

Boosting Iris Recognition by Margin-Based Loss Functions

Reihan Alinia Lat ¹, Sebelan Danishvar ², Hamed Heravi ² and Morad Danishvar ^{2,*}

¹ Biomedical Engineering Department, Faculty of Electrical and Computer Engineering, University of Tabriz, Tabriz 51368 Iran; reyhana.alinya96@ms.tabrizu.ac.ir

² College of Engineering, Design and Physical Sciences, Brunel University London, Uxbridge UB8 3PH, UK; sebelan.danishvar@brunel.ac.uk (S.D.); hamed.heravi@brunel.ac.uk (H.H.)

* Correspondence: morad.danishvar@brunel.ac.uk

Abstract: In recent years, the topic of contactless biometric identification has gained considerable traction due to the COVID-19 pandemic. One of the most well-known identification technologies is iris recognition. Determining the classification threshold for large datasets of iris images remains challenging. To solve this issue, it is essential to extract more discriminatory features from iris images. Choosing the appropriate loss function to enhance discrimination power is one of the most significant factors in deep learning networks. This paper proposes a novel iris identification framework that integrates the light-weight MobileNet architecture with customized ArcFace and Triplet loss functions. By combining two loss functions, it is possible to improve the compactness within a class and the discrepancies between classes. To reduce the amount of preprocessing, the normalization step is omitted and segmented iris images are used directly. In contrast to the original SoftMax loss, the EER for the combined loss from ArcFace and Triplet is decreased from 1.11% to 0.45%, and the TPR is increased from 99.77% to 100%. In CASIA-Iris-Thousand, EER decreased from 4.8% to 1.87%, while TPR improved from 97.42% to 99.66%. Experiments have demonstrated that the proposed approach with customized loss using ArcFace and Triplet can significantly improve state-of-the-art and achieve outstanding results.



Citation: Alinia Lat, R.; Danishvar, S.; Heravi, H.; Danishvar, M. Boosting Iris Recognition by Margin-Based Loss Functions. *Algorithms* **2022**, *15*, 118. <https://doi.org/10.3390/a15040118>

Academic Editor: Frank Werner

Received: 19 January 2022

Accepted: 26 March 2022

Published: 29 March 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: biometrics; machine learning; convolutional neural networks; deep learning; iris recognition; margin-based loss functions

1. Introduction

During the pandemic of COVID-19, non-contact or contactless biometric-based identification methods led the industry in 2020. There is a growing interest in contactless biometric recognition, whether it is using the face [1] and ears [2] or using the iris [3] and voice [4] as the most functional biometric techniques. Contact-based biometrics, such as fingerprints [5], where sensors require that the finger be tightly held, are secure for a person's data, but they are not as secure for their health. Wearing a facial mask is the new norm among people in this pandemic, and applications of face recognition technology [6,7] have encountered severe problems [8]. A recent work [9] focused on the system for recognizing faces with and without masks. Due to these challenges, iris recognition technology may become more widely adopted since it is contactless and not impacted by masks.

Based on Daugman [10], the four fundamental stages of the iris recognition system are segmentation, normalization, feature extraction, and feature matching. Although accurate iris segmentation is necessary, extracting discriminative features of iris images is the most challenging aspect of these systems. Convolutional neural networks (CNNs) can learn and extract features from highly complex datasets, such as iris images. A variety of solutions can improve CNNs, such as batch normalization, data augmentation, regularization, dropout, and different activation functions. Considering reducing the computational parameters, researchers designed improved structures of the pre-trained network.

Researchers have paid great attention to optimizing the loss functions to obtain distinctive feature distributions. The SoftMax loss function is one popular loss function for recognition tasks. In terms of inter-class separation, SoftMax loss performs well but is ineffective in minimizing intra-class compactness. For iris recognition systems, where only a few samples are available for each class, the generation of discriminating features is crucial to achieving high classification accuracy. There are differences in the way that SoftMax-based loss functions apply margin penalties. The Angular Softmax Loss (SphereFace) [11] applies a multiplicative angular margin penalty. The Large Margin Cosine Loss (CosFace) [12] adds a cosine margin penalty to the logit directly. Additive Angular Margin Loss (ArcFace) [13] places the additive cosine margin within a geometric angular space.

Using the MobileNet architecture with the combination of the ArcFace and Triplet loss functions in this paper, we train an iris recognition network taking advantage of both loss functions. We also compare the performance of this improved loss function to three others, consisting of original SoftMax loss, CosFace, and ArcFace. Notably, we try to optimize the network with segmented iris images without normalization. The testing datasets, especially CASIA-Iris-Thousand, have many classes with reflection or eyeglasses that will aid in clarifying the effect of challenging irises on loss functions by analyzing these two iris datasets. This work follows the main contributions:

- To extract discriminative iris features, our study focuses on implementing state-of-the-art loss functions, especially a combination of ArcFace and Triplet loss with a light-weight MobileNet architecture for the training approach;
- The performance of the proposed system is tested on two public datasets that interfere with the various noises: CASIA-Iris-Thousand [14] and IITD [15] datasets.

The remaining sections in this paper are arranged as follows: In Section 2, we review a couple of related works with more focus on deep learning-based methods. Section 3 presents an overview of the proposed deep learning approach with described combined loss function. Section 4 demonstrates the proposed system's experimental results. The last section includes conclusions and future directions.

2. Related Work

Recent studies by Adjabi et al. [1] reviewed 2D and 3D face recognition from traditional to deep-learning models. For the recognition of single samples of faces, Adjabi et al. [16] used multiblock color-binarized statistical images. A machine learning technique called Deep Unsupervised Active Learning (DUAL) was proposed by Khaldi et al. [2], to determine the feasibility of active learning that could be used in ear recognition and biometrics in general. Researchers have exploited neural networks or deep learning approaches not only to identification, but also with biometric trait classification. Classification is often used in large-scale datasets to reduce the search space, busting the performance. According to Militello et al. [17], the first study has compared the most common CNN architectures by dividing fingerprints into four, five, and eight categories. The purpose of this work is to demonstrate the potential of CNN architectures for fingerprint classification, by reducing the amount of data needed to be processed during identification and search.

Most recent iris recognition work focuses on segmenting the iris and extracting iris features. Noisy iris images, such as those occluded by eyelids or eyelashes, specular reflections, and off-angles, are behind the main problem in segmenting and localizing them [18]. Denoising images using first and second generation wavelets [19,20] or fast matching algorithms [21] are some ways of preprocessing images.

Processing speed and recognition time are two critical parameters of any real-time biometric system. The normalization process consumes a significant amount of time. In work [22], authors claimed that by skipping the classical normalization step, they could improve iris recognition by saving time comparable to that spent on it. Because of the normalization errors, the normalized samples showed inconsistencies, leading to unstable recognition. Considering the capability of deep learning to model highly complex distributions and nonlinear data changes, Lozej et al. in [23] tried to investigate how crucial

the use of traditional segmentation techniques is in a deep learning setting. Chen et al. [24] presented a novel non-normalized preprocessing method based on dynamic path search for iris segmentation. Using two public iris datasets, which are the same as in this work, the authors proved that iris segmentation images without normalization are a better choice for evaluating iris recognition techniques based on deep learning techniques.

For feature extraction, two categories of previous approaches are handcrafted approaches and learning-based approaches. Daugman [10] proposed the first study in handcrafted approaches in 1993, which used an integral-differential operator to segment the iris, followed by a Gabor filter to extract features. Wildes used the Laplacian pyramid and Fisher's linear analysis [25]. Krichen et al. [26] used Gabor wavelet packets. Miyazawa et al. [27] used 2D Discrete Fourier Transforms (DFTs).

In the field of learning-based methods, CNNs are effective techniques. Liu et al. [28] provided the first use of CNN for iris recognition in 2016. They used a simple CNN with a multiphase filter bank. According to Minaee et al. [29], using more layers is ineffective for recognizing more distinctive features. Their method used the VGG network architecture. Gangwar and Joshi [30] proposed DeepIrisNet in two categories. The first case used standard convolution layers, and the second case incorporated inception layers. The model can recognize images from cross-sensors with high accuracy. Nguyen et al. [31] extracted features using some popular pre-trained networks and classified them using Support Vector Machine (SVM). DensNet architecture achieved the best result in their work. Al-Waisy et al. [32] developed a deep learning system called IrisCONvNet, which trained a CNN with a SoftMax classifier. Using the ranking-level fusion method, they integrated the classification results. Zhang et al. [33] fused iris and periocular features with a weighted concatenation. They claimed that maxout units do better than ReLUs in their study. Using capsule networks, Zhao et al. [34] developed a deep learning algorithm for iris recognition. These networks can learn part-whole deep features relationships and lead to high classification accuracy.

The SoftMax loss is used to supervise the training of the reviewed iris recognition systems. Zhao and Kumar [35] proposed a network called UniNet. They designed an Extended Triplet Loss (ETL) function inspired by work in [36] and fed the samples to the model in triplet sets. Ahmad and Fuller [37] trained a network using ResNet with SoftMax classifier, thus optimizing the model using the triple loss function. Their vital target was to evaluate the network performance by eliminating the normalization stage. On constrained datasets, their network performed well. The primary demand of the triplet loss function is triplet mining or selecting semi-hard samples. In addition, comparisons among triplet samples within mini-batch are local. Inspired from the approach proposed in [38], Chen et al. in [39] introduced a novel loss function called Tight Center (T-Center). The optimized center loss function solved the problem of insufficient discrimination caused by SoftMax loss, but there is an incompatibility between Softmax loss and center-based loss functions. The SoftMax loss has an intrinsic angular distribution, while the center loss applies the Euclidean margin to penalize the distance between the deep features and their corresponding.

To angular discriminative deep feature learning, some loss functions such as SphereFace [11], CosFace [12], and Arcface [13] have been proposed recently. Hsu et al. [40] analyzed the performance of these loss functions on several face datasets containing challenge factors. Our paper uses a similar algorithm as in [40] for iris datasets.

3. Proposed Approach

Figure 1 provides an overview of the proposed framework. First, a preprocessing method removes the specular reflection from the input images. Then, a deep learning algorithm is applied to detect iris location and segment it into iris/noniris. The segmented irises are presented without being normalized to a feature extraction framework. In Section 3.1, we provide more details on image processing. During the training phase, a MobileNet architecture is implemented using a combined loss function with Triplet and

ArcFace. Through the testing phase, embedding is the model output before the classifier header, which is regarded as the image feature map. The embedding vector is used for evaluation, by calculating cosine similarities between each feature. First, we pick an image most similar to others in a single class as registered base embedding, which is the yellow dot in Figure 1. Then, we test all of the left-over images if classified correctly, which in Figure 1 are the red and blue dots representing the positive and negative samples. The task we present in this paper focuses on both angular measurement between two features and distance metric, so we prefer cosine similarities to compute extracted embedding values.

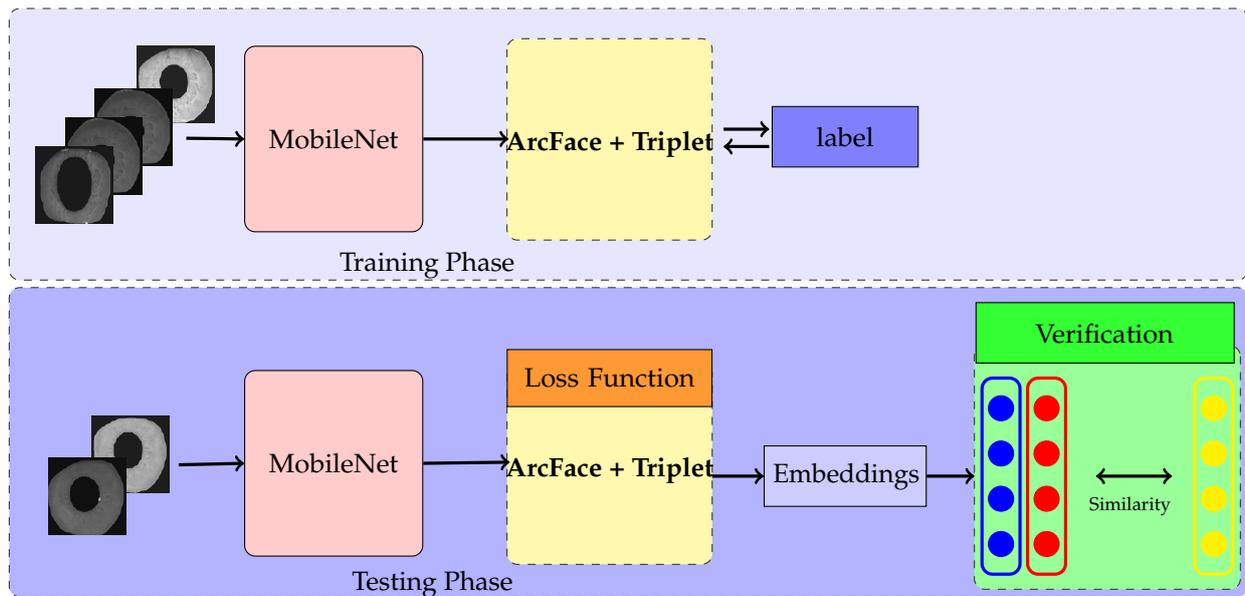


Figure 1. Overview of the proposed framework.

3.1. Preprocessing and Segmentation

With only a slight modification, we use a deep learning segmentation method that has been proposed in [41]. It trained a Mask R-CNN framework with segmentation ground truth datasets prepared by Hofbauer et al. [42]. For the images with reflections, the proposed method does not segment accurately. We solve this issue by adding a preprocessing stage, which uses an expansion operator and a fast-matching algorithm [21] to remove the existing reflections. Figure 2 shows examples of the iris bounding box detection and mask segmentation implementing this preprocessing. It can be observed that the removal of reflection points leads to accurate segmentation. In this work, we eliminate the normalization step common in most recent works after segmentation. With data augmentation, it is possible to increase the variation in the training datasets and help overcome over-fitting. In our augmentation, we randomly randomly half of all iris images horizontally and vertically. A vertical flip is equivalent to rotating an image by 180 degrees and then performing a horizontal flip.

3.2. Combined Loss of ArcFace and Triplet

The selected loss function to combine, ArcFace, is based on changes in the SoftMax with a triplet loss function. Softmax loss can be expressed as follows:

$$L_{Softmax} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_{y_j}}}, \tag{1}$$

where $x_i \in \mathbb{R}^d$ represents the deep feature vector or output from the last CNN layer to the fully connected layer of the i -th sample belonging to the y_i -th class. $W_j \in \mathbb{R}^d$ is the j -th column of the weight matrix $W_j \in \mathbb{R}^{d \times n}$ and $b_j \in \mathbb{R}^n$ is the bias term. n is the class number,

and N is the batch size. The SoftMax loss function cannot optimize deep feature vectors to achieve better intra-class similarity and the scattered distribution of samples can cause problems for clustering.

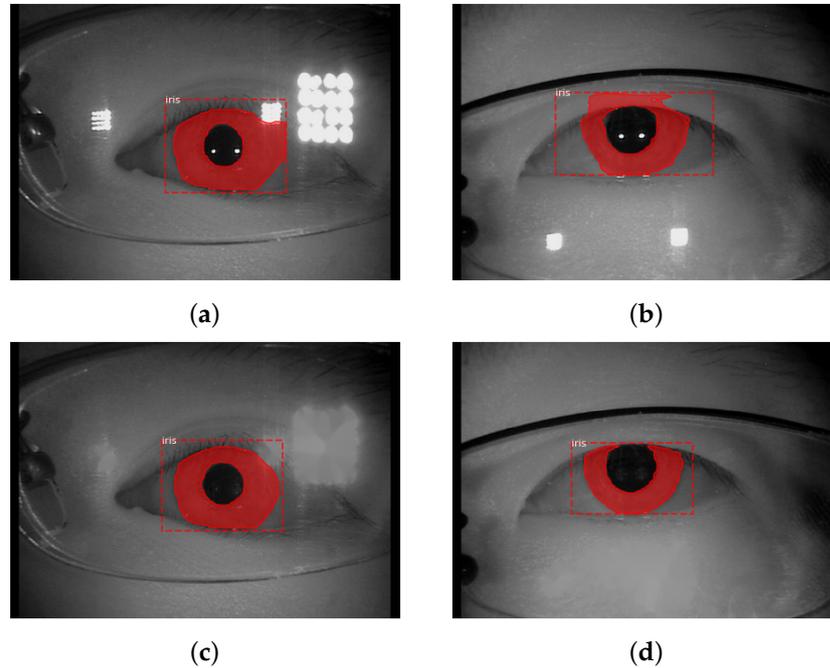


Figure 2. Sample results for the iris bounding box detection and mask segmentation from the method that was proposed in [41] (a,c) and with the proposed method to remove reflections (b,d).

Large Margin Cosine Loss [12] followed [11] initially set the bias value to zero. In $W_j^T x_i = \|W_j\| \|x_i\| \cos \theta_j$, where θ_j is the angle between feature vectors x_i and the ground truth weight W_j , the authors fixed $\|W_j\| = 1$ and $\|x_i\|$ by l_2 normalization and re-scaled $\|x_i\|$ to s . Normalizing the weights and features yields predictions based on only the angle between them. Each embedding has a radius s and is distributed on the hypersphere. They named it “CosFace”.

$$L_{co} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i})-m)}}{e^{s(\cos(\theta_{y_i})-m)} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}} \tag{2}$$

By taking advantage of the fact that the embedding features are distributed around each center on the hypersphere, the work in [13] added an additive angular margin penalty between x_i and W_{y_i} to enhance the intra-class compactness and inter-class dispersion simultaneously. The authors named their method “ArcFace”.

$$L_{aa} = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i}+m))}}{e^{s(\cos(\theta_{y_i}+m))} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}} \tag{3}$$

Compared with CosFace and similar loss functions, ArcFace has a better geometric property because of the exact alignment of angular margin with geodesic distance. Table 1 provides us with geometric comparisons differences of different decision margins. In binary classification, ArcFace has a fixed linear margin across the decision boundaries, whereas CosFace creates only a nonlinear margin. ArcFace loss functions can classify more accurately, since $\theta \in [0, \pi - m]$, $\cos(\theta + m) < \cos(\theta)$. In addition, because of $\sin(\theta)$ in $\cos(\theta + m) = \cos \theta \cos m - \sin \theta \sin m$, the new margin is more dynamic.

Table 1. Comparing binary classification decision boundaries.

Loss Functions	Decision Boundaries
Softmax	$(W_1 - W_2)x + b_1 - b_2 = 0$
CosFace	$s(\cos \theta_1 - m - \cos \theta_2) = 0$
ArcFace	$s(\cos(\theta_1 + m) - \cos \theta_2) = 0$

FaceNet [36] first uses triplet loss to learn deep representations of faces. In the presence of triplet loss, the positive pairs tend to have smaller distances than the negative pairs. In terms of mathematics, the loss can be described as follows using the Euclidean distance function:

$$L = \frac{1}{N} \sum_i^N \max\left(\|f(x_i^a) - f(x_i^p)\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha, 0\right) \quad (4)$$

$$\forall (f(x_i^a), f(x_i^p), f(x_i^n)) \in \tau$$

where $(f(x_i^a), f(x_i^p))$ and $f(x_i^n)$ represent the feature vectors of the positive pairs (x_i^a, x_i^p) and the negative pairs (x_i^a, x_i^n) and α is the distance margin. Training samples with too few samples limit the effectiveness of feature embedding. One approach [43] proposes hard sample mining to help diminish this problem by utilizing the effective information in each batch, which is necessary for maximizing the performance of feature embedding.

ArcFace enlarges the margin between different classes (the so-called inter-class margin), while Triplet Loss also narrows the margin for the same class (the so-called intra-class margin). Moreover, the semi-hard triplet focuses only on the more complex samples. Combining losses allow deep for feature learning with high optimization flexibility and more accurate convergence targeting.

3.3. Network Architecture

Choosing a network architecture is still a challenge and depends on its specific application. One of the most significant light-weight networks is MobileNet [44]. Google designed MobileNet and optimized it to implement on mobile devices. The key feature of this structure is the low number of learning parameters with high classification accuracy. Blocks that build MobileNet take advantage of depthwise convolution. By applying only one convolutional filter for each input channel, depthwise convolution performs light-weight filtering. Also, MobileNet blocks have a 1×1 convolution layer called a pointwise convolution. This component builds new features through linear combinations of input channels. These layers form a depthwise separable convolutional block. This block is illustrated in Figure 3. Each convolution layer is followed by batch normalization, which leads to faster convergence for deep learning models. MobileNet also benefits from ReLU6 activation function. ReLU6 activation is identical to the original ReLU, except that it is limited to six activations. In this paper, we use MobileNetV1. It begins with a regular 3×3 convolution and follows by 13 times the above building block.

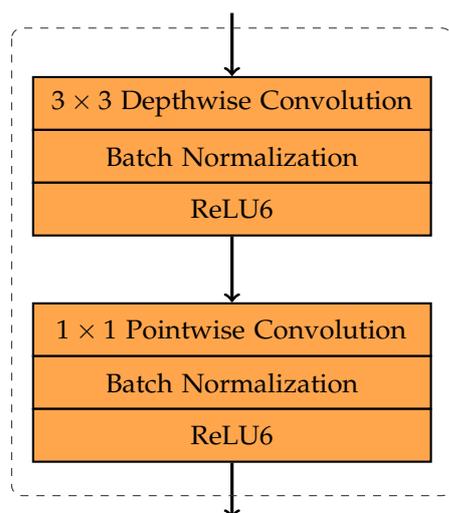


Figure 3. Depthwise Separable Convolutional block used in MobileNet.

4. Experiments and Results

In this section, we evaluate the performance of the proposed method on two selected datasets (CASIA-Iris-Thousand, IITD) and plot the Receiver Operating Characteristic (ROC) curves to calculate Area Under the ROC Curve (AUC) and Equal Error Rate (EER).

4.1. Datasets

- **CASIA-Iris-Thousand.** The CASIA-Iris-Thousand [14] includes 20,000 images collected from 1000 subjects using the IKEMB-100 camera. Eyeglasses and specular reflections cause the main challenges in this dataset. CASIA-Iris-Thousand is the first available iris dataset of 1000 people. This large number of classes causes a challenge for analyzing unique iris characteristics and requires new classification methods. The images in this collection were recorded in 640×480 resolution in jpg format. The resolution of the segmented image is 200×200 . we used the 10 left iris images from all subjects for the training set. We also select five random right iris images from each category to test the network performance;
- **IITD.** The IITD [15] contains 2240 images of the iris collected from 224 students and staff in Delhi, India. The database consists of 176 males and 48 females in the age range 14–55 years. There are different (iris) sizes and distributions of colors in this dataset. The images were recorded at 320×240 resolution and BMP format. The segmented images have a size of 200×200 . We use the five images from the left iris images for network training and the five right iris images for testing.

4.2. Experimental Settings

We implement the deep CNN model in Python, using the TensorFlow back-end, with Keras library, a stochastic gradient descent (SGD) optimizer, which is used to optimize the model's hyper-parameters. Momentum and weight decay are set at 0.9 and $5 \times e - 4$ for this optimizer. The initial learning rate is 0.009, divided by 10 on the 18th epoch of the 35 total epochs. The batch size is set to 64. According to [12], we would set the angular margin m of ArcFace at 0.5 and the feature scale s to 64.

4.3. Evaluation

To generate a stable evaluation result and reduce the calculating complexity, we first extract the embedding value from all images, and then sum up and normalize all embedding vectors in the same class as the registered base embedding vector. We use this vector as an embedding vector for this class and measure distance by calculating cosine distances between all embedding vectors and all registered base embedding vectors. Evaluation results consist of two parts: distances among all images and between all vectors

registered, a label that shows whether each score is positively or negatively classified. The final evaluating results contain two parts score for distances between all images and all registered vectors, and a label showing whether each score should be classified positive or negative.

We compare the performance of selected loss functions using ROC curves and the EER to quantify the verification process. In addition to ROC curves for each dataset, we also select TPR at FPR = 0.1% and calculate the AUC to provide other feedback. Figures 4 and 5 show the ROC curves obtained on IITD and CASIA-Iris-Thousand datasets, respectively. Tables 2 and 3 show the results based on TPR, EER, and AUC. The proposed CNN model that uses ArcFace + Triplet loss is more effective for recognizing iris patterns. To make our conclusion clearer, we plot the ROC curve for combined loss separately in Figures 6 and 7. Implementing ArcFace + Triplet loss achieves the lowest EER compared with the Arcface, CosFace, and SoftMax loss. The performance drops to the lower amount of AUC and TPR when evaluating the algorithm with Arcface and CosFace. Additionally, we plot the ROC of the SoftMax loss function to compare two margin-based loss functions with the SoftMax loss function. CASIA-Iris-Thousand is a more complex dataset than IITD because it contains images with eyeglasses and specular reflections. Therefore, the EER for CASIA-Iris-Thousand is higher than EER for IITD.

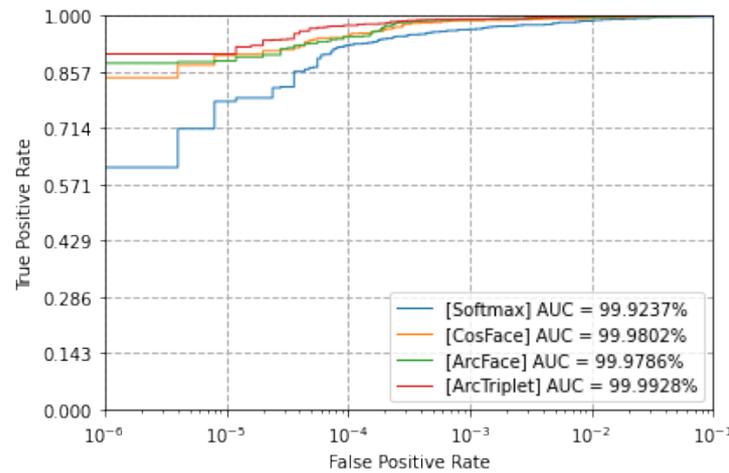


Figure 4. ROC curves of ArcFace + Triplet, ArcFace, CosFace and Sotmax on IITD.

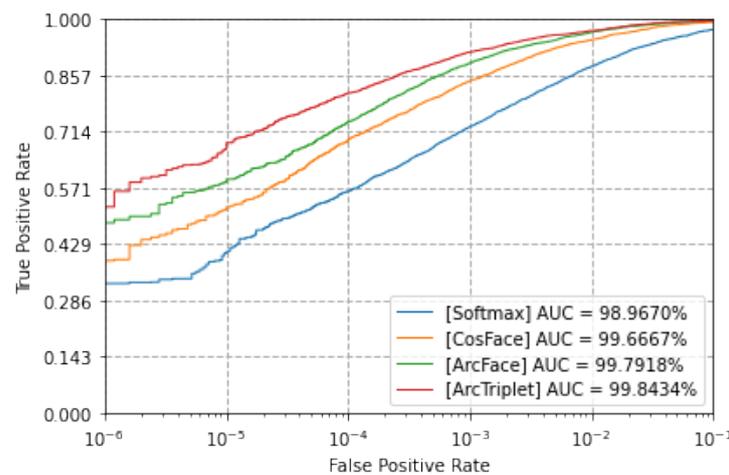


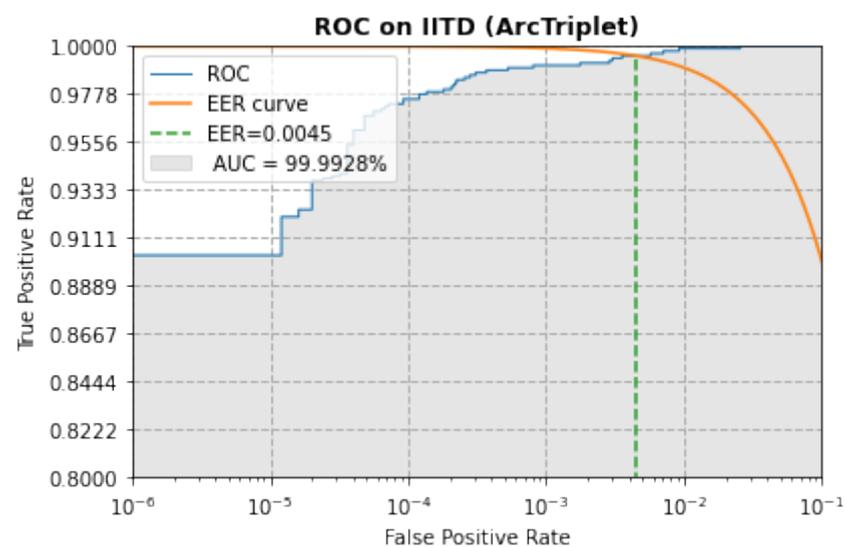
Figure 5. ROC curves of ArcFace + Triplet, ArcFace, CosFace and Sotmax on CASIA-Iris-Thousand.

Table 2. Testing results of different implementation on IIT Delhi.

Implementation	EER (%)	TPR (%)	AUC (%)
MobileNet + Softmax	1.11	99.77	99.92
MobileNet + Cosface	0.78	100	99.98
MobileNet + ArcFace	0.45	100	99.98
MobileNet + (ArcFace + Triplet)	0.45	100	99.99

Table 3. Testing results of different implementation on CASIA-Iris-Thousand.

Implementation	EER (%)	TPR (%)	AUC (%)
MobileNet + Softmax	4.80	97.42	98.97
MobileNet + Cosface	2.64	99.22	99.67
MobileNet + ArcFace	1.89	99.58	99.80
MobileNet + (ArcFace + Triplet)	1.87	99.66	99.84

**Figure 6.** ROC curve of ArcFace + Triplet loss on IITD.

On the evaluated results, we also plot the histograms of positive and negative distributions. This histogram visually compares how the proposed method with the margin-based loss functions can distinguish iris feature vectors. We can see a clear decision margin distinguished in Figure 8, where the positive similarities for ArcFace are more gathered to the right of the figure, and the overlapping area is smaller. Though Figure 9 is not as clear as Figure 8, we also see a more narrow distribution for CosFace and ArcFace positive similarities. As illustrated in Figures 8 and 9, ArcFace + Triplet is the best among all of them because there is a smaller overlap area between the red and blue distributions, and its positive distribution is more tended to the right than the other one. SoftMax loss is worst at separation. The margin between red and blue is smaller, and the overlapped area is broader. However, its negative red distribution is narrower and more concentrated, which means a smaller variance value. CosFace works similar to ArcFace, only adding a different margin between negative and positive. Thus, their results are identical, as only ArcFace + Triplet has a smaller overlapped area, making it a better separator.

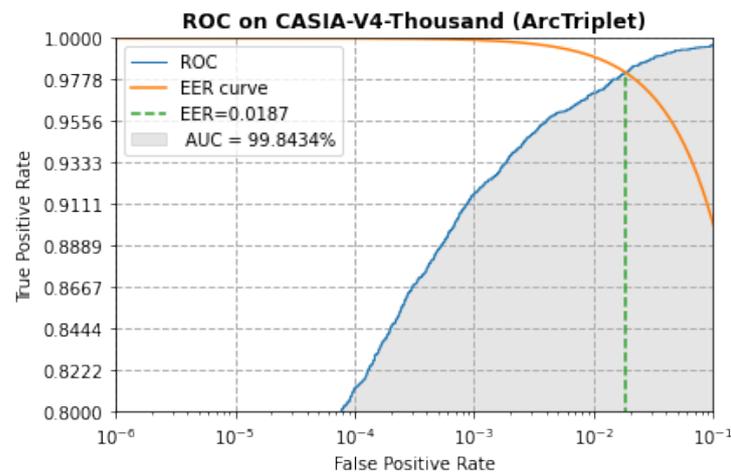


Figure 7. ROC curve of ArcFace + Triplet loss on CASIA-Iris-Thousand.

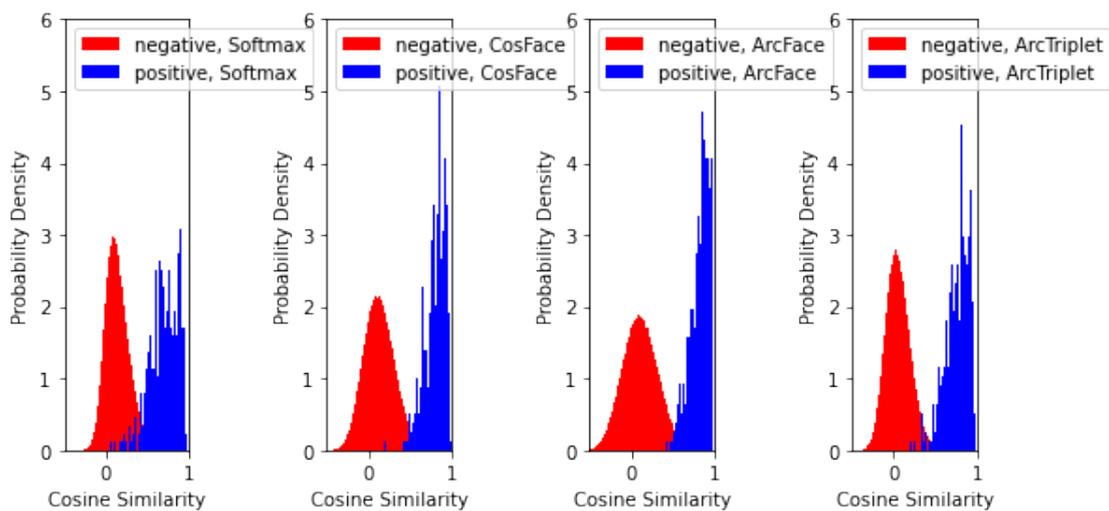


Figure 8. Histograms of positive and negative similarity distribution on the evaluating result for CASIA-Iris-Thousand.

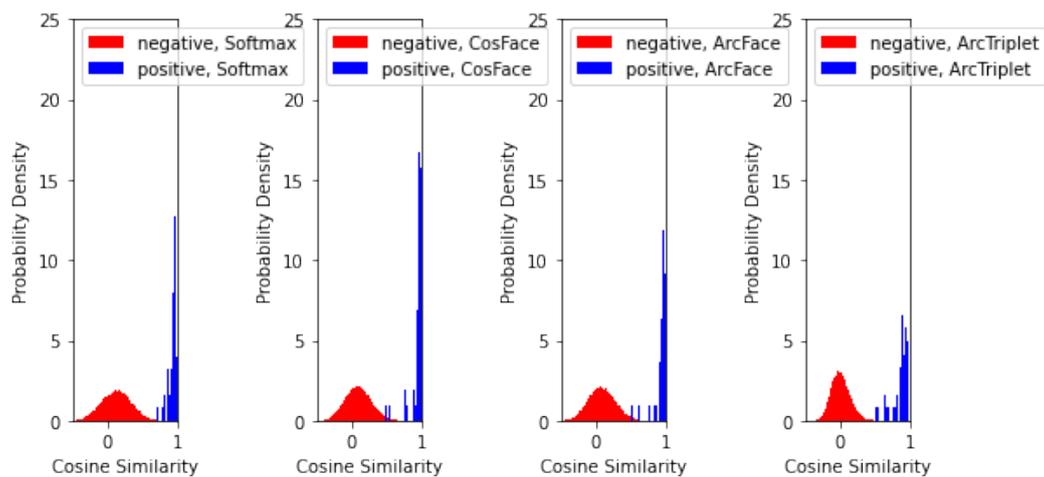


Figure 9. Histograms of positive and negative similarity distribution on the evaluating result for IITD.

We compare our model with three deep learning models, including DeepIrisNet [30], UniNet [35], CapsuleNet [34] and T-Center [39]. Table 4 shows that the proposed method with ArcFace + Triplet achieves the lowest value of EER for both datasets compared with

prior state-of-the-art works. It has greatly improved about 2.57% and 7% for IITD and CASIA-Iris-Thousand compared with T-Center. An important point to note is that the experiments we compared used normalized images, while our model does not involve normalization.

Table 4. Comparison of the performance on two public datasets.

	IITD			CASIA-Iris-Thousand		
	TPR (%)	EER (%)	AUC (%)	TPR (%)	EER (%)	AUC (%)
DeepIrisNet [30]	95.03	1.17	99.02	86.86	4.01	97.93
UniNet [35]	97.35	0.89	99.20	90.61	2.79	98.67
CapsuleNet [34]	97.39	0.76	99.27	91.04	3.05	98.60
T-Center [39]	97.43	0.74	99.30	92.54	2.36	98.75
Proposed method (MobileNet + (ArcFace + Triplet))	100	0.45	99.99	99.66	1.87	99.84

5. Conclusions and Future Work

Through the present article, the researchers proposed a deep learning framework for iris recognition based on MobileNet architecture by using a combined loss function with ArcFace and Triplet. ArcFace has a larger margin on semi-hard instances and keeps a small margin on easy ones; and increases the margins between classes, Triplet Loss narrows margins within a single class. This somehow makes ArcFace work better on iris recognition tasks. As expected, the robust discriminative proposed CNN architecture using the combined loss function achieved significant performance with the EER = 0.45% for IITD and EER = 1.87% for CASIA-Iris-Thousand without normalization step. The obtained results, especially for CASIA-Iris-Thousand, have proved that the combined loss function can effectively be applied to large-scale iris identities. According to the above finding, it can be realized that the results of the Arcface studied in the present investigation lead to better use of existing state-of-the-art loss functions for iris recognition systems.

Various limitations, such as noisy data, differences and similarities between classes, have led to the development of multimodal biometrics, which utilizes multiple sources of information to authenticate the person. Using a fusion of matching and decision score level method, the proposed approach could be applied in a multimodal context. As shown in Conti et al. [45], fusing iris and retina or periocular modalities represents a robust combination. For future developments of the proposed approach, we will focus on multimodal biometrics. In future research, we will apply the proposed method to various datasets collected in an unconstrained environment with more recent deep learning architecture. In addition, we will use other loss functions independent of determining the margins and scale.

Author Contributions: Conceptualization, R.A.L.; Data curation, S.D.; Formal analysis, R.A.L.; Methodology, H.H. and R.A.L.; Resources, H.H. and R.A.L.; Validation, S.D., R.A.L. and H.H.; Visualization, Morad Danishvar; Writing—original draft, R.A.L. and S.D.; Writing—review & editing, S.D., Morad Danishvar and R.A.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The analysed datasets are publicly available. Related references are reported in the References section.

Acknowledgments: The authors would like to thank Guowei Wang for providing the implementation of Keras_insightface, which is available on Github, accessed on April 2021 (https://github.com/leondgarse/Keras_insightface/ access on 25 April 2021).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Taleb-Ahmed, A. Past, present, and future of face recognition: A review. *Electronics* **2020**, *9*, 1188. [CrossRef]
2. Khaldi, Y.; Benzaoui, A.; Ouahabi, A.; Jacques, S.; Taleb-Ahmed, A. Ear recognition based on deep unsupervised active learning. *IEEE Sens. J.* **2021**, *21*, 20704–20713. [CrossRef]
3. Hu, J.; Wang, L.; Luo, Z.; Wang, Y.; Sun, Z. A Large-scale Database for Less Cooperative Iris Recognition. In Proceedings of the 2021 IEEE International Joint Conference on Biometrics (IJCB), Shenzhen, China, 4–7 August 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–6.
4. Gomez-Barrero, M.; Drozdowski, P.; Rathgeb, C.; Patino, J.; Todisco, M.; Nautsch, A.; Damer, N.; Priesnitz, J.; Evans, N.; Busch, C. Biometrics in the era of COVID-19: Challenges and opportunities. *arXiv* **2021**, arXiv:2102.09258.
5. Jain, A.K.; Arora, S.S.; Cao, K.; Best-Rowden, L.; Bhatnagar, A. Fingerprint recognition of young children. *IEEE Trans. Inf. Forensics Secur.* **2016**, *12*, 1501–1514. [CrossRef]
6. Bonnen, K.; Klare, B.F.; Jain, A.K. Component-based representation in automated face recognition. *IEEE Trans. Inf. Forensics Secur.* **2012**, *8*, 239–253. [CrossRef]
7. Mundial, I.Q.; Hassan, M.S.U.; Tiwana, M.I.; Qureshi, W.S.; Alanazi, E. Towards facial recognition problem in COVID-19 pandemic. In Proceedings of the 2020 4th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM), Medan, Indonesia, 3–4 September 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 210–214.
8. Rahman, A.; Hossain, M.S.; Alrajeh, N.A.; Alsolami, F. Adversarial Examples—Security threats to COVID-19 deep learning systems in medical IoT devices. *IEEE Internet Things J.* **2020**, *8*, 9603–9610. [CrossRef]
9. Talahua, J.S.; Buele, J.; Calvopiña, P.; Varela-Aldás, J. Facial recognition system for people with and without face mask in times of the covid-19 pandemic. *Sustainability* **2021**, *13*, 6900. [CrossRef]
10. Daugman, J.G. High confidence visual recognition of persons by a test of statistical independence. *IEEE Trans. Pattern Anal. Mach. Intell.* **1993**, *15*, 1148–1161. [CrossRef]
11. Liu, W.; Wen, Y.; Yu, Z.; Li, M.; Raj, B.; Song, L. Sphreface: Deep hypersphere embedding for face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 212–220.
12. Wang, H.; Wang, Y.; Zhou, Z.; Ji, X.; Gong, D.; Zhou, J.; Li, Z.; Liu, W. Cosface: Large margin cosine loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 5265–5274.
13. Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4690–4699.
14. Chinese Academy of Sciences Institute of Automation. Casia Iris Image Database. 2017. Available online: <http://www.cbsr.ia.ac.cn/english/IrisDatabase.asp> (accessed on 22 January 2021).
15. Kumar, A.; Passi, A. Comparison and combination of iris matchers for reliable personal authentication. *Pattern Recognit.* **2010**, *43*, 1016–1026. Available online: https://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Iris.htm (accessed on 25 April 2021). [CrossRef]
16. Adjabi, I.; Ouahabi, A.; Benzaoui, A.; Jacques, S. Multi-block color-binarized statistical images for single-sample face recognition. *Sensors* **2021**, *21*, 728. [CrossRef]
17. Militello, C.; Rundo, L.; Vitabile, S.; Conti, V. Fingerprint classification based on deep learning approaches: Experimental findings and comparisons. *Symmetry* **2021**, *13*, 750. [CrossRef]
18. Wang, C.; Wang, Y.; Zhang, K.; Muhammad, J.; Lu, T.; Zhang, Q.; Tian, Q.; He, Z.; Sun, Z.; Zhang, Y.; et al. NIR iris challenge evaluation in non-cooperative environments: Segmentation and localization. In Proceedings of the 2021 IEEE International Joint Conference on Biometrics (IJCB), Shenzhen, China, 4–7 August 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 1–10.
19. Ouahabi, A. *Signal and Image Multiresolution Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
20. Ouahabi, A. A review of wavelet denoising in medical imaging. In Proceedings of the 2013 8th International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, Algeria, 12–15 May 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 19–26.
21. Telea, A. An image inpainting technique based on the fast marching method. *J. Graph. Tools* **2004**, *9*, 23–34. [CrossRef]
22. Birgale, L.; Kokare, M. Iris Recognition without Iris Normalization. *J. Comput. Sci.* **2010**, *6*, 1042–1047. [CrossRef]
23. Lozej, J.; Štepec, D.; Štruc, V.; Peer, P. Influence of segmentation on deep iris recognition performance. In Proceedings of the 2019 7th International Workshop on Biometrics and Forensics (IWBF), Cancun, Mexico, 2–3 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
24. Chen, Y.; Wu, C.; Wang, Y. Whether normalized or not? Towards more robust iris recognition using dynamic programming. *Image Vis. Comput.* **2021**, *107*, 104112.
25. Wildes, R.P. Iris recognition: An emerging biometric technology. *Proc. IEEE* **1997**, *85*, 1348–1363. [CrossRef]
26. Krichen, E.; Mellakh, M.A.; Garcia-Salicetti, S.; Dorizzi, B. Iris identification using wavelet packets. In Proceedings of the 17th International Conference on Pattern Recognition, 2004, ICPR 2004, Cambridge, UK, 26 August 2004; IEEE: Piscataway, NJ, USA, 2004; Volume 4, pp. 335–338.
27. Miyazawa, K.; Ito, K.; Aoki, T.; Kobayashi, K.; Nakajima, H. An effective approach for iris recognition using phase-based image matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **2008**, *30*, 1741–1756. [CrossRef]

28. Liu, N.; Zhang, M.; Li, H.; Sun, Z.; Tan, T. DeepIris: Learning pairwise filter bank for heterogeneous iris verification. *Pattern Recognit. Lett.* **2016**, *82*, 154–161. [[CrossRef](#)]
29. Minaee, S.; Abdolrashidiy, A.; Wang, Y. An experimental study of deep convolutional features for iris recognition. In Proceedings of the 2016 IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA, 3 December 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 1–6.
30. Gangwar, A.; Joshi, A. DeepIrisNet: Deep iris representation with applications in iris recognition and cross-sensor iris recognition. In Proceedings of the 2016 IEEE international conference on image processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016; IEEE: Piscataway, NJ, USA, 2016; pp. 2301–2305.
31. Nguyen, K.; Fookes, C.; Ross, A.; Sridharan, S. Iris recognition with off-the-shelf CNN features: A deep learning perspective. *IEEE Access* **2017**, *6*, 18848–18855. [[CrossRef](#)]
32. Al-Waisy, A.S.; Qahwaji, R.; Ipson, S.; Al-Fahdawi, S.; Nagem, T.A. A multi-biometric iris recognition system based on a deep learning approach. *Pattern Anal. Appl.* **2018**, *21*, 783–802. [[CrossRef](#)]
33. Zhang, Q.; Li, H.; Sun, Z.; Tan, T. Deep feature fusion for iris and periocular biometrics on mobile devices. *IEEE Trans. Inf. Forensics Secur.* **2018**, *13*, 2897–2912. [[CrossRef](#)]
34. Zhao, T.; Liu, Y.; Huo, G.; Zhu, X. A deep learning iris recognition method based on capsule network architecture. *IEEE Access* **2019**, *7*, 49691–49701. [[CrossRef](#)]
35. Zhao, Z.; Kumar, A. A deep learning based unified framework to detect, segment and recognize irises using spatially corresponding features. *Pattern Recognit.* **2019**, *93*, 546–557. [[CrossRef](#)]
36. Schroff, F.; Kalenichenko, D.; Philbin, J. Facenet: A unified embedding for face recognition and clustering. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 815–823.
37. Ahmad, S.; Fuller, B. Thirdeye: Triplet based iris recognition without normalization. In Proceedings of the 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS), Tampa, FL, USA, 23–26 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–9.
38. Wen, Y.; Zhang, K.; Li, Z.; Qiao, Y. A discriminative feature learning approach for deep face recognition. In Proceedings of the European Conference on Computer Vision, Amsterdam, The Netherlands, 11–14 October 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 499–515.
39. Chen, Y.; Wu, C.; Wang, Y. T-center: A novel feature extraction approach towards large-scale iris recognition. *IEEE Access* **2020**, *8*, 32365–32375. [[CrossRef](#)]
40. Hsu, G.S.J.; Wu, H.Y.; Yap, M.H. A comprehensive study on loss functions for cross-factor face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 826–827.
41. Ahmad, S.; Fuller, B. Unconstrained iris segmentation using convolutional neural networks. In Proceedings of the Asian Conference on Computer Vision, Kyoto, Japan, 2–4 December 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 450–466.
42. Hofbauer, H.; Alonso-Fernandez, F.; Wild, P.; Bigun, J.; Uhl, A. A ground truth for iris segmentation. In Proceedings of the 2014 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 24–28 August 2014; IEEE: Piscataway, NJ, USA, 2014; pp. 527–532.
43. Smirnov, E.; Melnikov, A.; Novoselov, S.; Luckyanets, E.; Lavrentyeva, G. Doppelganger mining for face representation learning. In Proceedings of the IEEE International Conference on Computer Vision Workshops, Venice, Italy, 27–29 October 2017; pp. 1916–1923.
44. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.
45. Conti, V.; Rundo, L.; Militello, C.; Salerno, V.M.; Vitabile, S.; Siniscalchi, S.M. A Multimodal Retina-Iris Biometric System Using the Levenshtein Distance for Spatial Feature Comparison. *IET Biom.* **2021**, *10*, 44–64. [[CrossRef](#)]