Hierarchical fusion network for periocular and iris by neural network approximation and sparse autoencoder

Faisal Algashaam1 · Kien Nguyen1 · Jasmine Banks1 · Vinod Chandran1 · Tuan-Anh Do3 · Mohamed Alkanhal2

Received: 2 June 2019 / Revised: 17 May 2020 / Accepted: 12 October 2020 / Published online: 4 November 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
The eye region is one of the most attractive sources for identification and verification due to the representative availability of such biometric modalities as periocular and iris. Many score-level fusion approaches have been proposed to combine these two modalities targeting to improve the robustness. The score-level approaches can be grouped into three categories: transformation-based, classification-based and density-based. Each category has its own benefits, if combined can lead to a robust fusion mechanism. In this paper, we propose a hierarchical fusion network to fuse multiple fusion approaches from transformation-based and classification-based categories into a unified framework for classification. The proposed hierarchical approach relies on the universal approximation theorem for neural networks to approximate each fusion approach as one child neural network and then ensemble them into a unified parent network. This mechanism takes advantage of both categories to improve the fusion performance, illustrated by an improved equal error rate of the multimodal biometric system. We subsequently force the parent network to learn the representation and interaction strategy between the child networks from the training data through a sparse autoencoder layer, leading to further improvements. Experiments on two public datasets (MBGC version 2 and CASIA-Iris-Thousand) and our own dataset validate the effectiveness of the proposed hierarchical fusion approach for periocular and iris modalities.

Keywords
Periocular recognition · Iris recognition · Hierarchical multimodal fusion · Neural network approximation · Sparse autoencoder

1 Introduction
Biometrics, the knowledge of determining the identity of a person through the use of physical and/or behavioural traits, have gaining significant attention considering the increasing incidents of fraud challenges in highly secure identity authentication systems. Unlike traditional knowledge-based (e.g. PINs, passwords) and token-based (e.g. cards, keys) approaches, biometrics rely on human’s intrinsic characteristics which cannot be easily lost, forgotten or shared. Several human physiological (e.g. fingerprint, face, iris, palm vein, retina and ear) and behavioural (e.g. keystroke dynamics, signature, hand–eye coordination and hand tremors) characteristics have been successfully used as biometrics [1]. However, these unimodal biometric systems are widely known to be vulnerable to the noisy input, non-universal to the large-scale population, limited by an upper bound on identification accuracy and susceptible to spoof attacks [2].

Multibiometric systems seek to address these challenges by combining the evidence presented by multiple biometric
modalities in different stages. These multimodal systems can significantly enhance the overall recognition performance together with enlarging the population coverage, better preventing fraud and reducing the failure-to-enrol rate [3]. The advantages of the multibiometric systems over the traditional unimodal systems are summarized as: wider population coverage, better anti-spoofing, more robust to noisy data in a single modality, continuous monitoring when a single trait is not sufficient, fault tolerant when certain biometric sources become unreliable due to sensor or software malfunction or deliberate user manipulation [3].

The eye region provides a rich source of details for biometrics, where we can find multiple interesting traits including periocular, iris, sclera, retina and eye movement as shown in Fig. 1. In this paper, we only focus on the first two traits since they can be visibly captured non-invasively with high recognition rates, while the other three are either invasive or less accurate [3]. The two traits (periocular and iris) have their own pros and cons, but can complement each other. Iris is superior in recognition accuracy, but is limited in the imaging distance from the subject to the camera due to their small sizes. In contrast, periocular can be imaged from long distances due to its larger size, but is limited in the recognition accuracy. These two biometric modalities, when working together in a multimodal eye biometric system, can provide complementary information to improve the robustness of the system.

Depending on the levels in which the fusion is performed, the fusion approaches can be categorized into four groups: sensor level, feature level, score level and decision level [4]. While early fusion such as sensor level retains more information, it is also vulnerable to noise and superfluous details which are not used for classification. In contrast, late fusion such as decision level will work with more compact representations which are more directly used by the classification task, but lose details due to their reduced sizes. Feature-level fusion normally requires access to the feature representation, which may not be available, especially in commercial systems. Decision-level fusion is only performed on the rank orders, which retain limited information about each modality. Score level has been one of the most popular choices since it provides a balance among the amount of information, the availability of information and the robustness to noise. The existing score-level fusion approaches in the literature can be divided into three categories [5]:

- **Transformation-based**: The component scores are first normalized to a common domain and then fused using simple fusion operators such as weighted sum, product or order statistics. Most of approaches fusing periocular and iris in the literature use this category due to its simplicity [6–11].
- **Classification-based**: The component scores are used to train a classifier such as decision tree or support vector machines (SVMs) that discriminates between genuine and impostor features. This category is very popular in the general multimodal biometric fusion context [5], but there are surprisingly very few approaches discussing this category in the periocular and iris fusion context, except for [12] with AdaBoost.
- **Density-based**: These approaches are based on the likelihood ratio test, and they require estimating the density distribution of genuine and impostor match scores [13].

Each category has its own benefits, which if combined have the potential for a more robust fusion mechanism. In this paper, we propose a new approach to fuse two categories: transformation-based and classification-based into a unified framework for classification. The density-based approaches are not considered since they usually require a large number of match scores available to estimate the genuine and impostor distributions, which may not always feasible in real-life scenarios. Our proposed hierarchical fusion approach relies on the universal approximation theorem for neural networks, which states that for any continuous function, \( f \), on a compact set, \( K \), there exists a feedforward neural network with only a single hidden layer, which uniformly approximates \( f \) to within an arbitrary \( \varepsilon > 0 \) on \( K \) [14]. We approximate each category (transformation-based and classification-based) with a neural network and then ensemble these child neural networks in a larger parent neural network. This hierarchical mechanism takes advantage of both categories to improve the fusion performance, illustrated by an improved equal error rate (EER) of the multimodal biometric system. We introduce hierarchical fusion network for fusing periocular and iris. The hierarchical fusion network has three hidden layers; each is interpreted meaningfully in the multibiometric context.

The remainder of this paper is structured as follows: Section II describes our baselines for periocular and iris recognition; Section III introduces the proposed hierarchical fusion by neural network approximation approach; Section...
2 Iris and periocular recognition

This section will describe our recognition baseline approaches for periocular and iris.

2.1 Periocular recognition

The human periocular is often referred to as skin textures and anatomical features of the face region close to the eye, probably including the eye, eyelids, eyelashes and eyebrows [15, 16]. Periocular has recently emerged as a promising biometric trait for human identification due to a potentially attractive trade-off between the iris alone and the entire face, in terms of accuracy and visibility, especially for cases where neither the iris nor a full facial image can be acquired.

Many approaches have been proposed for periocular recognition in the literature [17–19], ranging from handcrafted features (e.g. [20–23]) to deep learning features (e.g. [24–26]). In this work, we choose to demonstrate our proposal on the elliptical higher-order-spectra periocular code features introduced by Faisal et al. in [27] due to the effectiveness of their approach. The proposed approach first samples the periocular region using an elliptical coordinate and then encodes the sampled region with higher-order spectral features to extract a feature vector called eHPC, which achieves robustness in scale, translation and head rotation. Figure 2 illustrates the elliptical sampling scheme and its eHPC feature vector. This approach obtains state-of-the-art recognition accuracy in the two periocular datasets: 97.71% for the Japanese Female Facial Expression (JAFFE) [28] and 99.52% for the Face Recognition Grand Challenge (FRGC) [29]. Details can be found in [27, 30].

2.2 Iris recognition

The iris is a protected internal organ of the eye, located behind the anterior cavity and the cornea, but in front of the lens [31]. The iris consists of muscle tissue that comprises a sphincter muscle that causes the pupil to contract and a group of dilator muscles that cause the pupil to dilate. Iris has been one of the most accurate biometric traits with very low false match rates and high processing speeds in large-scale datasets [32]. Such an observation is supported by the complex textural pattern of its stroma that varies across individuals, the perceived permanence of its distinguishing attributes, wide universality and limited genetic penetrance [31, 33].

Similar to periocular recognition, a plentiful number of approaches have been proposed for iris recognition in the literature [18, 32, 34], evolving from the classic IrisCode to handcrafted features (e.g. [35–37]) and more recently to modern deep learning features (e.g. [38–41]). In this work, we choose to demonstrate our proposal on the Gabor phase-quadrant features, which were first introduced by Daugman [42], but are still dominating the commercial iris recognition market today. The approach we employ is presented in Fig. 3. The iris region is first segmented from the eye image by two non-concentric boundary circles: the inner circle (or pupil circle) and the outer circle (or limbus circle). These two circular contours are detected based on an integro-differential operator [42, 43]. Then the segmented region is normalized to a fixed rectangular dimension so that it can be used for comparison. The normalization process relies on a rubber sheet model to transform the iris texture from Cartesian to polar coordinates. Each normalized iris is subsequently demodulated to extract the phase information using quadrature 2D Gabor wavelets to generate IrisCode as a final feature vector. Finally, the similarity between two IrisCodes is calculated based on their Hamming distance.

It is worth noting that any periocular and iris recognition approaches can be used to generate the match scores, the two chosen approaches here are only for illustration. More advanced approaches can be trivially plugged into achieve more advanced fusion results. These two approaches for periocular recognition and iris recognition will generate two match scores when comparing a new subject with one subject in the database. These two scores will be combined by approaches in the transformation-based category, the classification-based category and our proposed hierarchical fusion approach.

3 The proposed hierarchical fusion approach

In this section, we discuss the proposed approach to fuse one fusion approach in the transformation-based category
Fig. 3 Daugman’s iris recognition approach [42]

Fig. 4 Approximation of the transformation-based fusion category (sum, average) with a neural network. The $N$ inputs are $N$ match scores coming from biometric modalities of the $c$ identity classes. Two hidden neurons $HT_1$ and $HT_2$ represent weighted sum and average fusion approaches. $HT$ stands for a hidden neuron to approximate transformation category.

with one fusion approach in the classification-based category. We first explain the methods to approximate transformation-based category and classification-based category by neural networks (NNs). Then these two child NNs are fused to create a parent NN, which functions as hierarchical fusion.

3.1 Approximating the transformation-based category with a neural network

There are a number of fusion operators for the transformation-based category including sum of the match scores, product of the match scores and order statistics such as maximum, minimum or median of the match scores [5]. The sum has also been reinforced with weighted mechanisms such as uniform weighting (averaging) and matcher qualities (EERs). These two mechanisms can be trivially simulated by a one-layer neural network as shown in Fig. 4. The first hidden neuron, $HT_1$, functions like a weighted sum with the weights estimated by the EERs of matchers. The second hidden neuron, $HT_2$, functions like an averaging sum of the input match scores,

$$HT_1 = \sum_{i=1}^{N} w_i S_i \quad \text{and} \quad HT_2 = \sum_{i=1}^{N} \frac{1}{N} S_i.$$ 

3.2 Approximating the classification-based category with a neural network

The approaches in the classification-based category treat the input match scores as a multidimensional input vector and then classify it using classifiers such as SVM. Interestingly, Kang et al. have shown that a SVM can be explicitly approximated as a neural network [44]. Approximating a SVM with a neural network can significantly accelerate the computation with little or no prediction loss. A hybrid neural network (HNN) has been shown to be able to approximate support vector classification. We adapt their proposal to approximate our score fusion SVM as one-hidden-layer network as follows. Given a training dataset $D = \{s^i, y^i\}_{i=1}^{M}$, $y^i \in \{1, \ldots, c\}$ where $s^i$ is one score feature vector, $y^i$ is the identity class to which those scores belong and $M$ is the number of training samples, the algorithm for finding an approximated neural network has the following stages:

- **Stage 1** Train the SVC for the dataset, $D$. We train a multiclass SVM using the selected multiclass classification approach. Consequently, several decision functions $f_j \in \mathbb{R}$ of the SVM are obtained.
- **Stage 2** Construct a new training dataset, $D'$, relying on the decision functions of the SVM learned in the first stage. We first compute the decision function values, $f_j(s^i)$, for the original training dataset, $D$. After that, the new training dataset is constructed as $D' = \{s^i, \{f_1(s^i), \ldots, f_c(s^i)\}\}_{i=1}^{M}$.
- **Stage 3** Train the neural network for the new dataset, $D'$. A neural network with multiple outputs is trained with the new dataset $D'$, where the $j$th output node of the neural network corresponds to the decision function, $f_j$. The neural network is designed to have the number of output nodes equal to the number of decision functions.

The approximation process is illustrated in Fig. 5.

3.3 Proposed hierarchical fusion network

Once the child NN approximating the transformation-based category and the child NN approximating the classification-based category have been constructed, it is very interesting to observe that both are in the form of a one-hidden-layer neural network, with the same input and output. Observing that, it is straightforward to combine these two child NNs
Fig. 5 Approximation of the classification-based fusion category (SVM) with a neural network [44]. The $N$ inputs are $N$ match scores coming from biometric modalities of the $c$ identity classes. The $M$ hidden neurons $HC_1, \ldots, HC_M$ represent the decision fusion of the SVM modelled by the neural network. HC stands for a hidden neuron to approximate classification category.

Fig. 6 Proposed hierarchical fusion network with three novel hidden layers. The summary layer approximates and fuses the transformation-based category approaches and the classification-based category approach into a single hidden layer with the red and blue hidden neurons, respectively. The sparse autoencoder layer imposes a sparsity constraint to learn a compact representation of the input score, which facilitates extraction of meaningful combination of inputs. The interaction layer further forces neurons to interact with each other to enrich the representation capability, allowing the network to better model the interactions within the inputs to output accurate corresponding classes.

The major benefit of the proposed network is an effective way to combine the advantages of approaches from different categories in the score-level fusion by allowing them to concatenate (the summary layer) and to encode meaningful representation (the sparse autoencoder layer), and to encourage interaction within neurons (the third hidden layer). This hierarchical interaction mechanism leads to a unified framework for fusing fusion approaches. The proposed hierarchical fusion by network approximation and ensemble is applicable to multiple input modalities. It can also be trivially extended to include other fusion approaches from any categories if they can be approximated with neural networks (which is possible according to the universal approximation capability of neural network [14]). In the scenario of two modalities: periocular and iris, the fusion approach is simplified with only two-dimensional inputs.

It is noteworthy that combining and fusing multiple inputs using a neural network has been used before. However, compared with the conventional approaches where all layers and hidden neurons are manually decided, we introduce a novel methodology to make the network design meaningful. The insights from the neural network approximation theory of the modern fusion categories allow us to build intuition in the design of the summary layer. The insights from the sparsity representation learning allow us to build confidence in the design of the sparse autoencoder layer. While it may appear
similarly to other conventional fully connected neural networks, the hierarchical fusion network differs in nature due to the incorporation of the specific domain knowledge from the multibiometric field to make them achieve optimal performance for the multibiometric fusion task.

4 Experimental results

We choose two most relevant datasets from the literature for experiments.

- CASIA-Iris-Thousand [47]: contains NIR eye images captured by cooperative subjects in close-up distances. The images were captured by the dual-eye iris IKEMB-100 camera from IrisKing at the Institute of Automation, Chinese Academy of Sciences (CASIA). There are 20,000 NIR iris images captured in several sessions from 1000 subjects. This is the largest public periocular and iris dataset in the literature in terms of the number of subjects. Some example images are illustrated in Fig. 7.

- MBGC version 2 [48]: contains NIR face images captured by less cooperative subjects at standoff distances when the subjects walked through a portal from a 3-meter distance. There are 628 near infrared eye videos of 147 subjects, acquired by the Iris-On-the-Move (IOM) system. The videos are recorded, while the subjects are walking through a portal, which mounts the IOM system and NIR illuminators [49]. This is a challenging dataset since the subjects are long distance away from the camera and noise factors: reflections, luminosity, contrast, eyelid and eyelash iris obstruction and focus characteristics. These uncontrolled image acquisition conditions make this dataset well simulate the real-life imaging conditions. Some example images are illustrated in Fig. 8.

For the MBGC dataset, since there are a number of videos where the eyes are completely out of view or blurred, we removed those videos. For the rest, we selected high-quality frames to generate a database of 3000 frames from 120 subjects where the irises and periocular are clearly presented. For each subject, two frames are chosen to serve as templates, and the rest serve as the query frames to compare with the templates to search for the match. In total, there are \((3000 - 2 \times 120) \times (2 \times 120) = 662,400\) comparisons of periocular to be conducted. The left iris and right iris are treated as two different classes.

We also collect our own dataset to validate the experiments. We collected 212 images pertaining to 53 individuals captured using NIR spectrum using a IP2M-842B wireless network IP surveillance camera. The size of each image is 800 × 600 pixels. There is a balance between genders and age groups in the collected dataset. Images from our own dataset are not presented here due to the privacy.

4.1 Performance metrics

Two metrics are used in this work: genuine acceptance rate (GAR) at false acceptance rate (FAR) = 0.01% and equal error rate (EER). These two metrics have been widely used in the biometric context to analyse the performance of biometric systems. FAR presents the probability (the percentage of times) at which the system incorrectly declares that a biometric sample belongs to the claimed identity when the sample actually belongs to a different subject (impostor). Conversely, FRR indicates the probability at which the system incorrectly rejects a claimed identity when the sample actually belongs to the subject (genuine user). The first metric, GAR at FAR = 0.01%, gives us an estimate of how well the algorithm performs at a specific setting. The second metric, EER, is the rate at which FAR and FRR are equal. EER is usually estimated from the receiver operating characteristic (ROC) curve, which is a graph that illustrates the relationship between the FAR and the FRR. In general, the approach with the lowest EER is the most accurate.

4.2 Periocular alone performance

There are several parameters of the periocular recognition algorithm to be determined, including the periocular window size, the number of angles and radius local points. Similar to the intuition in [27], we choose a rectangular size of \((380 \times 340)\) for the periocular region, 8 angles and 12 local radial points. The chosen periocular algorithm achieves a GAR = 90.5% @FAR = 0.01% and an EER of 2.5% on the MBGC dataset.

4.3 Iris alone performance

There are several parameters of the iris recognition algorithm to be determined for segmentation, normalization and feature extraction. For segmentation and normalization, a rectangular size of \((256 \times 64)\) is chosen for the normalization process. For feature extraction, the base wavelength of the Gabor filter in pixel is 6 and the remaining power of the Gauss part of the filter relative to the centre is 0.01. We implemented our algorithm using an open-source software, USIT version 2.2, from the University of Salzburg [50]. The chosen iris recognition algorithm achieves a GAR = 89.4% @FAR = 0.01% and an EER of 3.1% on the MBGC dataset.

4.4 Hyper-parameter tuning

We first investigated the effect of the parameter choice for the hierarchical fusion network. A portion of the dataset
Hierarchical fusion network for periocular and iris by neural network approximation

Table 1 Effect of the number L of hidden neurons in the interaction layer on the modelling capacity of the parent neural network, in terms of genuine acceptance rate (GAR) at false acceptance rate (FAR) = 0.01%

| Validation subset | 3   | 4   | 5   | 6   | 7   | 8   | 9   |
|-------------------|-----|-----|-----|-----|-----|-----|-----|
| 53.8              | 68.7| 89.2| 91.5| 93.7| 92.7| 91.1|

The performance comparison is given in Table 2.

Fig. 7 Some example images from two subject S5001 and S5019 from the CASIA-Iris-Thousand dataset. This is to demonstrate the capacity of the proposed approach to work with cooperative subjects and close-up imaging distances.

Fig. 8 Some example frames from one video of the subject 05308d66 walking through the IOM portal in the first row. The second row zooms in the periocular and iris regions of the centre frame. This is to demonstrate the capacity of the proposed approach to work with less cooperative subjects and standoff imaging distances.

4.5 Hierarchical fusion

Each fusion approach has its own strengths. When combined, they can be complimentary to each other, leading to better fusion performance. Two approaches in the transformation-based category (weighted sum and product) have been implemented in this work, achieving GARs of 93.97% and 95.13% @FAR = 0.01%, respectively, on the MBGC dataset. The SVM approach in the classification-based category has also been implemented with a GAR of 95.8% @FAR = 0.01%. Our proposed fusion-by-fusion approach fuses these approaches from the two categories, leading to improvement in the overall GAR of 97.94% @FAR = 0.01% on the MBGC dataset.

Similar improvements can be observed on the CASIA-Iris-Thousand dataset. Two approaches in the transformation-based category (weighted sum and product) achieved GARs of 94.55% and 96.68% @FAR = 0.01%, respectively. The classification-based approach achieved a GAR of 97.36% @FAR = 0.01%. The proposed hierarchical fusion approach outperformed both, achieving a GAR of 98.27% @FAR = 0.01%.

The same improvements can also be achieved on our own dataset. Two approaches in the transformation-based category (weighted sum and product) achieved GARs of 89.69% and 88.01% @FAR = 0.01%, respectively. The classification-based approach achieved a GAR of 91.41% @FAR = 0.01%. The proposed hierarchical fusion approach outperformed both, achieving a GAR of 94.15% @FAR = 0.01%.

The performance comparison is given in Table 2.
Table 2: Performance comparison of various fusion approaches in terms of genuine acceptance rate (FAR) and equal error rate (EER) for on the MBGC version 2, CASIA-Iris-Thousand and our own dataset

| Datasets         | Transformation-based fusion | Classification-based fusion | The proposed hierarchical fusion |
|------------------|-----------------------------|-----------------------------|---------------------------------|
|                  | Weighted sum | Product | Weighted sum | Product | Weighted sum | Product |
| MBGC version 2   | GAR@0.01%    | EER     | GAR@0.01%    | EER     | GAR@0.01%    | EER     |
|                  | 93.97        | 2.13    | 95.13        | 1.92    | 95.82        | 1.87    |
| CASIA            | 94.55        | 1.98    | 96.68        | 1.87    | 97.36        | 1.45    |
| Ours             | 89.69        | 3.52    | 88.01        | 3.76    | 91.41        | 3.27    |

5 Conclusions

In this paper, we have proposed a new mechanism to combine multiple fusion approaches in a unified framework. We employ a hierarchical fusion scheme, which relies on the universal approximation capability of neural network to approximate each fusion approach with a child neural network and then combine them into a parent neural network. The proposed hierarchical fusion network has three hidden layers: a summary layer, a sparse autoencoder layer and an interaction layer. The neural network-based approximation mechanism allows the system automatically to explore and discover the best strategy to combine the component scores, leading to a better verification accuracy in comparison with each component approach itself. The proposed approach has been validated in the MBGC dataset, the CASIA-Iris-Thousand dataset and our own dataset with promising results, outperforming other baselines. The proposed approach can also be extended to fuse multiple biometric modalities other than simple periocular and iris.

References

1. Jain, A.K., Ross, A., Prabhakar, S.: An Introduction to Biometric Recognition. IEEE Trans. Circuits Syst. Video Technol. 14, 4–20 (2004)
2. Ross, A., Nandakumar, K., Jain, A.: Introduction to multibiometrics. In: Jain, A., Flynn, P., Ross, A. (eds.) Handbook of Biometrics, pp. 271–292. Springer, Berlin (2008)
3. Ross, A., Nandakumar, K., Jain, A.: Handbook of Multibiometric. Springer, Berlin (2006)
4. Ross, A., Jain, A., Nandakumar, K.: Levels of fusion in biometrics. In: Handbook of Multibiometrics, pp. 59–90. (2006)
5. Ross, A., Nandakumar, K.: Fusion, score-level. In: Li, S.Z., Jain, A. (eds.) Encyclopedia of Biometrics, pp. 611–616. Springer, Boston (2009)
6. Woodard, D.L., Pundkil, S., Miller, P., Jillela, R., Ross, A.: On the fusion of periocular and iris biometrics in non-ideal imagery. In: 20th International Conference on Pattern Recognition (ICPR), pp. 201–204. (2010)
7. Alonso-Fernandez, F., Mikaelyan, A., Bigun, J.: Comparison and fusion of multiple iris and periocular matchers using near-infrared and visible images. In: 3rd International Workshop on Biometrics and Forensics (IWBF 2015), pp. 1–6. (2015)
8. Trokielewicz, M., Szadkowski, M.: Iris and periocular recognition in arabian race horses using deep convolutional neural networks. In: 2017 IEEE International Joint Conference on Biometrics (IJCB), 1–4 Oct. 2017, Piscataway, NJ, USA, pp. 510–16. (2017)
9. Ahmed, N.U., Cvetkovic, S., Siddiqui, E.H., Nikiforov, A., Nikiforov, I.: Combining iris and periocular biometric for matching visible spectrum eye images. Pattern Recogn. Lett. 91, 11–16 (2017)
10. Chun-Wei, T., Kumar, A.: Human identification from at-a-distance images by simultaneously exploiting iris and periocular features. In: 2012 21st International Conference on Pattern Recognition (ICPR 2012), 11–15 Nov. 2012, Piscataway, pp. 553–556. (2012)
11. Raja, K.B., Raghavendra, R., Stokkenes, M., Busch, C.: Multimodal authentication system for smartphones using face, iris and periocular. In: 8th IAPR International Conference on Biometrics, ICB 2015, May 19, 2015–May 22, 2015, Phuket, Thailand, pp. 143–150. (2015)
12. Oishi, S., Ichino, M., Yoshiura, H.: Fusion of iris and periocular user authentication using AdaBoost for mobile devices. In: 2015 IEEE International Conference on Consumer Electronics (ICCE), 9-12 Jan. 2015, Piscataway, NJ, USA, pp. 428–429. (2015)
13. Nandakumar, K., Chen, Y., Dass, S.C., Jain, A.: Likelihood ratio-based biometric score fusion. IEEE Trans. Pattern Anal. Mach. Intell. 30, 342–347 (2008)
14. Hornik, K.: Approximation capabilities of multilayer feedforward networks. Neural Networks 4, 251–257 (1991)
15. Jillela, R., Ross, A.A., Boddei, V.N., Kumar, B.V.K.V., Hu, X., Plemmons, R., et al.: Iris segmentation for challenging periocular images. In: Burge, J.M., Bowyer, W.K. (eds.) Handbook of Iris Recognition, pp. 281–308. Springer, London (2013)
16. Ambika, R.D., Radhika, R.K., Seshachalam, D.: Periocular Recognition Using Reduced Features. Int. J. Comput. Sci. Inf. Technol. 5, 151–158 (2013)
17. Alonso-Fernandez, F., Bigun, J.: A survey on periocular biometrics research. Pattern Recogn. Lett. 82(2), 92–105 (2016)
18. Nigam, I., Vatsa, M., Singh, R.: Ocular biometrics: a survey of modalities and fusion approaches. Information Fusion 26, 1–35 (2015)
19. Algashaam, F.M., Nguyen, K., Alkanhal, M., Chandran, V., Boles, W., Banks, J.: Multispectral periocular classification with multimodal compact multi-linear pooling. IEEE Access 5, 14572–14578 (2017)
20. Ambika, D.R., Radhika, K.R., Seshachalam, D.: Periocular authentication based on FEM using Laplace-Beltrami eigenvalues. Pattern Recogn. 50, 178–194 (2016)
21. Park, U., Jillela, R.R., Ross, A., Jain, A.K.: Periocular biometrics in the visible spectrum. IEEE Trans. Inf. Forensics Secur. 6, 96–106 (2011)
22. Juefei-Xu, F., Savvides, M.: Subspace-based discrete transform encoded local binary patterns representations for robust perioc-
ular matching on NIST’s face recognition grand challenge. IEEE Trans. Image Process. 23, 3490–3505 (2014).

23. Proença, H., Briceno, J.C.: Periocular biometrics: constraining the elastic graph matching algorithm to biologically plausible distortions. IET Biometrics 3, 167–175 (2014)

24. Proença, H., Neves, J.C.: Deep-PWIS: periocular recognition without the iris and sclera using deep learning frameworks. IEEE Trans. Inf. Forensics Secur. 13, 888–896 (2018)

25. Zhao, Z., Kumar, A.: Accurate periocular recognition under less constrained environment using semantics-assisted convolutional neural network. IEEE Trans. Inf. Forensics Secur. 12, 1017–1030 (2017)

26. Raghavendra, R., Raja, K. B., Busch, C.: On comparison score fusion of deep autoencoders and relaxed collaborative representation for smartphone based accurate periocular verification. In: 2016 19th International Conference on Information Fusion (FUSION), 5–8 July 2016, Piscataway, NJ, USA, pp. 2221–2228. (2016)

27. Algashaam, F., Nguyen, K., Chandran, V., Banks, J.: Elliptical higher-order-spectra periocular code. IEEE Access 5, 6978–6988 (2017)

28. Daugman, J.G.: High confidence visual recognition of persons by a test of statistical independence. IEEE Trans. Pattern Anal. Mach. Intell. 15, 1148–1161 (1993)

29. Jiang, S., Cho, S.: Approximating support vector machine with artificial neural network for fast prediction. Expert Syst. Appl. 41, 4989–4995 (2014)

30. Shao, Z., Wang, L., Wang, Z., Deng, J.: Remote sensing image super-resolution using sparse representation and coupled sparse autoencoder. IEEE J. Select. Topics Appl. Earth Observ. Remote Sens. 12(8), 2663–2674 (2019). https://doi.org/10.1109/JSTARS.2019.2925456

31. Nianfeng, L., Man, Z., Haiqing, L., Zhenan, S., Tieniu, T.: Iris Recognition and human-interpretable features. In: 2015 International Conference on Image and Vision Computing New Zealand (IVCNZ), pp. 1–6. (2015)

32. Ross, A.: Iris recognition: the path forward. Computer 43, 30–35 (2010)

33. Bowyer, K.W., Hollingsworth, K., Flynn, P.J.: Image understanding for iris biometrics: a survey. Comput. Vis. Image Underst. 110, 281–307 (2008)

34. Nguyen, K., Fookes, C., Jillela, R., Sridharan, S., Ross, A.: Long range iris recognition: a survey. Pattern Recogn. 72, 123–143 (2017)

35. Bowyer, K.W., Burge, M.J.: Handbook of Iris Recognition. Springer, London (2016)

36. Shao, Z., Wang, L., Wang, Z., Deng, J.: Remote sensing image super-resolution using sparse representation and coupled sparse autoencoder. IEEE J. Select. Topics Appl. Earth Observ. Remote Sens. 12(8), 2663–2674 (2019). https://doi.org/10.1109/JSTARS.2019.2925456

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Faisal Algashaam is pursuing his Ph.D. at the Queensland University of Technology. His research focuses on periocular biometrics. His research interests are in the application of computer vision and machine learning techniques to biometrics and scene understanding.

Dr. Kien Nguyen is a Research Fellow at the Queensland University of Technology. He has been conducting research in the area of iris recognition and biometrics for the last 10 years and has published his research in high-quality journals and conferences in the area. His research interests are in the application of computer vision and deep learning techniques to the areas of biometrics, surveillance, and scene understanding. He has been serving as an Associate Editor of the journal IEEE Access in the area of biometrics since 2016.

Dr. Jasmine Banks received the B.Eng. degree in electronics and information technology and the Ph.D. degree from the Queensland University of Technology, in 1993 and 2000, respectively. She is currently a Lecturer with the School of Engineering Systems, Queensland University of Technology. Her research interests include artificial intelligence and image processing, computer hardware, and electrical and electronic engineering.

Dr. Vinod Chandran received a Ph.D. degree in electrical and computer engineering from Washington State University in 1990. He retired as
a Professor with the Queensland University of Technology (QUT), Brisbane, QLD, Australia, in 2016. He has authored or coauthored over 200 international journal articles and conference papers. His research interests include span signal processing, image processing, and machine learning—applied to biometrics and biomedical systems.

Tuan-Anh Do received the masters degree from Hanoi University of Technology in 2008, and he is currently a doctoral graduate student at Hanoi University of Technology. He is a lecturer in the Department of Computer Science, School of Information and Communication Technology, Hanoi University of Technology. His research interests include algorithm analysis, combinatorics optimization, software testing and debugging, and AI. He also has extensive experience in technology, as he has worked for many years in the business environment.

Mohamed Alkanhal received the B.Sc. degree (Hons.) in computer engineering from King Saud University, Riyadh, Saudi Arabia, the M.Sc. degree from the University of Southern California, Los Angeles, CA, USA, and the Ph.D. degree from Carnegie Mellon University, Pittsburgh, PA, USA. He was with the Computer Research Institute, King Abdulaziz City for Science and Technology (KACST), Riyadh, as a Research Faculty Member from 2001 to 2006 and the Director from 2006 to 2014. He is currently the Director of the Communications and Information Technology Research Institute, KACST. He has authored or co-authored many technical papers and has three patents pending. He has done extensive research in several areas, including pattern recognition, digital signal processing, and natural language processing.