Advancing Active Authentication for User Privacy and Revocability with BioCapsule

Edwin Sanchez*
Anthony Weyer*
Joseph Palackal
Kai Wang
Tyler Phillips
Xukai Zou
edwsanch@iu.edu
antweyer@iu.edu
joppakool@gmail.com
kwang@georgiasouthern.edu
tyler.phillips@alumni.iu.edu
xzou@iupui.edu
Indiana University-Purdue University Indianapolis
Indianapolis, Indiana, USA

Abstract

Biometric Facial Authentication has become a widely used mode of authentication in recent years, which can be attributed to the ever-growing popularity of mobile devices. With this growth in popularity, there is also a growth in concern over privacy for biometrics. Along with the issue of template revocability with biometric data, there is a need for a system that can provide for these issues while remaining easy to use and practical. BioCapsule is a system designed to solve these issues. While BioCapsule has been tested for its face authentication capabilities, this paper extends the scheme to Active Authentication, where a user is continuously authenticated throughout a session on a mobile device. The MOBIO dataset is used for testing, which contains video recordings of 150 individuals using mobile devices over several sessions. We find that the BioCapsule system not only performs comparably to the baseline system performance, but in some cases exceeds baseline performance in terms of False Acceptance Rate, False Rejection Rate, and Equal Error Rate. We examine these findings to learn about both the nature of the Active Authentication task and how BioCapsule interacts with this system. We also examine hyperparameters such as time interval for sampling user facial features, and window size, referring to how many past samples to average over with the current sample to determine user authenticity. With this, we determine BioCapsule to be a powerful and practical addition to improve system security for face-based Active Authentication systems.

Keywords: Continuous Authentication, Active Authentication, Deep Neural Networks, Face Authentication, Biometrics, Mobile Authentication

ACM Reference Format:
Edwin Sanchez, Anthony Weyer, Joseph Palackal, Kai Wang, Tyler Phillips, and Xukai Zou. 2023. Advancing Active Authentication for User Privacy and Revocability with BioCapsule. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation emai (Conference acronym 'XX).* ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

In recent years, biometrics has proven to be a useful method of authenticating users [6]. The reasons are three-fold. Sensors capable of sampling biometrics are now in the hands of many users as mobile devices and come in the form of cameras, microphones, and touch screens. Biometrics are also less obtrusive in a work environment, where there is less of a burden on the user to remember a password or maintain a physical token on their person. The final reason is the introduction and proliferation of Deep Learning, which takes away the need for system engineers to manually extract important features from sampled biometrics and replaces it with a statistically trained model that identifies key patterns in the data [19], [16], [4].

*Both authors contributed equally to this research.

Conference acronym 'XX, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

While the main hindrances to the use of biometrics have been addressed, new limitations have taken their place in the form of *user privacy* and *revocability*. As more user devices come with biometric sensors and most notably face identification and recognition software pre-installed, some have begun to push back against these systems due to fears of societal implications [2], [11]. Another key issue is revocability. If biometric data is compromised, you cannot simply change the biometrics as you might in password-based or physical token-based system, as biometrics are drawn from physiological details of an individual person. This makes some biometric authentication schemes extremely rigid and a risk to implement as a primary authentication method [14]. These are the issues that state-of-the-art biometric authentication systems aim to resolve.

This paper aims to apply a biometric authentication scheme, called BioCapsule [12], [13], to the domain of facial recognition and authentication. More specifically, we measure the performance of the scheme in the context of Active Authentication (AA), or continuously authenticating the user throughout the user's active use of a resource during a session. This category of authentication provides increased security in comparison to single authentication at the beginning of a session, helping to defend against post-login attacks. This paper also works to identify how the BioCapsule scheme affects the feature embeddings of extracted facial features during the authentication process.

2 Related Work

Much of the recent previous work done with regard to biometric authentication focuses on the two concerns of *privacy* and *revocability*. The main categories of state-of-the-art schemes that have emerged with regard to these issues have been Biometric Cryptosystems (BCS) and Cancelable Biometrics (CB). While these categories focus on securing biometrics in a more general sense, some other schemes have been developed that focus more specifically on Continuous Authentication.

2.1 Biometric Cryptosystems and Cancelable Biometrics

BCS schemes generate keys from sampled biometrics [18], [10]. These generated keys can then be used to authenticate the user into the system. CB schemes attempt to apply a set of transformations to given biometric features in a secure manner [15]. However, there remain concerns with these schemes. BCS schemes are brittle, generating wildly different keys from only minor changes in sampled biometrics and requiring a stabilizing mechanism to be usable [18], [7]. CB schemes are susceptible to the same problems, while also having issues where the user's biometrics can be recovered from the original template. Additionally these schemes can reduce system performance [18], [7].

Another glaring issue comes down to the method of implementing these schemes. The schemes often rely on the entire system being designed with the scheme in mind from the start. This is an unattractive prospect for system designers as it may require recreating their existing system from the ground up, resulting in less flexibility and more effort when implementing these systems.

2.2 Continuous Authentication

Several schemes have been developed to handle multi-modal biometric authentication for continuous authentication. These systems may use a blend of multiple biometrics, such as face and voice recognition, as well as gyroscopic information or screen touch patterns [9], [3]. However, these schemes can put excessive strain on the device's battery due to the processing power needed. The demand of these schemes can be lessened by widening the time intervals between authentication checks. However, this comes with the risk of allowing for an attacker to access sensitive information in the event of a post-login attack.

In 2023, Keykhaie and Pierre [8] propose a face-based continuous authentication system using SIM/eSIM technology to protect biometric templates. Their proposed system uses a combination of modern deep learning face preprocessing and feature extraction models. To fit the biometric templates and authentication process onto the small footprint of a SIM card, their system performs a process called quantization. The preprocessing and feature extraction steps are done on the mobile device, with the authentication check happening on the SIM card. However, quantizing the model has measurable effects on the performance of the system.

3 Overview of BioCapsule Scheme

The BioCapsule scheme is designed to solve the current issues of privacy and revocability, while also maintaining a simplistic and easy to implement structure that is provably secure [17], [12], [**?**]. BioCapsule can be broken down into 3 main steps, performed during enrollment and at authentication time: signature extraction, key generation, and secure fusion.

Before biocapsule can begin, we must have a feature vector representing the user's biometrics. This feature vector must be of constant length, but this input length can be adjusted throughout the system to account for different types of biometric features. We must also have features from another subject, called the reference subject (RS). The biometrics generated from this subject is extracted in the same way as the user and will be fused with the user's in order to provide privacy and revocability for the user. Later, if the generated biocapsules are compromised, the authentication system administrator can 'revoke' the biocapsule, or replace the current biocapsules for a user with new ones using a new RS.

After obtaining a user's feature vector, we begin signature extraction. This process involves taking a feature vector and reshaping it to be two dimensional (a matrix). The elements of each row of the matrix get averaged into a single value and are then rescaled, resulting in a signature vector. This process is one-way, as there are no single set of elements that can result in any final element in the signature vector - there are multiple possible solutions to the averaged result per row [17].

The next step is key generation. The subject's signature vector generated from signature extraction is used to generate keys using a pseudo-random number generator (PRNG). A PRNG is seeded with an element from the signature vector, and is then used to generate an array of new numbers with the given seed. This is repeated for every element of the signature vector. This will result in a matrix of the same shape as the reshaped input features from the signature extraction step. The matrix is then reshaped back into a vector, resulting in the same exact shape and size as the original input feature vector, but with psuedo-randomly generated numbers. These elements are them binarized, with positive elements set to 1 and negative elements set to -1. This vector with the same shape and size as the input feature vector is refered to as the key.

This process of signature extraction and key generation is repeated for both the user and the RS selected for the user. To perform the final step referred to as fusion, a user's input feature vector is element-wise multiplied with the RS's generated key, and the RS's input feature vector is element-wise multiplied with the user's generated key. The two output vectors are then element-wise added together, resulting in a biocapsule generated for the given user with their assigned RS.

There are several benefits to this approach. All of the steps (signature extraction, key generation, and secure fusion) are one-way - either the result from the step could have multiple possible inputs, meaning there is no single solution for any given result, or some information is transformed in a way that requires more information than the result itself to recover the input [17]. This approach is also simpler to use, as it can be easily modified for different types of biometrics. While this paper looks at continuous face authentication, the scheme is designed to be applied to any type of features extracted from a user that can be expressed as a vector. Changing vector length is a matter of changing parameters of the biocapsule process. Signature extraction can be modified to address the type of data being authenticated. Since the BioCapsule scheme generates feature vectors of the same size and shape as the input, the system can be added into existing systems with little modification and retraining.

4 Continuous Authentication System

Figure 1 shows an example continuous authentication system, and the basis for testing in this paper. From a high level, the system begins when it is signaled to continuously authenticating a specific user. The system then captures an image from the device's camera. This image gets passed to the preprocessing step, which finds any faces in the image and generates a cropped frame for each detected face. If there are no detected faces, the process ends here and the system locks the screen. If multiple faces are detected, the center most face is selected for the authentication decision. The cropped frame of the detected face is then passed to the feature extractor, which generates a vector representing the face in euclidean space; the feature vector. The next step is to apply biocapsule if applicable, then pass the generated biocapsule to the binary classifier for the authentication step. If biocapsule is not used, then the raw feature vector is passed to the binary classifier. The binary classifier generates a probability for whether the cropped face should be authenticated. If the generated probability is above a certain threshold, then the authentication system waits idly for a set amount of time before repeating the process from the image capturing step. If the classifier's generated probability is below a certain threshold, the the screen is locked. There can also be an additional step after generating a probability, where probability of the last n frames are stored to be averaged over using a given averaging method.

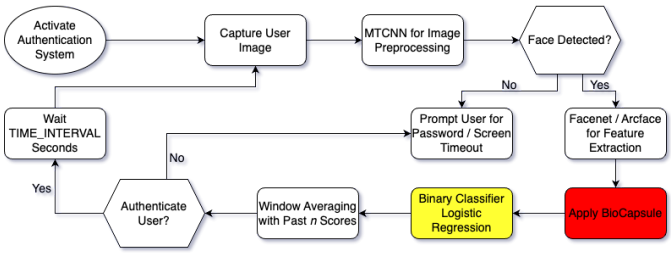

Figure 1. The continuous authentication system used as the basis for testing. The red rectangle represents the biocapsule process, which the system is tested with and without. The yellow rectangle represents the binary classifier, which is affected by whether or not biocapsule is used. The rest of the system does not directly rely upon biocapsule, and so does not need to be adjusted when adding biocapsule into the system.

For enrollment, the binary classifier is the only system that needs to be trained. A set of positive authentication samples are collected from the subject the classifier is trained for, along with samples from other subjects to be negative samples. These positive and negative samples are passed

through the preprocessing and feature extraction steps, resulting in the feature vectors to be given to the binary classifier for training. If biocapsule is used, then a RS is also passed through the preprocessing and feature extraction steps to create its own feature vector. With a subject's feature vector and a RS's feature vector, the biocapsule generation process can be completed. If biocapsule is not used, the system will simply use the positive and negative samples generated from feeding images through the preprocessing and the feature extraction steps.

Since the authentication decision is based on a threshold, the enrollment process which trains the classifier also tunes the threshold used for distinguishing between a positive authentication and a negative authentication. A fraction of the training data is used for threshold tuning, where the system aims to tune the threshold to maximize or minimize a certain target metric.

For testing purposes, the lock screen is ignored. The probabilities and authentication decision are recorded. Testing also does not capture from an actual camera, and instead pulls from a preprocessed dataset for continuous authentication.

The following subsections describe the dataset used for testing, as well as gives details about the preprocessing, feature extraction, biocapsule, window averaging, and wait time steps in the system.

4.1 Preprocessing

The preprocessing step is the first step in the continuous authentication process. Here, preprocessing refers to taking a raw input image and finding all of the faces that are in the image. At this stage the system is not concerned with the identity of the faces, only with finding all faces within a given image. The preprocessing step also detects key points with respect to all of the faces, finding both eyes, the nose, and the left and right corners of the mouth. The model used to perform preprocessing in the system is MTCNN [19].

Since MTCNN will generate cropped faces for each face found in the image, we need to decide what to do when there's multiple faces or no faces at all. If there are no faces, the authentication process ends here and the system should lock the device. If there are multiple faces, the authentication system selects the center-most face, making the assumption that the user of the device is the most likely to be centered in the sampled image. A less strict system may wish to leave the screen unlocked if no faces are detected, and a stricter system may wish to lock the screen if more than one face is detected.

4.2 Feature Extraction

The feature extraction process is the second step in the AA process. This step takes the input of the previous step, preprocessing, and generates feature vectors from the input images. These features can then be used for the authentication decision. The feature vectors generated should contain important features from the cropped face image given, with similar feature vectors generated for the same face, but different feature vectors generated for different faces. The feature extraction step uses one of two state-of-the-art face feature extraction models: FaceNet [16] and ArcFace [4]. Testing results with each feature extraction model are given in the Experiments section.

4.3 BioCapsule Generation

The next step in the process is to generate a biocapsule using the biocapsule scheme, as described in the section `Overview of BioCapsule`. If using the baseline system architecture, this biocapsule step can be skipped. The feature vector generated from the feature extraction step can be directly passed to the binary classifier for either enrollment training or AA if the classifier is already trained. If the biocapsule scheme is used, then a RS is needed to generate resulting biocapsules. As described in [17], [12], [**?**], a single RS can be used with all enrolled users, or a new RS can be selected for each user. During testing, faces from the Labeled Faces in the Wild (LFW) [5] dataset are used as RSs.

4.4 Binary Classifier

The binary classifier is the next step in the process, in charge of generating a probability for a given sample. During the enrollment process, the classifier is trained on feature vectors from the feature extraction process. If biocapsule is used, then the classifier is trained on the generated biocapsules instead. Positive samples refer to feature vectors or biocapsules generated for the positive subject. Negative samples refer to feature vectors or biocapsules generated from all other subject samples used in training, in a one-vs-all method. Also during enrollment, a fraction of the samples are used for threshold tuning. The threshold tuning finds a threshold that maximizes or minimizes the system's performance with respect to the samples given, targeting a performance metric.

4.5 Window Averaging

The next step is an additional step used to potentially improve the performance of the overall system, balancing intrusiveness with security. Window averaging refers to taking the past n probabilities generated by the classifier and averaging them using a selected averaging algorithm. This averaged score is then used for the authentication decision rather than a single probability generated from the currently sampled features.

With a bigger window size, a single sampled frame has a fraction of the influence over the overall authentication decision made. If there is a single bad frame, such as when a user turns away from their camera for a moment, averaging the score over many past probabilities may keep the authentication system from locking the screen unnecessarily. But this comes with significant security risks, as a post-login

attacker may have access to the device for longer before the system denies access.

4.6 Authentication Decision

The final stop in the AA process is the authentication decision, using the tuned threshold. If the generated probability from the classifier is less than a threshold value, the frame is considered to have failed the authentication challenge and the system is locked. If the probability is above the threshold, then the user has passed the authentication challenge and the system does nothing. As long as the user passes the authentication check, the user will not be bothered.

Here, the authentication system is set to idle until it is time to authenticate the user again. While one could set the system to run as fast as the camera can capture frames, this may result in a system that constantly denies the user access to the device for natural motions that move the face out of frame or in an unrecognizable orientation (ex. stretching, leaning over, looking away from the camera, rubbing one's face, etc.). To combat this, the system will only authenticate at a specified rate of time, selected to strike a balance between obtrusiveness and security. Similar to window averaging, the longer the time between authentications, the more likely an attacker can gain access to the system for long enough to a security breach. While a shorter time between authentications provides stronger security, constant authentications can result in an obtrusive system, resulting in lower productivity for the user.

From here, the system begins again, continuously authenticating the user by repeating the above steps. The system stops when the authentication check is failed or the system is no longer required.

5 Experiment

The following subsections describe notable experiment setup details and the results found during testing. The tests aim to answer the following questions: (1) Does the plain system perform reasonably during AA? and (2) Does the system with BioCapsule applied perform comparably during AA? Code for the replicating this experiment can be found at https://github.com/Edwin-Sanchez2003/BioCapsule.

5.1 The MOBIO Dataset

The MOBIO dataset is a dataset designed for bi-modal continuous authentication. The dataset [9] includes videos of 150 different subjects over a 2-year time span. The videos were taken in 5 different countries and 6 different locations in total. The devices used in this dataset for recording were a 2008 MacBook and a Nokia phone. There were 12 sessions recorded in total. The first session was recorded on both laptop and phone. The remaining 11 sessions were recorded on phone only. Each session has 21 videos, where a subject speaks directly into the camera of the mobile device or laptop, reading from a prompt given to the subject. The tests in this paper ignore the audio data and focus solely on the videos.

Figure 2. Sample images of subjects from the MOBIO dataset.

For our purposes, we attempt to mimic the testing setup in [8] with the MOBIO dataset. The first session of each subject is reserved for training each subject's respective binary classifier. The remaining sessions are used for testing. To run a test on a single subject, the single subject's sessions are regarded as positive samples, while the remaining subject's sessions are used as negative samples. Testing was done with two different settings with respect to the MOBIO dataset: single platform and cross platform. Single platform refers to training on samples extracted from the same device as testing is run on. This simulates the scenario where the classifier is trained using the same camera that is used during AA. Cross platform refers to training on samples extracted using a different camera during training than the camera used during AA. This simulates the scenario where a user's data is sampled for training the system using a different camera than what the user would normally use for AA.

Subjects *f-210* and *f-218* were removed due to an insufficient number of sessions. The former has only the laptop session, while the latter has all session but the laptop session.

5.2 Metrics

The metrics used for testing are False Acceptance Rate (FAR) (equation 1), False Rejection Rate (FRR) (equation2), and Equal Error Rate (EER). FAR is a ratio measuring the number of false positives (a person passes the authentication check who should not have) that were let in by the system with respect to the total number of negative samples tested. The total negative samples are expressed in the equation as $FP + TN$. FRR is a ratio measuring the number of false negatives (a person who does not pass the authentication check but should have) that the system failed to let in with respect to the total number of positive samples tested. The total positive samples are expressed in the equation as $FN + TP$. The EER represents where these two metrics are equal in a system. This can be found by generating the probabilities for a set of data samples, then sliding the threshold to where these two rates are equal.

$$FAR(\%) = (FP/(FP + TN)) * 100 \quad (1)$$

$$FRR(\%) = (FN/(FN + TP)) * 100 \quad (2)$$

These metrics allow us to measure the systems performance with respect to the questions "What rate does the system allow unauthorized users in?" (FAR) and "What rate does the system keep authorized users out?" (FRR). The lower the FAR and FRR scores, the better the system performance. EER is a way to find the most optimistic performance of the model when deployed in the wild.

5.3 Results

Table 1 shows the performance of the system with and without BioCapsule applied, as well as with different feature extraction models, RS settings, and single vs. cross platform training and testing. Across the board, cross platform EER performance is worse when compared to their single platform counterparts. This can be attributed to the classifier fitting to the camera's image quality on one device during training, and then under-performing when being showed samples from another camera's images. When comparing *No BC* settings to *BC with Single RS*, we see performance slight performance degradation, although with still comparable performance. The degradation can be attributed to the BioCapsule process losing some information when the fusion process happens, as the BioCapsule scheme takes in two feature vectors of equal length and outputs one. Additionally, with single RS, the transformations applied to the user's features are the same, resulting in lower inter-class separation as the RS features are weighted equally with the user's [17]. As seen with the results using ArcFace for the feature extractor, the performance degradation is negligible. Table 2 shows the performance of the system when compared to the SIM card based systems designed in [8], further proving the system's effectiveness when compared to other template security systems.

Surprisingly we see that the performance of systems set with *Multi RS* outperform both *Single RS* systems and the *No BC* baseline systems. This, again, can be attributed to BioCapsule's fusion process. While using a single RS has the effect of reducing inter-class variation by fusing every subject with the same RS, multi RS has the opposite affect as each new reference subjects applies different transformations on the user's feature vector, increasing inter-class variation. In data with low inter-class variation, BioCapsule can artificially increase inter-class variation by selecting diverse RSs for each user, as the variation in RS will be reflected in the output biocapsule. Figure 5 demonstrates this phenominon images of similar looking people from the VGG2 dataset [1]. Figure 5 also shows that BioCapsule also reduces the intra-class variation of the generated feature vectors compared to the original feature vectors. This again due to the RS applying a constant transformation on the input feature vectors, smoothing out high variation in input images.

Model Type	BC	RS	Platform	FAR (%)	FRR (%)	EER (Test) (%)
ArcFace	No BC	N/A	Single	0.419	0.873	0.010
			Single	0.527	0.865	0.020
	BC	Single	Single	0.440	0.892	0.021
			Multi	0.361	0.920	0.031
		Multi	Single	0.458	0.854	0.007
			Multi	0.587	0.862	0.013
FaceNet	No BC	N/A	Single	0.393	6.349	0.863
			Multi	0.396	9.521	1.332
	BC	Single	Single	0.365	13.106	1.960
			Multi	0.423	17.902	2.549
		Multi	Single	0.488	1.140	0.086
			Multi	0.532	1.715	0.180

Table 1. The performance of the system with different settings, where each score represents the mean performance of the system averaged across all subjects in MOBIO. Test EER is found by using the probabilities predicted on the test set and moving the threshold for authentication to where the FAR and FRR scores are equal +=0.01%

When comparing the baseline feature vectors (the top two graphs of figure 5) with both feature vectors generated using BC with both single and multi RS (the bottom four graphs), we see that the clusters for each person end up in a new location in the space, implying that BC performed some sort of transformation on the images, as expected. However, we can also see the clusters tighten when biocapsule is applied in both the single and multi RS variations (lower intra-class variation). We can further see that multi RS further separates clusters (higher inter-class variation) while single RS brings them closer together (lower inter-class variation).

Model Type	Use BC	Ref. Subj.	Platform	FAR (%)	FRR (%)	EER (TEST) (%)
Arcface	No BC	N/A	Single	0.419	0.873	0.010
			Multi	0.527	0.865	0.020
	BC	Single RS	Single	0.440	0.892	0.021
			Multi	**0.361**	**0.920**	**0.031**
		Multi RS	Single	0.458	**0.854**	**0.007**
			Multi	**0.587**	0.862	0.013
Facenet	No BC	N/A	Single	0.393	6.349	0.863
			Multi	0.396	9.521	1.332
	BC	Single RS	Single	**0.365**	13.106	1.960
			Multi	0.423	**17.902**	**2.549**
		Multi RS	Single	0.488	**1.140**	**0.086**
			Multi	**0.532**	1.715	0.180

Figure 3. Use new table

6 Future Work

While we have struck at the main question of how viable biocapsule is in the domain of AA with a realistic application of the system, there are more questions that need to be answered. While the MOBIO dataset provides a useful baseline for comparing systems, it lacks the real-world issues that may be found when deploying this system in the wild. Most subjects remain within the camera's field of view, whereas

Site	Architecture	No BC | Single/Multi RS BC	Alg.	Single Platform	Cross Platform
BUT	CA-MMOC	No BC	L-SVM	0.1(0.4)	0.2(0.4)
			LDA	0.3(0.4)	3.5(3.1)
	F-MMOC		LR	0.1(0.4)	0.2(0.4)
			L1	0.1(0.2)	0.1(0.2)
			L2	0.1(0.1)	0.1(0.1)
	D-CSLDA		L_inf	0.8(1.6)	0.9(1.0)
			CSLDA	13.5(4.2)	21.9(5.2)
	ArcFace	Single RS-BC	LR	0.0000876(0.00495)	0.00219(0.0105)
		Multi RS-BC		0.0(0.0)	0.0 (0.0)
	FaceNet	Single RS-BC		1.63(1.42)	2.32(2.33)
		Multi RS-BC		0.136(0.748)	0.195(0.713)
IDIAP	CA-MMOC	No BC	L-SVM	0.0(0.0)	0.5(0.8)
			LDA	0.2(0.2)	11.5(12.3)
	F-MMOC		LR	0.0(0.0)	0.3(0.4)
			L1	0.1(0.1)	2.6(9.1)
			L2	0.1(0.1)	2.4(8.9)
	D-CSLDA		L_inf	0.2(0.1)	2.8(2.1)
			CSLDA	12.8(6.2)	27.1(10.2)
	ArcFace	Single RS-BC	LR	0.0(0.0)	0.00451(0.0230)
		Multi RS-BC		0.0(0.0)	0.0 (0.0)
	FaceNet	Single RS-BC		1.18(1.71)	1.78(2.68)
		Multi RS-BC		0.00161(0.00531)	0.0372(0.120)
LIA	CA-MMOC	No BC	L-SVM	1.4(4.2)	1.5(3.2)
			LDA	1.6(3.0)	2.1(3.2)
	F-MMOC		LR	1.4(3.8)	1.5(3.0)
			L1	1.2(2.5)	1.3(2.3)
			L2	1.1(3.0)	1.2(2.9)
	D-CSLDA		L_inf	1.3(2.6)	1.4(2.6)
			CSLDA	19.1(8.2)	24.7(8.7)
	ArcFace	Single RS-BC	LR	0.0819(0.293)	0.137(0.501)
		Multi RS-BC		0.0335(0.151)	0.0600 (0.218)
	FaceNet	Single RS-BC		2.66(3.15)	2.49(2.27)
		Multi RS-BC		0.103(0.322)	0.128(0.388)
UMAN	CA-MMOC	No BC	L-SVM	0.1(0.1)	1.1(0.2)
			LDA	0.4(0.4)	3.0(2.7)
	F-MMOC		LR	0.1(0.1)	0.1(0.2)
			L1	0.1(0.2)	0.1(0.1)
			L2	0.1(0.1)	0.1(0.1)
	D-CSLDA		L_inf	0.7(1.3)	1.1(1.1)
			CSLDA	16.1(4.6)	23.1(10.2)
	ArcFace	Single RS-BC	LR	0.0295(0.116)	0.00297(0.0391)
		Multi RS-BC		0.0(0.0)	0.0 (0.0)
	FaceNet	Single RS-BC		2.64(4.43)	3.63(6.64)
		Multi RS-BC		0.147(0.614)	0.404(1.754)
UNIS	CA-MMOC	No BC	L-SVM	0.1(0.2)	0.1(0.2)
			LDA	0.4(0.4)	0.3(0.4)
	F-MMOC		LR	0.1(0.2)	0.1(0.2)
			L1	0.3(0.3)	0.5(0.8)
			L2	0.2(0.2)	0.4(0.7)
	D-CSLDA		L_inf	0.7(1.1)	1.0(1.1)
			CSLDA	15.1(7.1)	21.0(8.4)
	ArcFace	Single RS-BC	LR	0.000242(0.00116)	0.000606(0.00291)
		Multi RS-BC		0.0(0.0)	0.0 (0.0)
	FaceNet	Single RS-BC		1.56(3.16)	1.66(2.81)
		Multi RS-BC		0.0352(0.125)	0.0844(0.337)
UOULU	CA-MMOC	No BC	L-SVM	0.1(0.1)	0.8(0.6)
			LDA	0.6(0.4)	9.3(6.5)
	F-MMOC		LR	0.1(0.1)	0.5(0.9)
			L1	0.2(0.1)	7.3(11.1)
			L2	0.1(0.1)	6.8(13.1)
	D-CSLDA		L_inf	0.8(1.6)	0.9(1.0)
			CSLDA	13.5(4.2)	21.9(5.2)
	ArcFace	Single RS-BC	LR	0.0000876(0.00495)	0.00219(0.0105)
		Multi RS-BC		0.0(0.0)	0.0 (0.0)
	FaceNet	Single RS-BC		1.63(1.42)	2.32(2.33)
		Multi RS-BC		0.136(0.748)	0.195(0.713)

Table 2. The performance of the systems designed in [8] compared to ours (BioCapsule schemes). The table shows the performance of the system with respect to subjects at each of the 6 MOBIO locations [9]

.

in the real world a subject may have a variety of movements that momentarily put the user's face out of frame or hide the user's face from the camera. Also, more variety in environments and longer videos are needed to test the robustness of these systems.

7 Conclusion

Biometrics has grown to play an important role in user authentication in recent years, due to the rise of mobile devices. And with this, there is a need for better security around these mobile devices. Active authentication is one way to provide a stronger layer of security. While this works well on its own, continuous authentication using biometrics does not address the privacy concerns of today's users, nor does it deal with the problem of template revocability that hinders the proliferation of biometric authentication.

Biocapsule can solve these issues, while remaining secure and preserving the representational power of the underlying system. Our tests find that biocapsule can be applied to continuous face authentication systems and can perform comparably to these systems, while providing more security benefits for the system and users.

TABLE 1
System's EER in Single Platform and Cross Platform Scenarios With Standard Deviation in Parentheses

				Mean Test EER (std) %	
Site	Architecture	No BC | Single RS BC | Multi RS BC	Classification Alg.	Single Platform	Cross Platform
BUT	CA-MMOC	No BC	L-SVM	0.1 (0.4)	0.2 (0.4)
		No BC	LDA	0.3 (0.4)	3.5 (3.1)
		No BC	LR	0.1 (0.4)	0.2 (0.4)
	F-MMOC	No BC	L1	0.1 (0.2)	0.1 (0.2)
		No BC	L2	0.1 (0.1)	0.1 (0.1)
		No BC	L_inf	0.8 (1.6)	0.9 (1.0)
	D-CSLDA	No BC	CSLDA	13.5 (4.2)	21.9 (5.2)
	ArcFace	Single RS-BC	Logistic Regression	0.0000876 (0.000495)	0.00219 (0.0105)
		Multi RS-BC	Logistic Regression	0.0 (0.0)	0.0 (0.0)
	FaceNet	Single RS-BC	Logistic Regression	1.63 (1.42)	2.32 (2.33)
		Multi RS-BC	Logistic Regression	0.136 (0.748)	0.195 (0.713)
IDIAP	CA-MMOC	No BC	L-SVM	0.0 (0.0)	0.5 (0.8)
		No BC	LDA	0.2 (0.2)	11.5 (12.3)
		No BC	LR	0.0 (0.0)	0.3 (0.4)
	F-MMOC	No BC	L1	0.1 (0.1)	2.6 (9.1)
		No BC	L2	0.1 (0.1)	2.4 (8.9)
		No BC	L_inf	0.2 (0.1)	2.8 (2.1)
	D-CSLDA	No BC	CSLDA	12.8 (6.2)	27.1 (10.2)
	ArcFace	Single RS-BC	Logistic Regression	0.0 (0.0)	0.00451 (0.0230)
		Multi RS-BC	Logistic Regression	0.0 (0.0)	0.0 (0.0)
	FaceNet	Single RS-BC	Logistic Regression	1.18 (1.71)	1.78 (2.68)
		Multi RS-BC	Logistic Regression	0.00161 (0.00531)	0.0372 (0.120)
LIA	CA-MMOC	No BC	L-SVM	1.4 (4.2)	1.5 (3.2)
		No BC	LDA	1.6 (3.0)	2.1 (3.2)
		No BC	LR	1.4 (3.8)	1.5 (3.0)
	F-MMOC	No BC	L1	1.2 (2.5)	1.3 (2.3)
		No BC	L2	1.1 (3.0)	1.2 (2.9)
		No BC	L_inf	1.3 (2.6)	1.4 (2.6)
	D-CSLDA	No BC	CSLDA	19.1 (8.2)	24.7 (8.7)
	ArcFace	Single RS-BC	Logistic Regression	0.0819 (0.293)	0.137 (0.501)
		Multi RS-BC	Logistic Regression	0.0335 (0.151)	0.0600 (0.218)
	FaceNet	Single RS-BC	Logistic Regression	2.66 (3.15)	2.49 (2.27)
		Multi RS-BC	Logistic Regression	0.103 (0.322)	0.128 (0.388)
UMAN	CA-MMOC	No BC	L-SVM	0.1 (0.1)	1.1 (0.2)
		No BC	LDA	0.4 (0.4)	3.0 (2.7)
		No BC	LR	0.1 (0.1)	0.1 (0.2)
	F-MMOC	No BC	L1	0.1 (0.2)	0.1 (0.1)
		No BC	L2	0.1 (0.1)	0.1 (0.1)
		No BC	L_inf	0.7 (1.3)	1.1 (1.1)
	D-CSLDA	No BC	CSLDA	16.1 (4.6)	23.1 (10.2)
	ArcFace	Single RS-BC	Logistic Regression	0.0295 (0.116)	0.00927 (0.0391)
		Multi RS-BC	Logistic Regression	0.0 (0.0)	0.0 (0.0)
	FaceNet	Single RS-BC	Logistic Regression	2.64 (4.43)	3.63 (6.64)
		Multi RS-BC	Logistic Regression	0.147 (0.614)	0.404 (1.754)
UNIS	CA-MMOC	No BC	L-SVM	0.1 (0.2)	0.1 (0.2)
		No BC	LDA	0.4 (0.4)	0.3 (0.4)
		No BC	LR	0.1 (0.2)	0.1 (0.2)
	F-MMOC	No BC	L1	0.3 (0.3)	0.5 (0.8)
		No BC	L2	0.2 (0.2)	0.4 (0.7)
		No BC	L_inf	0.7 (1.1)	1.0 (1.1)
	D-CSLDA	No BC	CSLDA	15.1 (7.1)	21.0 (8.4)
	ArcFace	Single RS-BC	Logistic Regression	0.000242 (0.00116)	0.000606 (0.00291)
		Multi RS-BC	Logistic Regression	0.0 (0.0)	0.0 (0.0)
	FaceNet	Single RS-BC	Logistic Regression	1.56 (3.16)	1.66 (2.81)
		Multi RS-BC	Logistic Regression	0.0352 (0.125)	0.0844 (0.337)
UOULU	CA-MMOC	No BC	L-SVM	0.1 (0.1)	0.8 (0.6)
		No BC	LDA	0.6 (0.4)	9.3 (6.5)
		No BC	LR	0.1 (0.1)	0.5 (0.9)
	F-MMOC	No BC	L1	0.2 (0.1)	7.3 (11.1)
		No BC	L2	0.1 (0.1)	6.8 (13.1)
		No BC	L_inf	0.5 (0.6)	7.1 (10.8)
	D-CSLDA	No BC	CSLDA	22.5 (9.2)	31.1 (11.5)
	ArcFace	Single RS-BC	Logistic Regression	0.0174 (0.510)	0.0393 (0.0880)
		Multi RS-BC	Logistic Regression	0.0125 (0.0515)	0.0235 (0.0970)
	FaceNet	Single RS-BC	Logistic Regression	2.24 (2.59)	3.88 (4.29)
		Multi RS-BC	Logistic Regression	0.0749 (0.252)	0.252 (0.477)

Figure 4. Use new Table

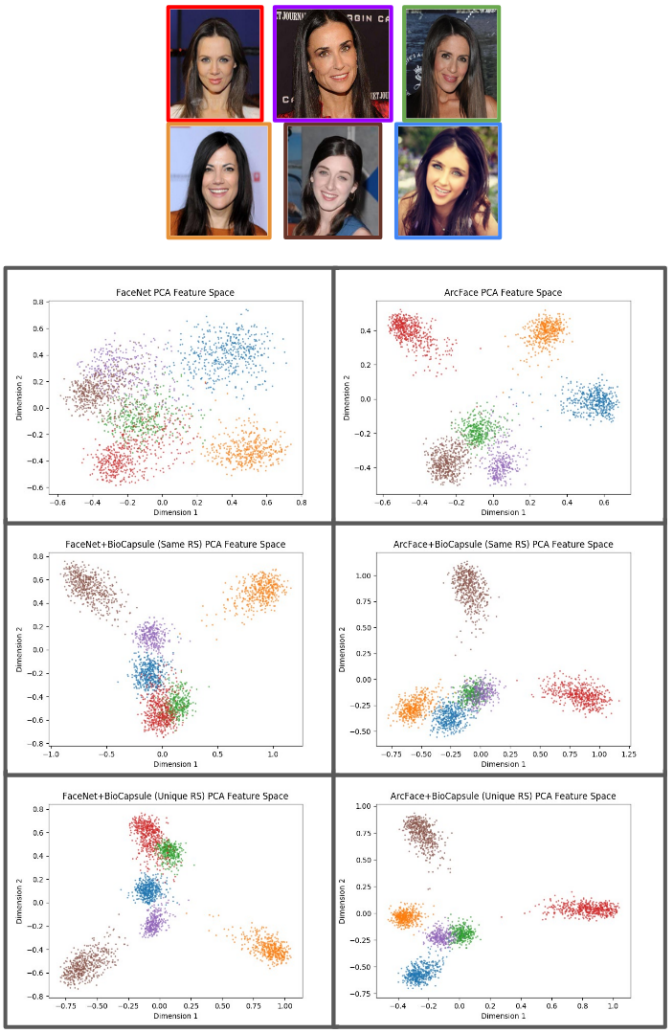

Figure 5. A figure visualizing feature vectors generated from face images, using FaceNet and ArcFace. Feature vectors are reduced from 512 dimensions to 2 for visualization purposes. Each dot corresponds to an image from one of the six people above, with each person distinguished by a different color.

8 Acknowledgements

Our reseach was made possible by the support of IUPUI's REU program (CN-1852105). This research was supported by the National Science Foundation REU program and the National Institute of Health (grants 1R01AR077273-01 and 1R15GM139094-01A1).

References

[1] Cao, Q., Shen, L., Xie, W., Parkhi, O. M., and Zisserman, A. Vggface2: A dataset for recognising faces across pose and age. *CoRR abs/1710.08092* (2017).

[2] Chamary, J. No, apple's face id is not a "secure password", Sep 2017.

[3] Crouse, D., Han, H., Chandra, D., Barbello, B., and Jain, A. K. Continuous authentication of mobile user: Fusion of face image and inertial measurement unit data. In *2015 International Conference on Biometrics (ICB)* (2015), pp. 135–142.

[4] Deng, J., Guo, J., Xue, N., and Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2019).

[5] Huang, G. B., Ramesh, M., Berg, T., and Learned-Miller, E. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Tech. Rep. 07-49, University of Massachusetts, Amherst, October 2007.

[6] Jain, A., Ross, A., and Prabhakar, S. An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology 14*, 1 (2004), 4–20.

[7] Karimovich, G. S., and Turakulovich, K. Z. Biometric cryptosystems: Open issues and challenges. In *2016 International Conference on Information Science and Communications Technologies (ICISCT)* (2016), pp. 1–3.

[8] Keykhaie, S., and Pierre, S. Lightweight and secure face-based active authentication for mobile users. *IEEE Transactions on Mobile Computing* (2021).

[9] Khoury, E., El Shafey, L., McCool, C., Günther, M., and Marcel, S. Bi-modal biometric authentication on mobile phones in challenging conditions. *Image and Vision Computing 32*, 12 (2014), 1147–1160.

[10] Lu, H., Martin, K., Bui, F., Plataniotis, K. N., and Hatzinakos, D. Face recognition with biometric encryption for privacy-enhancing self-exclusion. In *2009 16th International Conference on Digital Signal Processing* (2009), pp. 1–8.

[11] Mozur, P. Inside china's dystopian dreams: A.i., shame and lots of cameras, Jul 2018.

[12] Phillips, T., Yu, X., Haakenson, B., Goyal, S., Zou, X., Purkayastha, S., and Wu, H. Authn-authz: Integrated, user-friendly and privacy-preserving authentication and authorization. In *2020 Second IEEE International Conference on Trust, Privacy and Security in Intelligent Systems and Applications (TPS-ISA)* (2020), pp. 189–198.

[13] Phillips, T., Zou, X., Li, F., and Li, N. Enhancing biometric-capsule-based authentication and facial recognition via deep learning. In *Proceedings of the 24th ACM Symposium on Access Control Models and Technologies* (New York, NY, USA, 2019), SACMAT '19, Association for Computing Machinery, p. 141–146.

[14] Prabhakar, S., Pankanti, S., and Jain, A. Biometric recognition: security and privacy concerns. *IEEE Security Privacy 1*, 2 (2003), 33–42.

[15] Savvides, M., Vijaya Kumar, B., and Khosla, P. Cancelable biometric filters for face recognition. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.* (2004), vol. 3, pp. 922–925 Vol.3.

[16] Schroff, F., Kalenichenko, D., and Philbin, J. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (June 2015).

[17] Sui, Y., Zou, X., Du, E. Y., and Li, F. Design and analysis of a highly user-friendly, secure, privacy-preserving, and revocable authentication method. *IEEE Transactions on Computers 63*, 4 (2014), 902–916.

[18] Uludag, U., Pankanti, S., Prabhakar, S., and Jain, A. Biometric cryptosystems: issues and challenges. *Proceedings of the IEEE 92*, 6 (2004), 948–960.

[19] Zhang, K., Zhang, Z., Li, Z., and Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters 23*, 10 (2016), 1499–1503.

Received 20 February 2007; revised 12 March 2009; accepted 5 June 2009