.

[7] Z. Zhao and A. Kumar, "A deep learning based unified framework to detect, segment and recognize irises using spatially corresponding features," *Pattern Recognition*, vol. 93, pp. 546–557, 9 2019.

[8] Z. Zhao and A. Kumar, "Towards More Accurate Iris Recognition Using Deeply Learned Spatially Corresponding Features," in *Proceedings of the IEEE International Conference on Computer Vision*, 12 2017, pp. 3829–3838.

[9] K. Wang and A. Kumar, "Toward More Accurate Iris Recognition Using Dilated Residual Features," *IEEE Transactions on Information Forensics and Security*, vol. 14, pp. 3233–3245, 2019.

[10] 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.

[11] J. Daugman, "How Iris Recognition Works," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, 2004.

[12] J. K. Pillai, V. M. Patel, R. Chellappa, and N. K. Ratha, "Secure and robust iris recognition using random projections and sparse representations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 9, pp. 1877–1893, 2011.

[13] K. Miyazawa, K. Ito, T. Aoki, K. Kobayashi, and H. Nakajima, "An effective approach for iris recognition using phase-based image matching," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1741–1756, 10 2008.

[14] J. Thornton, M. Savvides, and B. V. Kumar, "A bayesian approach to deformed pattern matching of iris images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[15] H. Proença, "Quality assessment of degraded iris images acquired in the visible wavelength," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 1, pp. 82–95, 2010.

[16] J. Daugman, "New methods in iris recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 37, no. 5, pp. 1167–1175, 2007.

[17] S. A. Schuckers, N. A. Schmid, A. Abhyankar, V. Dorairaj, C. K. Boyce, and L. A. Hornak, "On techniques for angle compensation in nonideal iris recognition," *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 37, no. 5, pp. 1176–1190, 10 2007.

[18] M. Karakaya, D. Barstow, H. Santos-Villalobos, J. Thompson, D. Bolme, and C. Boehnen, "Limbus impact on off-angle iris degradation," *Proceedings - 2013 International Conference on Biometrics, ICB 2013*, 2013.

[19] M. Karakaya, "Deep learning frameworks for off-angle iris recognition," in *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems, BTAS 2018*. Institute of Electrical and Electronics Engineers Inc., 7 2018.

[20] A. Gangwar and A. Joshi, "DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition," in *International Conference on Image Processing*, 2016, pp. 2301–2305.

[21] K. Nguyen, C. Fookes, A. Ross, and S. Sridharan, "Iris recognition with off-the-shelf CNN features: A deep learning perspective," *IEEE Access*, vol. 6, pp. 18 848–18 855, 2018.

[22] J. Long, E. Shelhamer, and T. Darrell, "Fully Convolutional Networks for Semantic Segmentation," in *IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 6 2015, pp. 3431–3440.

[23] K. He, G. Gkioxari, P. Dollar, and R. Girshick, "Mask R-CNN," in *Proceedings of the IEEE International Conference on Computer Vision*, 3 2017, pp. 2980–2988.

[24] F. Boutros, N. Damer, K. Raja, R. Ramachandra, F. Kirchbuchner, and A. Kuijper, "On benchmarking iris recognition within a head-mounted display for ar/vr applications," in *2020 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2020, pp. 1–10.

[25] C. Palmero, A. Sharma, K. Behrendt, K. Krishnakumar, O. V. Komogortsev, and S. S. Talathi, "Openeds2020: open eyes dataset," *arXiv preprint arXiv:2005.03876*, 2020.

[26] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 770–778.

[27] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the*

IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 4700–4708.

[28] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan *et al.*, "Searching for MobileNetV3," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1314–1324.

[29] I. Nigam, M. Vatsa, and R. Singh, "Ocular biometrics: A survey of modalities and fusion approaches," *Information Fusion*, vol. 26, pp. 1–35, 5 2015.

[30] Z. Zhao and A. Kumar, "Improving periocular recognition by explicit attention to critical regions in deep neural network," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 12, pp. 2937–2952, 2018.

[31] S. Verma, P. Mittal, M. Vatsa, and R. Singh, "At-a-distance person recognition via combining ocular features," in *International Conference on Image Processing*. IEEE, 9 2016, pp. 3131–3135.

[32] J. M. Smereka, V. N. Boddeti, and B. V. Kumar, "Probabilistic deformation models for challenging periocular image verification," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1875–1890, 2015.

[33] X. Wang, R. Zhang, C. Shen, T. Kong, and L. Li, "Solo: A simple framework for instance segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[34] "Weblink to access new dataset developed during this work," https://web.comp.polyu.edu.hk/csajaykr/metairis.html.

[35] W. J. Ryan, D. L. Woodard, A. T. Duchowski, and S. T. Birchfield, "Adapting starburst for elliptical iris segmentation," in *2008 IEEE Second International Conference on Biometrics: Theory, Applications and Systems*. IEEE, 2008, pp. 1–7.

[36] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Computer Vision and Image Understanding*, vol. 110, no. 2, pp. 281–307, 2008.

[37] C. Lu, S. Xia, M. Shao, and Y. Fu, "Arc-Support Line Segments Revisited: An Efficient High-Quality Ellipse Detection," *IEEE Transactions on Image Processing*, vol. 29, pp. 768–781, 10 2020.

[38] O. M. Kürtüncü and M. Karakaya, "Limbus impact removal for off-angle iris recognition using eye models," in *2015 IEEE 7th International Conference on Biometrics Theory, Applications and Systems, BTAS 2015*. Institute of Electrical and Electronics Engineers Inc., 12 2015.

[39] X. Li, L. Wang, Z. Sun, and T. Tan, "A feature-level solution to off-angle iris recognition," in *Proceedings - 2013 International Conference on Biometrics, ICB 2013*. IEEE Computer Society, 2013.

[40] A. Abhyankar and S. Schuckers, "A novel biorthogonal wavelet network system for off-angle iris recognition," *Pattern Recognition*, vol. 43, no. 3, pp. 987–1007, 3 2010.

[41] W. Chen, X. Chen, J. Zhang, and K. Huang, "Beyond triplet loss: a deep quadruplet network for person re-identification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 403–412.

[42] H. Zhang, C. Wu, Z. Zhang, Y. Zhu, H. Lin, Z. Zhang, Y. Sun, T. He, J. Mueller, R. Manmatha *et al.*, "ResNeSt: Split-attention networks," *arXiv preprint arXiv:2004.08955*, 2020.

[43] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International Journal of Computer Vision*, vol. 115, no. 3, pp. 211–252, 2015.

[44] J. Deng, J. Guo, N. Xue, and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 4690–4699.

[45] G. Quinn, J. Matey, E. Tabassi, and P. Grother, "IREX V: Guidance for Iris Image Collection," 2014.

[46] P. A. Johnson, P. Lopez-Meyer, N. Sazonova, F. Hua, and S. Schuckers, "Quality in face and iris research ensemble (q-fire)," in *2010 Fourth IEEE International Conference on Biometrics: Theory, Applications and Systems (BTAS)*. IEEE, 2010, pp. 1–6.

[47] R. Girshick, "Fast r-cnn," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 1440–1448.

[48] J. Hu, L. Wang, Z. Luo, Y. Wang, and Z. Sun, "A large-scale database for less cooperative iris recognition," in *2021 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 2021, pp. 1–6.

[49] A. Kumar, V. Kanhangad, and D. Zhang, "A new framework for adaptive multimodal biometrics management," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 1, pp. 92–102, 3 2010.

[50] Z. Zhao and A. Kumar, "An accurate iris segmentation framework under relaxed imaging constraints using total variation model," in *Proceedings of the IEEE International Conference on Computer Vision*, 2015, pp. 3828–3836.

[51] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, 2012.

[52] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *International Conference on Learning Representations*. International Conference on Learning Representations, ICLR, 2015.

[53] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[54] "Verieye sdk 12.0." [Online]. Available: http://www.neurotechnology.com/verieye.html