Enhanced Biometric Access Control for Mobile Devices

Dane Brown*†, Karen Bradshaw †
*Modelling and Digital Science, Council for Scientific and Industrial Research, Pretoria, South Africa.
[1]dbrown@csir.co.za
†Department of Computer Science, Rhodes University, Grahamstown, South Africa.
[2]k.bradshaw@ru.ac.za

***Abstract*—In the new Digital Economy, mobile devices are increasingly being used for tasks that involve sensitive and/or financial data. Hitherto, security on smartphones has not been a priority and furthermore, users tend to ignore the security features in favour of more rapid access to the device. We propose an authentication system that can provide enhanced security by utilizing multi-modal biometrics from a single image, captured at arm's length, containing unique face and iris data. The system is compared to state-of-the-art face and iris recognition systems, in related studies using the CASIA-Iris-Distance dataset and the IITD iris dataset. The proposed system outperforms the related studies in all experiments and shows promising advancements to at-a-distance iris recognition on mobile devices.**

***Index Terms*—at-a-distance, face, iris, mobile, multi-modal**

I. Introduction

Since their invention four decades ago, mobile phones have evolved from being purely tools for making voice calls to being used for an ever increasing variety of tasks, such that they have become indispensable in everyday life. With the advent of the Digital Economy, the use of smartphones has expanded specifically to include tasks such as email, banking, social networking, web surfing, e-commerce and more recently bitcoin wallets [1].

A potential threat has arisen as a result of the transformation of the role of mobile phones in that when it was solely used for communication, security was not a high priority. Nowadays, however, due to the sensitive nature of the tasks performed using smartphones and the fact that much personal and/or company information is stored on the device itself, access control (typically in the form of encrypted text passwords that require memorization by the user) has become critical [1]. One-time pins (OTPs) do provide an extra layer of protection, but are useless when the mobile device itself is captured by the attacker and do not provide a solution for on-device tasks.

Moreover, studies have shown that human behaviour in the use of smartphones tends towards convenience and rapid access of applications installed thereon, with the result that often security features protecting access to the device are compromised [1]. For example, swipe access patterns are simplified and passwords are shortened to facilitate rapid access to the device. Moreover, security built into the applications is often bypassed when passwords are saved within their respective applications on the device. Although biometric authentication using a single biometric, such as fingerprint or face, is growing for mobile devices, there is still a risk of compromise due to forgeries [2] [3].

Multi-modal biometrics can provide a solution to both smartphone authentication and rapid access to the applications on the device by combining multiple sources of biometric data. However, leveraging the combined data is not appropriate to every application, for instance, when user convenience is compromised. A growing trend toward mobile and at-a-distance biometric acquisition has been observed due to the recent focus on frontal sensor variety and quality [4]. This includes the use of the face and iris region to unlock the mobile device [5]. Since these two biometrics can be acquired simultaneously at-a-distance, the inconvenience to the user is similar to utilizing a single biometric but with improved accuracy and security. On the other hand, this provides opportunities for robustness in the form of capturing the face, iris or both based on the angle at which the mobile device is held, thereby reducing user inconvenience by replacing text passwords.

Face recognition is a highly visible and user-friendly biometric used for authentication, or as a secondary biometric as a profile picture [6]. As a primary means of identification, the face is challenging in uncontrolled applications due to varied pose angles and occlusions. Recent advancements include an automatic face alignment system requiring milliseconds per face [7] and the handling of pose angles up to 60° by modelling a single frontal database image. While the latter is still not feasible on a mobile device without high-performance cloud computing, the former enables direct authentication, thereby lowering the risk of interception. This paper thus assumes that the majority of face data will be near-frontal, which we define as a head pose of less than 30°in any direction.

The use of iris texture analysis for biometric recognition is well established. The United Arab Emirates have used iris recognition for border control since 2001 and estimate that 2.7 billion iris cross-comparisons are done per day [8]. Its recent growth is attributed to it being the most accurate external image biometric [9]. Recently, iris sensors have attained improved capturing range and reduced price for commercial use [5]. The increased capturing range, however, also increases human error and can result in a reduction of texture detail.

In this paper, we propose an authentication system for

mobile devices that can provide sufficient security based on the application. while at the same time minimizing inconvenience to the user. To achieve this, the face and iris are captured simultaneously by requiring the user to only gaze once at the front of the device. Furthermore, face and iris fusion combinations are investigated to establish the additional degree of security that can be obtained with the extra information. Finally, the emerging biometric trend of at-a-distance iris recognition is investigated to determine whether it is comparable to the well-established face biometric.

The rest of the paper is organized as follows: Section II presents related face, iris and fused systems found in the literature. Section III discusses the construction of the system optimized for mobile devices. The experimental analyses and results are discussed in Section IV. Section V concludes the paper and outlines future work.

II. Related Studies

This section consists of related studies of the face, iris and their combined features.

A. Face

Face segmentation prunes away dynamic features such as background and hair of an individual. Consistently removing these dynamic features during segmentation is a well-researched problem. Kazemi and Sullivan [7] introduced a new face segmentation technique that uses gradient boosting for learning an ensemble of regression trees. The system is not only accurate but surpasses real-time performance at approximately 1 ms per image on a single processor. The face landmarks are automatically computed from a sparse subset of pixel intensities. Furthermore, the system is capable of handling missing or partially labelled data. The resulting coordinates for face landmarks are used to segment the face.

Face recognition performs matching or classification on a segmented face. A well-known face recognition system, known as DeepFace [10], uses a nine-layer deep neural network with over 120 million parameters. It achieves a high accuracy by training a model with over 4 million labelled faces. This large training model is unfortunately not feasible on mobile devices.

A recent approach by Haghighat *et al.* [11] accurately performs face recognition using only a single training sample, without 3D modelling or deep learning. Their approach aims to frontalize all segmented faces using a base mesh per individual without any face detection preprocessing. The disadvantage of their approach is the added time complexity of 40 Gabor filters in five scales and eight orientations.

B. Iris

While it is well known that the iris is the most accurate external biometric, this is achieved on short-range sensors [9]. The challenge is improving iris recognition when using at-a-distance sensors.

There are two main iris segmentation methods namely, Integro-differential operator and Hough transform (HT) [4]. The Integro-differential operator applies an exhaustive search for both the centre and radius of the iris, independently, by calculating the maximum in the blurred derivative with respect to the increasing radius of normalized contour integrals along circular trajectories. On the other hand, HT uses binary edge maps to localize pupil and iris boundaries. Votes are accumulated to estimate the parameters of the boundary concentric circles.

Umer *et al.* [12] used an HT algorithm suitable for iris extraction. The algorithm, known as Restricted Circular Hough transformation (RCHT), searches for circles bounded by the upper and lower eyelids. Texture within two concentric circles are extracted, constituting the feature vectors of the iris. The resulting feature vectors of each eye are combined into a larger vector, which is classified using support vector machines (SVMs).

C. Face and Iris Fusion

Eskandari *et al.* designed a robust fusion scheme for the face and iris at the feature-level [13]. They emphasize that feature alignment plays a pivotal role during the pre-processing step. A comprehensive set of experiments show the effect alignment has on recognition accuracy relative to the biometric data. The face is segmented and aligned using the two detected eye positions as reference. Irises are segmented and rotated for alignment based on the face's pose position. Furthermore, their system is made robust to noise by applying particle swarm optimization. However, this substantially increases computational complexity.

Verification accuracy was evaluated on the CASIA-Iris-Distance dataset, containing close-up near-infrared face images. Feature alignment and low false matches was the focus of evaluation, measured at 0.01% FAR. Both irises benefited significantly from alignment with a 44% improvement to verification accuracy. Aligning the face improved verification accuracy by 7%. Furthermore, fusing the face and both irises improved the accuracy by 14%. The face, fused irises and all three fused achieved a GAR of 87.33%, 71.55% and 94.44%, respectively.

III. Proposed System

The following subsections detail the proposed face and iris biometric recognition method. The system was coded in C++ using the OpenCV and Dlib image processing libraries. The proposed method demonstrates a low complexity biometric recognition algorithm suitable for mobile devices. Fig. 1 provides an overview of the algorithm, separating the face on the left and the iris on the right of the diagram. This diagram is referred to throughout this section, in which the different phases of the system are explained.

A. Feature Detection

1) Face: An initial region of interest (ROI) is determined by detecting the face by classifying Histogram of Gaussian (HoG) features using a linear SVM and a sliding window resulting in a set of regressors. Face detection is repeated until

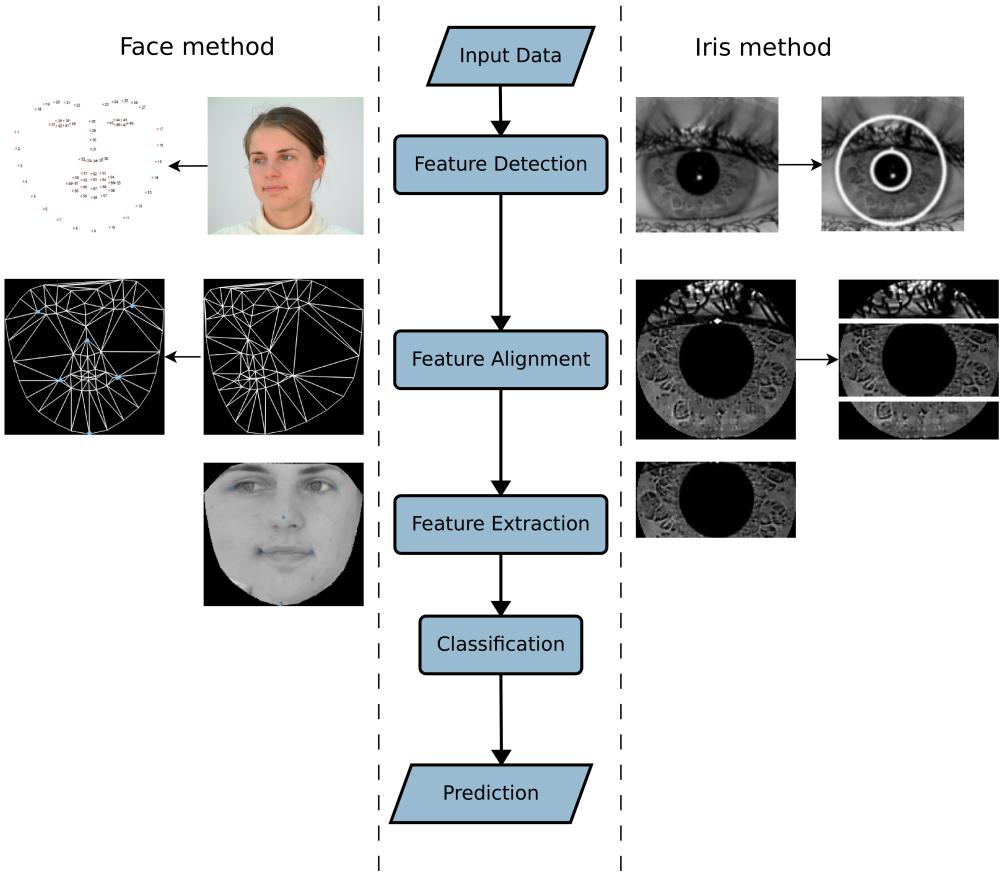

Fig. 1. Proposed face and iris recognition solution.

it is successful by rotating about the x-axis in 5° increments in both directions.

68 interpolated landmarks are detected, within the initial face ROI, based on a model as shown in Fig. 1. The landmark model used in this paper was trained on the iBUG 300-W face annotated dataset in the form of a cascade of regression trees [14]. The time complexity is reduced when training and testing the landmark detector by calculating the transform and warping only once at each level of the cascade. Furthermore, training is only performed once, allowing the landmark detector to predict on unseen faces. This allows for real-time face processing directly on the mobile device.

2) Iris: The RCHT method is applied on the input data used to capture the ROI around the pupil of the ocular image [12]. The consistency of the RCHT method is improved by applying a Laplacian of Gaussian (LoG) filter with 5×5 and 17×17 kernels, respectively. The iris boundary is thus determined by contrasting the sclera and the iris region below or horizontal to the pupil, depending on boundary threshold. This threshold caters for squinting and off-centre eye gazing. The result is shown by the bounding concentric circles.

B. Feature Alignment

To ensure autonomous feature alignment, a fall-back mechanism is used by relying on the confidence score produced by the fast local binary pattern histogram (LBPH) method combined with a sliding window [15]. This threshold can also be adjusted for strict rejection of low quality data when requiring high security. The following feature alignment methods are used for the face and iris respectively before the fall-back mechanism is employed.

1) Face: A face mesh is constructed based on the detected landmarks. The face mesh is constructed using Delaunay triangulation, such that no landmark is inside the circumcircle of any triangle. The mesh posing toward the left is corrected to a frontal pose as shown in the diagram. This method allows for fast face frontalization at angles up to 30°.

2) Iris: Inverted Gaussian filter with 11×11 kernel is used to sharpen the ROI, which consists of the concentric circles. The top and bottom of the iris region are pruned at the edge of the pupil to once again cater for squinting and off-centre eye gazing. However, before the data is pruned, the iris texture is aligned by rotating the image of the query image according to the image. This enables the alignment of non-occluded texture.

C. Feature Extraction

The previous steps minimized intra-class variation, which is imperative for good classification. However, maximizing inter-class variation is also necessary as it plays a significant role in reduces false matches. The following feature extraction and normalization techniques are used for the face and iris.

A combination of LoG and circular local binary patterns

(LLBP) was found to be particularly effective at maximizing inter-class variation for image-based biometrics, in previous research [15]–[17]. The LoG filter removes unwanted features on the low and high frequency spectrum before enhancing the remaining features, effectively increasing the mean signal. The Gaussian and Laplacian kernels were 15×15 and 7×7, respectively. This also further lowers intra-class variation by reducing subtle differences in images of the same person often caused by facial expressions and distortions due inconsistent gazing at the mobile device. This was applied to all of the normalized and segmented data. When the LLBP operator is used in this way, it reduces lighting differences without a noise side effect. The resulting images are resized to 75×75 before classification.

D. Classification

The Eigen classification method was found to produce optimal results for image-based biometrics. This classifier maximizes the total variance in data based on a linear combination of features. The largest variance in data is contained within the first few principal components which are modelled into classes. This significant reduction in data allows for fast verification on mobile devices. The trained and test models are compared based on the distances between eigenvalues during matching.

Given N sample images x_k the total scatter matrix is defined as [18]:

$$S_t = \sum_{k=1}^{N} (x_k - \mu)(x_k - \mu)^T,$$

where $m \in \mathbb{R}^n$ is the mean image obtained from the samples.

IV. Experimental Analysis and Results

In this section user convenience and security are assessed in two separate experiments on appropriate datasets discussed below and in Section II. The state-of-the-art related studies that was discussed contain datasets relevant to applications on mobile devices. Therefore, both verification and identification results are included where possible. The genuine acceptance rate (GAR) is the evaluation metric in all given cases. The false acceptance rate (FAR) is $100 - GAR$ in percentage unless otherwise stated.

A. Experimental Datasets

The following three datasets were used for empirical testing.

CASIA-Iris-Distance is a challenging dataset used in the experiments discussed in Section IV [5]. The data was captured by an at-a-distance sensor – developed inhouse by the authors – at 2352×1728 and at three metres away. The dataset was collected in an effort to show that the iris can be a robust biometric.

This dataset is organized as follows: 90 individuals are selected for comparison against Eskandari *et al.*'s system [13]. The eyes are first separated from the face, enabling the capturing of the iris during feature detection. Iris data is extracted from the face by Haar cascades that are trained to differentiate between face and iris features [19]. This serves as experimental data for handheld camera sensors capable of capturing near-infrared data of the face and iris simultaneously, which is an emerging trend on smartphones. Five samples are used for training and five samples are used for testing per 90 individuals.

TABLE I
Verification results on the CASIA-Iris-Distance dataset.

Method	Face (%)	Fused irises (%)	Face and irises fused (%)
Eskandari *et al.*	87.33	71.55	94.44
Proposed	92.05	78.87	97.88

The FERET b-series [20] face dataset consists of 200 individuals captured at 256×384 and is limited to five pose angles that are less than 30°.

The IITD iris dataset was captured using a JPC1000 Iris Recognition Camera sensor at 320×240 and handheld distance [21]. This sensor is referred to as: "The smallest, the fastest and the cheapest Iris recognition camera (sensor) module in the world." Therefore, the IITD dataset serves as an ideal test for the viability of mass produced iris sensors on mobile devices. Four samples are used for training and one sample is used for testing.

B. Experiment 1

Verification performs authentication by checking if the eigenvalue distance between a training sample of a known class and a test sample are below a threshold. Table I compares the verification performance of the proposed face and iris systems to Eskandari *et al.*'s approach on the CASIA-Iris-Distance dataset. The verification accuracy is evaluated as 0.01% FAR. The face, two irises and the fusion of all three perform better when using the proposed method. The improvement on irises is slightly more pronounced than that on the face. This is attributed to the iris algorithm focussing on at-a-distance performance.

C. Experiment 2

1) Face: This experiment demonstrates the face results on the FERET b-series [20] dataset. Haghighat *et al.* and the proposed system achieve 100% and 98.5% verification accuracy, respectively, each using only a single training sample per individual. This slight increase is attributed to the difference in complexity between the two systems. The proposed system is approximately 10 times faster than Haghighat *et al.*'s system as determined in previous research [15].

Identification finds the best match within N number of classes by finding the closest eigenvalue distance between a test sample and the training samples of all N classes, if it is below a threshold.

TABLE II
VERIFICATION RESULTS ON THE IITD IRIS DATASET.

Method	Left (%)	Right (%)	Fused (%)
Umer *et al.*	98.12	98.23	99.55
Proposed	99.55	99.55	100

TABLE III
IDENTIFICATION RESULTS ON THE IITD IRIS DATASET.

Method	Left (%)	Right (%)	Fused (%)
Umer *et al.*	97.75	97.05	98.37
Proposed	98.66	97.32	100

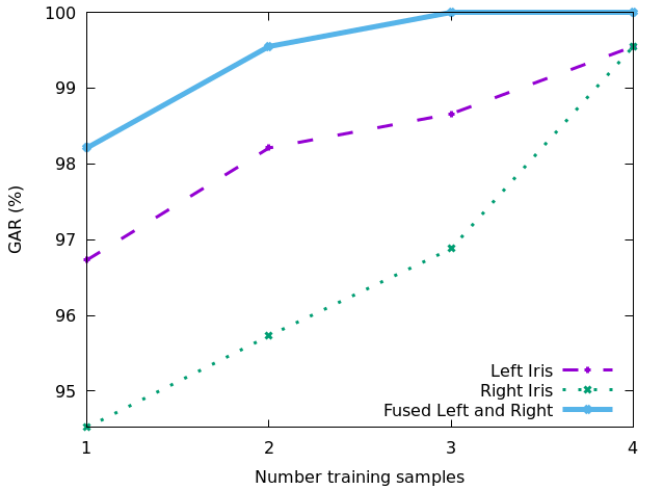

Fig. 2. Verification results on the IITD iris dataset using one to four training samples.

The proposed system achieved an 84.38% identification accuracy using a single sample. Haghighat *et al.* did not test identification accuracy.

2) Iris: Table II and III compare the respective verification and identification performance of the proposed iris system to Umer *et al.*'s related system on the IITD dataset. The proposed system outperforms all of Umer *et al.*'s results for both verification and identification. However, since they only use a single test sample per person, and the rest for training, the comparison is limited. The proposed system's verification results are thus elaborated in Fig. 2.

Fig. 2 shows that promising results are obtained by using a separate sensor for the iris. This is beneficial to the security of financial applications that require the highest security. The face is already known to perform well on the mobile phone and was thus not investigated further. Our iris algorithm produces promising results on handheld iris sensors. This is evident by the fact that over 98% GAR is achieved when combining the irises despite only requiring a single training sample. This is the same result as the face under similar conditions. Furthermore, a perfect accuracy is a achieved by using three or more training samples. In contrast, the state-of-the-art system does not achieve a perfect accuracy even when using four training samples.

V. CONCLUSION AND FUTURE WORK

A secure means of conducting monetary transactions in this Digital Economy has become increasingly important. While passwords and OTPs are required for banking transactions, they can be intercepted by an attacker and do not solve on-device tasks that require security. The addition of single biometrics to mobile devices provides an extra layer of security but at the risk of forgery.

This paper shows the benefits of multi-modal biometrics on mobile devices by demonstrating an improved recognition of both genuine and impostor users, thereby improving high-security applications. Moreover, user convenience can also be improved, particularly when using the face and iris due to their synergy for ease of acquisition and supplementation of unique features. Iris recognition at-a-distance was shown to be as effective as the well-established face recognition.

Two experiments were set up to appropriately demonstrate the performance of the face and iris on mobile devices. The CASIA-Iris-Distance dataset was evaluated using an iris algorithm optimized for real-time and at-a-distance use on mobile devices. The proposed system significantly outperformed the recent state-of-the-art iris verification system by Eskandari *et al.* The face was also evaluated on the same dataset and also outperformed Eskandari *et al.*'s system. Finally, fusing the face and iris yielded a near-perfect accuracy, while attaining the advantage of increased forgery mitigation.

Another experiment was conducted on the larger IITD dataset consisting of a cost-effective handheld dedicated iris sensor to further emphasize the mobile device optimized iris algorithm when focusing on enhanced security. This experiment showed that fusing both irises can achieve a perfect accuracy on larger datasets without the help of the face or other well-established at-a-distance biometrics. This is promising as it eliminates the need for typing passwords, thereby speeding up security applications to be instantaneous with a simple gaze at the camera of the mobile device. The results of at-a-distance iris recognition in this paper are positive whether combining it with the face, fusing two irises or simple using a single iris.

In future, more security tests will be conducted such as spoof attacks and liveness detection. While the system can run at real-time on mid to top range smartphones, another feasibility test can be conducted in the form of a mobile app that interfaces as an extra security layer instead of integrating it directly into financial-based apps.

ACKNOWLEDGEMENT

The authors would like to thank the CSIR, Information Security department for their financial support. Thank you to the authors of the publicly available datasets used in this paper. This work was undertaken in the Distributed Multimedia CoE at Rhodes University, with financial support from Telkom SA, Tellabs/CORIANT, Easttel, Bright Ideas 39, THRIP and NRF SA (UID 90243). The authors acknowledge that opinions, findings and conclusions or recommendations expressed here are those of the authors and that none of the

above mentioned sponsors accept liability whatsoever in this regard.

REFERENCES

[1] W. Melicher, D. Kurilova, S. M. Segreti, P. Kalvani, R. Shay, B. Ur, L. Bauer, N. Christin, L. F. Cranor, and M. L. Mazurek, "Usability and security of text passwords on mobile devices," in *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, ser. CHI '16. New York, NY, USA: ACM, 2016, pp. 527–539. [Online]. Available: http://doi.acm.org/10.1145/2858036.2858384

[2] S. Bharadwaj, M. Vatsa, and R. Singh, "Biometric quality: from assessment to multibiometrics," 2015.

[3] D. Kaur and G. Kaur, "Level of fusion in multimodal biometrics: A review," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 3, no. 2, pp. 242–246, 2013.

[4] C. Rathgeb, A. Uhl, and P. Wild, *Iris biometrics: from segmentation to template security*. Springer Science & Business Media, 2012, vol. 59.

[5] Casia iris image database. [Online]. Available: http://biometrics.idealtest.org/

[6] A. Jain, L. Hong, and S. Pankanti, "Biometric identification," *Communications of the ACM*, vol. 43, no. 2, pp. 90–98, 2000.

[7] V. Kazemi and J. Sullivan, "One millisecond face alignment with an ensemble of regression trees," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1867–1874.

[8] J. Daugman, "Iris recognition border-crossing system in the uae," *International Airport Review*, vol. 8, no. 2, 2004.

[9] G. Kaur and C. Verma, "Comparative analysis of biometric modalities," *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 4, no. 4, pp. 603–613, 2014.

[10] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 1701–1708.

[11] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi, "Fully automatic face normalization and single sample face recognition in unconstrained environments," *Expert Systems with Applications*, vol. 47, pp. 23–34, 2016.

[12] S. Umer, B. C. Dhara, and B. Chanda, "Iris recognition using multiscale morphologic features," *Pattern Recognition Letters*, vol. 65, pp. 67–74, 2015.

[13] M. Eskandari and Ö. Toygar, "Selection of optimized features and weights on face-iris fusion using distance images," *Computer Vision and Image Understanding*, vol. 137, pp. 63–75, 2015.

[14] C. Sagonas, E. Antonakos, G. Tzimiropoulos, S. Zafeiriou, and M. Pantic, "300 faces in-the-wild challenge: Database and results," *Image and Vision Computing*, vol. 47, pp. 3–18, 2016.

[15] D. Brown and K. Bradshaw, "An investigation of face and fingerprint feature-fusion guidelines," in *Beyond Databases, Architectures and Structures. Advanced Technologies for Data Mining and Knowledge Discovery*. Springer, 2015, pp. 585–599.

[16] ——, "A multi-biometric feature-fusion framework for improved unimodal and multi-modal human identification," in *Technologies for Homeland Security (HST), 2016 IEEE Symposium on*. IEEE, 2016, pp. 1–6.

[17] ——, "Extended feature-fusion guidelines to improve image-based multi-modal biometrics," in *Proceedings of the Annual Conference of the South African Institute of Computer Scientists and Information Technologists*. ACM, 2016, p. 7.

[18] P. N. Belhumeur, J. P. Hespanha, and D. Kriegman, "Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 711–720, 1997.

[19] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR 2001*, vol. 1. IEEE, 2001, pp. I–511.

[20] P. J. Phillips, H. Wechsler, J. Huang, and P. J. Rauss, "The feret database and evaluation procedure for face-recognition algorithms," *Image and vision computing*, vol. 16, no. 5, pp. 295–306, 1998.

[21] A. Kumar and A. Passi, "Comparison and combination of iris matchers for reliable personal authentication," *Pattern recognition*, vol. 43, no. 3, pp. 1016–1026, 2010.

Dane Brown is currently pursuing his PhD. in Computer Science at Rhodes University. The focus of his research is computer vision, information security and biometrics.

Karen Bradshaw obtained her PhD from Cambridge University and is an Associate Professor at Rhodes University, Department of Computer Science. Her research interests are computer vision, distributed computing and robotics frameworks.