Biometrics Fusion with Applications in Passenger Re-authentication for Automated Border Control Systems

Hoang (Mark) Nguyen[1], Ajita Rattani[2] and Reza Derakhshani[1]
[1]Department of Computer Science and Electrical Engineering
University of Missouri at Kansas City
Email: hdnf39@mail.umkc.edu
[2]Department of Electrical Engineering and Computer Science
Wichita State University
Wichita, Kansas
Email: ajita.rattani@wichita.edu

***Abstract*—Over 4.1 billion aircraft passengers flew in 2017. This number is expected to be nearly double by 2037. Due to the significant growth in airline services and passenger traffic, automated border control (ABC) systems have been installed to control the air traffic flow in an automatic manner while maintaining the security. However, there are multiple security checkpoints from airport entry to boarding the flight which requires passenger's re-authentication multiple times. We propose a deep representation based learning to combine face and soft-biometrics. The proposed model has applications in automated border control to further ease the traffic flow while maintaining high security at various checkpoints. Using a deep learning-based feature fusion framework, our method obtains 99.02% genuine match rate (GMR) at false match rate (FMR) = 10^{-5} using ResNet-18 model.**

I. Introduction

It is reported that[1], from 2018 to 2037, the number of aircraft passengers will grow up 4.7% worldwide annually, and in some regions, the increase is expected to reach 7.8%. In countries like China, air travel traffic is increasing faster than airport facility capabilities. On average, the border guards only have 12 seconds to verify a passenger's identity to authorize border crossing [1]. Thus, to manage airport traffic flow while maintaining high security level, automated border control (ABC) systems are gaining traction.

Automated border control (ABC) can be defined as an automated system to verify passenger identity without human involvement [2], [3]. Recently, a number of airports are allowing U.S and Canadian citizens to use a mobile passport app to expedite their entry into the United States, expanding ABC's reach through mobile platforms[2]. European Union (EU) project FastPass[3] has established and demonstrated a harmonized and modular approach for ABC gates for Schengen external borders with a focus on convenience, speed, and security.

[1]http://www.boeing.com/commercial/market/commercial-market-outlook/
[2]https://www.airsidemobile.com/
[3]https://www.fastpass-project.eu/

The basic steps of verification process at the ABC e-gate systems are shown in Figure 1. First, the e-gate reads the digital data from the identity document (ID). Second, passenger's biometric samples are captured. Third, the system compares the live captured data with the data in the identity document. After a successful match, ABC system can perform re-authentication for the rest of the security checkpoints, mitigating the need for repeated passenger identity document checks [2].

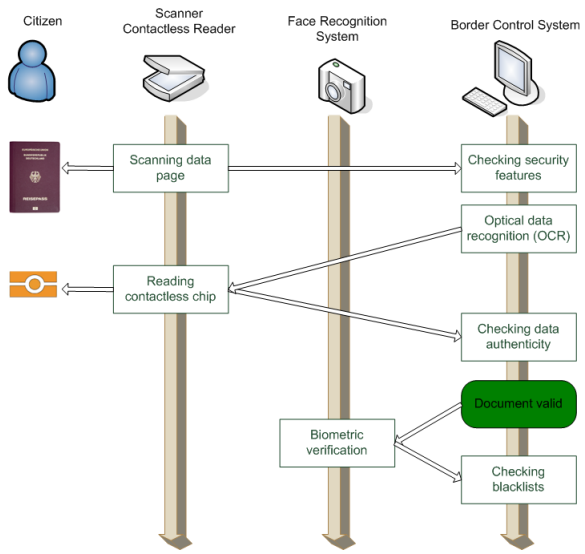

Fig. 1. Steps involved in a typical Automated Border Control e-gate System.

Face and fingerprint recognition techniques are common in ABC e-gates. Some systems may use iris biometrics as well [4]. Three types of automated border crossing procedures are possible [2], [5]: (i) one step process; (ii) integrated two-step process; (iii) segregated two-step process. In a one-step process, the document, the identity, and the authorization are verified at the same time inside the e-gate. In an integrated

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

two-step process, the validity of the document is checked before letting the traveler go inside the e-gate for identity verification and travel authorization. In a segregated two-step process, the validity of the document and the travel authorization can be checked at a different time and place during a border crossing. In order to perform the required steps, four subsystems are involved: (i) the Document Authentication System (DAS), which checks the validity of the document; (ii) the Biometric Verification System (BVS), which captures live biometric samples and compares them with the ones contained in the document; (iii) the Central Systems Interface (CSI), which handles communication with external systems; (iv) and the Border Guard Maintenance System (BGMS), which is used by the officers to monitor the ABC system [6].

In recent years, an increasing number of companies have started providing e-gate services, often in combination with e-passport and biometric visa technologies. In such systems, each passenger's transaction is submitted to a monitoring system that provides timely and complete information on passengers identity and transaction status. The complex software framework is also able to detect identity fraud, extract biometric and biographic information, perform real-time checks of intelligence and criminal databases, and instantly alert border control officers in case of anomalies/emergencies [6]. Such solutions should not only be fast and reliable for airports and border authorities, but also user friendly for travelers. *However, the state-of-the-art in ABC systems is yet to achieve the aforesaid ideals with only handful of papers published* [2]. There are still *open challenges* which include robust anti-spoofing techniques, compatibility between systems by adopting a common biometric data format, and a means for passenger re-authentication to facilitate smooth transition at various checkpoints.

The aim of this paper is to propose a passenger re-authentication method to increase the convenience and security requirements at various checkpoints of an ABC system. Figure 2 shows an overview of our proposed ABC system for airport security consisting of first-level verification based on matching passenger's face image with his or her identity document, followed by passenger's re-authentication stage. During the first checkpoint verification with face biometrics, the system also captures passenger's clothing information as soft-biometrics. After the initial successful authentication, the captured face image along with clothing information is stored and used for re-authentication at subsequent checkpoints allowing for more casual and convenient user interaction, speed and security. The contributions of this work are as follows:

1) Development of a multi-modal convolutional neural network [7] that combines face and clothing based soft-biometrics information at feature level.
2) A method for faster, more convenient, and yet secure face-based passenger re-authentication at various security checkpoints by adding clothing features as a short-term second factor. The proposed system also mitigates the need of passenger's document identity check at various checkpoints.

The rest of the paper is organized as follows. In section 2, we discuss related work in ABC systems and the use of face and clothing information for user re-authentication. In section 3, we describe our proposed method for passenger re-authentication. Section 4 provides our experimental validations. Conclusions and future work are given in section 5.

II. Related Work

Ruggero et al. [6], provided a comprehensive survey of biometric techniques and systems that enable automatic identity verification for ABC systems. In [8], J. Snchez del Ro discussed an ABC e-gate system that used a face recognition algorithm along with its performance and image quality metrics recommended by International Civil Aviation Organization[4]. The proposed e-gate system reportedly achieved a False Rejection Rate (FRR) below 5% at 0.1% False Acceptance Rate (FAR).

Existing literature also includes several face-based user re-authentication methods. In [9], authors introduced a face biometrics system for continuous user authentication in mobile devices. The proposed method used fast fourier transform for template matching and face tracking. Mahbub et al. in [10] investigated a continuous authentication system using partial faces. The proposed method attempted to detect 14 facial segments using cascade AdaBoost classifier trained using local binary patterns features.

Multiple research studies have exploited clothing information for person re-identification [11], [12] in surveillance systems where hard biometric traits such as face and fingerprint are not readily available. In [13], clothing attributes were used to enhance the performance of other soft biometric traits such as human gait [14]. In [15], Nguyen et al. evaluated user re-authentication using clothing information in conjunction with selfie face images captured using mobile devices [16], [17]. The reported best result of 0.032 EER was obtained using feature level fusion of hand-crafted and VGG-16 deep features.

III. Proposed Method for Re-authentication

Our proposed method include face and clothing segmentation, feature extraction, deep feature fusion and matching. Next, we discuss these steps in detail.

A. Face and Clothing Segmentation

Face and clothing information were segmented by two different modules. The face region was segmented by Dlib [18], an open source face analysis library. We used Dlib version 19.16 that uses convolutional neural network (CNN) and histogram of oriented gradients (HOG) for face detection. Because there was no viable and publicly available clothing detection and segmentation tool at the time of this study, we develop our own clothing segmentation algorithm using U-Net, a deep learning model for image segmentation.

U-Net [19] is a convolutional neural network that was originally developed for biomedical image segmentation. It is an encoder-decoder CNN which contains contracting and

[4]https://www.icao.int/Pages/default.aspx/

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

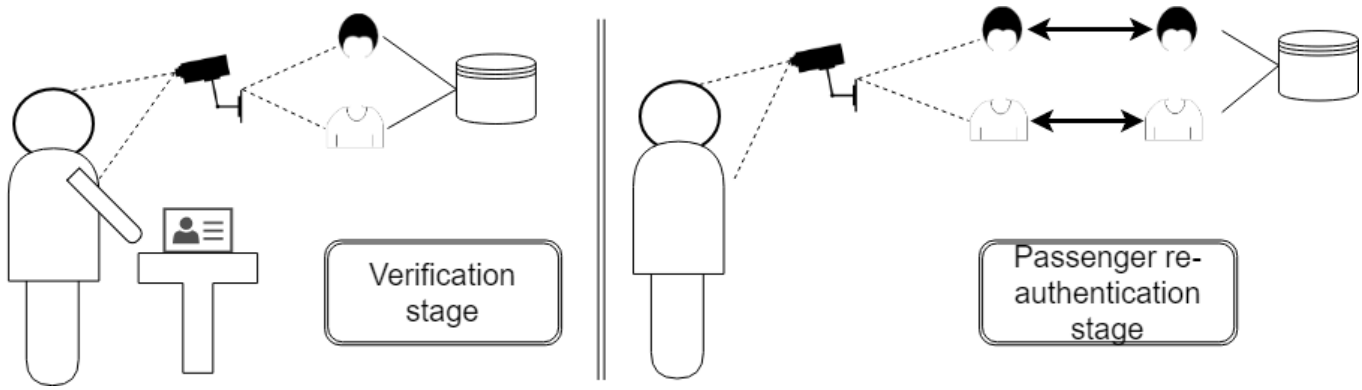

Fig. 2. Schema of the proposed model for passenger re-authentication at various security checkpoints in an ABC system.

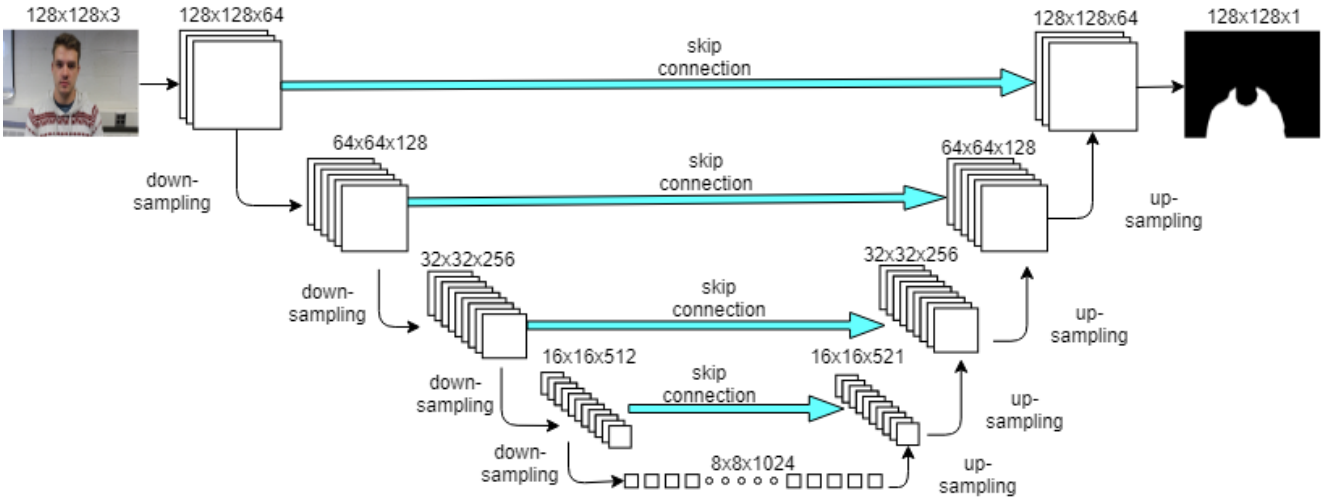

Fig. 3. Architecture of U-Net model used for clothing mask generation.

expansive paths. The encoder is a repeated application of two 3×3 convolutions, followed by rectified linear units (ReLU), and 2×2 max pooling operation. Similarly, each decoder layer consists of up-sampling using 2×2 up-convolution, a concatenation of corresponding feature maps from the contracting path, and two 3×3 convolutions followed by ReLUs. Figure III shows the U-Net architecture used in this paper.

About $1,100$ images, 700 from an in-house dataset and 400 from SiW dataset, were used to train the U-net model for clothing segmentation. Ground truth binary masks were obtained by manual pixel-wise labelling of clothing region using MATLAB. The data then was heavily augmented to obtain $39,000$ clothes images for image segmentation task. 80% of samples were chosen randomly for training and the rest were left for validation. Training images along with the binary masks were used to train the model for 200 epochs based on the performance on validation set. The model obtained 94% precision and recall on a separate test set (100 images) by comparing the ground truth masks with the clothing masks output by the U-net model. The ground truth masks and the binary masks output by the U-net model were compared using intersection over union of pixels between them. Figure 4 shows examples of (a) original images, (b) segmented face using Dlib [18], and (c) segmented clothing using U-Net model.

B. Deep Feature Extraction

An ABC system should have high throughput. Thus, we decided to adopt lightCNN [20] and Resnet-18 [21] for both clothing and facial feature extraction. Different from commonly used deep learning models such as VGG-16 (138M parameters) [22] and AlexNet (60M parameters) [23], our lightCNN and Resnet-18 models contain only 5.6M and 11.5M parameters, respectively. We modified the models to work with RGB images. The input image size is 128x128 for lightCNN and 224x224 for Resnet18. For the lightCNN model, the image was first resized to 144x144 and randomly cropped to 128x128 to fit the model input and to increase the data size. For Resnet-18, the image was resized to 256x256 and randomly cropped to 224x224. Lastly, a random horizontal flip was applied to augment the dataset. During matching, images were resized to 128x128 or 224x224 depending on the model.

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

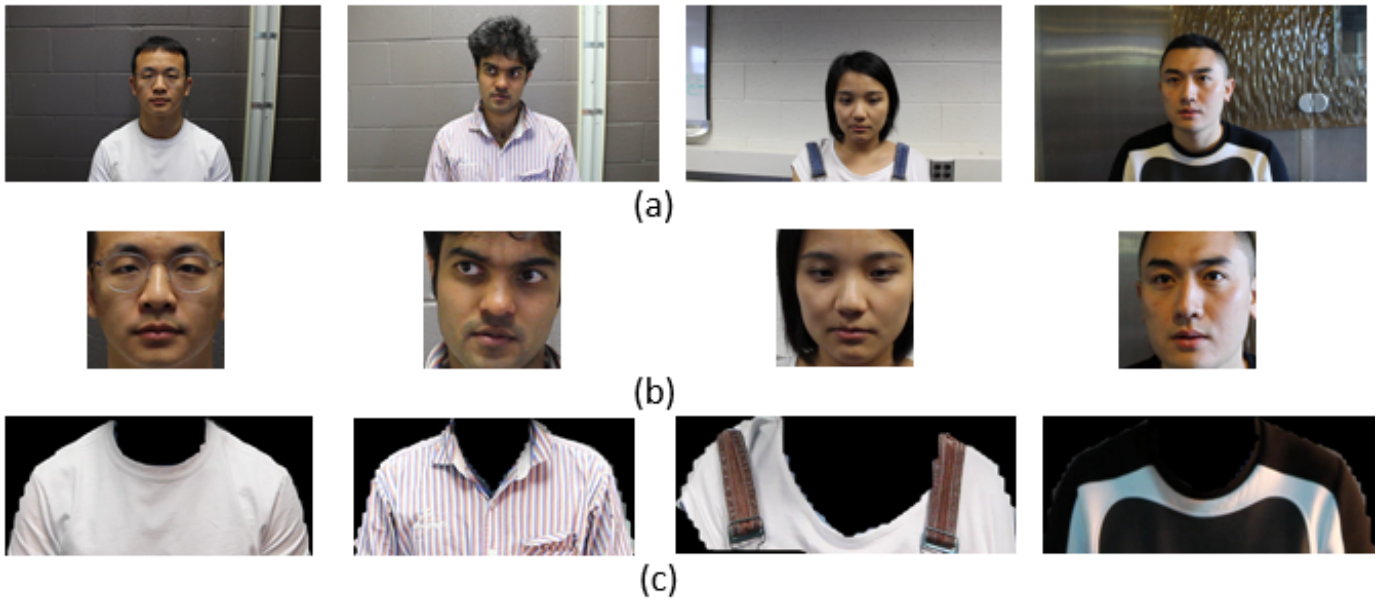

Fig. 4. Example of (a) original images, (b) detected faces, (c) and segmented clothes obtained by U-Net model.

LightCNN is a lightweight CNN structure designed for noisy labeled data. The architecture used in this work is a 9-layer version of lightCNN with 4 convolutional layers, 4 maxout units, and two fully connected layers. The network employs the Max-Feature-Map (MFM) function instead of general ReLU activation function in order to produce more compact and discriminant features. The MFN activation function is defined in equation 1. Regarding the last two layers, the first layer is a fully connected layer that generates a 256 dimensional feature vector, while the second is a softmax layer. Our lightCNN architecture is shown in Table I. The total number of parameters is 5.5M.

$$f_{ij}^{k} = \max_{1 \leq k \leq n} (C_{ij}^{k}, C_{ij}^{k+n}) \tag{1}$$

Residual networks or ResNets [21] employ a residual block to help the model to extract much deeper features without having the vanishing gradient problem. The block is defined in 2. x, y, and $F(x, W_i)$ are input, output, and learning residual mapping, respectively. Passing the input value x to the output y preserves the information which is potentially lost in convolution operation. We used ResNet-18, the lightest of the ResNet models, and trained it from scratch using our clothing and face dataset. One fully connected layer with feature-length of 256 and one softmax layer was added to the end of the model. The total number of parameters in ResNet-18 is 11.5M, which is twice that of a 9-layer lightCNN. Yet, it is still considered to be a relatively small model compared to models such as AlexNet or VGG-16.

$$y = F(x, W_i) + x \tag{2}$$

The feature dimensionality for both 9-layer lightCNN and ResNet18 is 256, which is significantly lower than AlexNet (4016) or VGG-16 (4016). To train each model, the initial learning rate was set to 0.001 and reduced slowly after each epoch. We trained for 20 epochs and only saved the parameters that yielded the best result. The momentum and weight decay were set to 0.9 and 10e-4, respectively.

TABLE I
9-LAYERED LIGHTCNN ARCHITECTURE FOR DEEP FEATURE EXTRACTION FROM FACE AND CLOTHING MODALITIES.

Type	Filter Size/ Stride, Pad	Output Size	#Param
Convolution	5 x 5/1, 2	128 x 128 x 96	7.3K
MFM	-	128 x 128 x 48	-
MaxPooling	2 x 2/2	64 x 64 x 48	-
Convolution	1 x 1/1	64 x 64 x 96	4.7K
MFM	-	64 x 64 x 48	-
Convolution	3 x 3/1, 1	64 x 64 x 192	83.1K
MFM	-	64 x 64 x 96	-
MaxPooling	2 x 2/2	32 x 32 x 96	-
Convolution	1 x 1/1	32 x 32 x 192	18.6K
MFM	-	32 x 32 x 96	-
Convolution	3 x 3/1, 1	32 x 32 x 384	332.2K
MFM	-	32 x 32 x 192	-
MaxPooling	2 x 2/2	16 x 16 x 192	-
Convolution	1 x 1/1	16 x 16 x 384	74.1K
MFM	-	16 x 16 x 192	-
Convolution	3 x 3/1, 1	16 x 16 x 256	442.6K
MFM	-	16 x 16 x 128	-
Convolution	1 x 1/1	16 x 16 x 256	33K
MFM	-	16 x 16 x 128	-
Convolution	3 x 3/1, 1	16 x 16 x 256	295.2K
MFM	-	16 x 16 x 128	-
MaxPooling	2 x 2/2	8 x 8 x 128	-
fc1	-	512	4194.8K
MFM_fc1	-	256	-
Total	-	-	5527K

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

C. Deep Feature Fusion

We used a deep learning-based fusion model in Figure 5 with three fully connected layers to combine the features from face and clothing modalities. The fusion network accepts two 256-dimensional feature vectors and forms a concatenated feature vector of length 512. This is followed by a fully-connected layer of 256 dimensional vector. The last layer is the softmax layer to train the model on the dataset. Table II shows the total number of parameters for lightCNN and ResNet-18 models for face and clothing information. The total number of parameter of the system using ResNet-18 is 22.96M, about twice as the lightCNN, which is 11.3M.

TABLE II
THE NUMBER OF PARAMETERS IN 9-LAYER LIGHTCNN AND RESNET-18 FOR FACE AND CLOTHING MODALITY ALONG WITH THEIR DEEP FEATURE FUSION.

	Modality	Input shape	#Param
lightCNN	Clothing	128x128x3	5527K
	Face	128x128x3	5527K
	Deep Fusion	512	262.6K
	Total	-	11316.6K
ResNet18	Cloth	224x224x3	11,350K
	Face	224x224x3	11,350K
	Deep Fusion	512	262.6K
	Total	-	22962.6K

When training the fused model, we froze the parameters for both face and clothing models. Similar to its constituent models, the momentum of the fusion model was set to 0.9, the weight decay to 10^{-4}, and the learning rate to 0.001, which was adjusted after every epoch. We trained the model for 20 epochs and only the best parameters were saved. Figure 5 shows the architecture of the multimodal CNN network for deep feature fusion of the face and clothing information.

D. Deep Feature Matching

Deep features from pair of training and testing images were obtained by activating one of the fully-connected layers of the trained models. The matching score between the deep features were obtained using cosine (equation 3), L1 (equation 4) and euclidean distance (equation 5) metrics given below.

$$d_{cos}(A,B) = \frac{\sum_{i=1}^{n} A_i B_i}{\sqrt{\sum_{i=1}^{n} A_i^2}\sqrt{\sum_{i=1}^{n} B_i^2}} \tag{3}$$

$$d_{L1}(A,B) = \sum_{i=1}^{n}(A_i - B_i) \tag{4}$$

$$d_{Euclidean}(A,B) = \sqrt{\sum_{i=1}^{n}(A_i - B_i)^2} \tag{5}$$

IV. EXPERIMENTAL EVALUATION

A. Dataset and Protocol

In this work, we used Spoof in the Wild (SiW) face anti-spoofing database [24] to evaluate our models. The dataset contains 8 live and up to 20 spoof videos of 165 subjects which were collected in 4 sessions with various distances, poses, illuminations, and facial expressions. Since the purpose of this work is not anti-spoofing, we used only the live videos to extract the image data for our tasks. For every 10 frames, we selected one and generated more than total of $100,000$ images this way.

We chose the SiW dataset for several reasons. First, the number of subjects (i.e., 165) is relatively high. Second, the dataset comes from video streams akin to security camera feeds in passenger re-authentication scenarios. Third, almost all images in the dataset contain both face and clothing information. Moreover, the participants wore the same clothes throughout the video. Popular face datasets such as Labeled Faces in the Wild (LFW) do not have the participants wearing same clothes, making it unsuitable for our task. Lastly, various distances, poses, and lighting of SiW dataset give us a wide range of variations.

We extracted frames from live videos and only chose the subset of images that contained both face and clothing information. In total, about $95,000$ images - 575 images per each participant - were picked for our experiment. Half of the images, stratified per participant, were used for training and the rest were set aside for testing. The training set and testing set are overlapped (closed-set). We used cosine, Euclidean and L1 distance metrics to obtain matching scores on deep features generated by lightCNN, ResNet and Deep feature-fusion models for face and clothing modalities.

The verification performance of our two modalities (face, clothing) and their deep fusion was evaluated via ROC Equal Error Rate (EER) and Genuine Match Rate (GMR) at various False Match Rate (FMR) levels. We also fused the scores of the models associated with clothing and face modalities via simple sum rule as the baseline.

B. Results and Discussion

Table III shows the EER of face and clothing modalities along with their deep feature fusion and simple sum rule, using three different distance metrics: cosine, Euclidean, and L1. We can see that ResNet-18 outperforms 9-layer lightCNN in all the experiments by around 2% EER. Cosine matching was the best per the above metrics for both models (lightCNN and ResNet) and modalities (face and clothing). The deep feature fusion significantly improved the performance of the re-authentication system. For example, with lightCNN, single modality achieved about 4.5% EER, while the deep fusion of features produced 1.47% EER with cosine distance. For ResNet model, sum rule outperformed deep feature fusion, which can be justified based on higher complexity of the ResNet and thus its theorized higher variance that could have been averaged out by the sum rule. On the other hand, deep feature fusion obtained slightly better results than sum rule for lightCNN. As such, larger training and development sets should improve deep feature fusion model across scenarios, especially when dealing with more complex models.

Table IV shows the GMR@10^{-4}FMR and GMR@10^{-5}FMR of ResNet-18 and 9-layer lightCNN ;

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

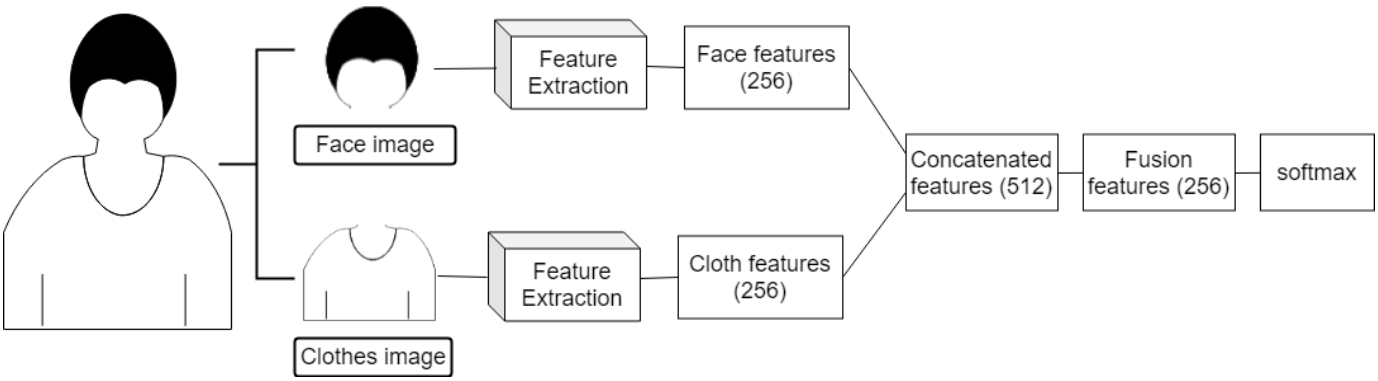

Fig. 5. Deep feature fusion of face and clothing information for passenger re-authentication at various checkpoints in an ABC system.

TABLE III
EER (%) FROM ON MATCHING DEEP FEATURES OBTAINED FROM RESNET18, 9-LAYER LIGHTCNN, DEEP FEATURE FUSION MODEL AND SIMPLE SUM-RULE FOR FACE AND CLOTHING MODALITIES. DEEP FEATURES WERE MATCHED USING COSINE, L1 AND EUCLIDEAN DISTANCE METRICS.

	Metric	Clothing	Face	Sum Rule	Deep Fusion
LightCNN	Cosine	4.913	4.373	1.540	1.470
	Euclidean	6.800	6.681	2.500	2.437
	L1	6.924	6.823	2.581	2.457
ResNet18	Cosine	0.29	0.34	0.007	0.238
	Euclidean	0.412	0.503	0.007	0.294
	L1	0.375	0.500	0.007	0.293

TABLE IV
GMR(%) AT FMR = 10E-5 AND FMR = 10E-4 WHEN MATCHING DEEP FEATURES OBTAINED FROM RESNET18, 9-LAYER LIGHTCNN, DEEP FEATURE FUSION MODEL AND SIMPLE SUM RULE FOR FACE AND CLOTHING MODALITIES. DEEP FEATURES WERE MATCHED USING COSINE, L1 AND EUCLIDEAN DISTANCE METRICS.

		@FMR = 10e-5				@FMR = 10e-4			
	Metric	Clothing	Face	Sum Rule	Deep Fusion	Clothing	Face	Sum Rule	Deep Fusion
LightCNN	Cosine	37.82	20.90	63.59	68.17	49.52	43.30	80.79	86.70
	Euclidean	35.55	16.50	44.86	59.84	43.32	31.00	69.74	76.49
	L1	36.57	16.10	45.50	59.07	41.94	31.39	68.66	75.50
ResNet	Cosine	67.45	91.80	100.00	99.02	97.62	96.45	100.00	99.16
	Euclidean	62.65	87.35	100.00	95.16	96.35	96.20	100.00	99.85
	L1	63.45	86.27	100.00	95.59	96.16	96.10	100.00	97.15

deep feature fusion and simple sum rule; for face and clothing information using three distance metrics (cosine, L1 and euclidean). For systems using ResNet-18, at FMR = 10^{-5}, face recognition achieved better accuracy than clothing with the highest GMR being 91.8%. At FMR = 10^{-4}, clothing modality slightly outperformed the face modality. In both cases, deep feature fusion produced best results with 99.02%GMR@10^{-5}FMR and 99.16%GMR@10^{-4}FMR. However, sum rule obtained the best performance for three distance metrics with 100% GMR.

Using lightCNN, clothing modality notably outperformed face modality at GMR@10^{-4} FMR and GMR@10^{-5}FMR. Deep feature fusion overwhelmingly improved the performance of the system. For example, with cosine distance metric, at FMR = 10^{-5}, clothing and face modalities achieved 37.82% and 20.9% GMR, respectively, but the fusion attained 68.17% GMR which is 30% higher compared to face alone. Deep feature fusion model obtained better results than simple sum rule. Again, cosine metric performed the best among other distance metrics.

V. CONCLUSION AND FUTURE WORK

In this paper, we proposed a biometric fusion system suitable for secure, fast and convenient passenger re-authentication at different security checkpoints in an ABC system. The proposed method takes advantage of (non-ideal) face and complements it with clothing information. We employed two lightweight CNN models: a 9-layer lightCNN and a ResNet18. We also introduced a deep feature fusion model that combines face and clothing information at one of the later intermediate layer of the CNN. Three different distance metrics were evaluated to match features from enrollment and verification samples. Deep feature fusion improved upon face and clothing scores by a wide margin. It was noted that clothing, though not a biometric, can be a strong visual identification signal (albeit only for the short term such as the case for an ABC system), given the differences in individuals' preferences for clothing items, combined with clothing's larger coverage. This,

978-1-7281-5092-5/19/$31.00 ©2019 IEEE

of course, needs to be moderated by a primary biometric such as face to form a reliable identity signal, a feat that we achieved via fusion.

As a part of future work, our models will be evaluated in an open-set environment where the subjects do not overlap between the training and testing set. Larger datasets will be used to improve the deep feature fusion model, especially when fusing larger, higher variance models. We would also like to evaluate other soft biometric traits such as glasses, eyebrows, and gender to further enhance the performance of the biometric fusion re-authentication system. Lastly, an adaptive fusion system for combining the primary biometric (face) with secondary soft biometric traits or auxiliary signals of identity may be developed to improve system performance, especially in the case of partial face information.

References

[1] J. Fergusson, E. U. E. A. for the Management of Operational Cooperation at the External Borders, and E. A. Z. W. O. na Granicach Zewnetrznych Państw Członkowskich Unii Europejskiej, *Twelve Seconds to Decide: In Search of Excellence: Frontex and the Principle of 'best Practice'*. Publications Office, 2014. [Online]. Available: https://books.google.com/books?id=hpkEoQEACAAJ

[2] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Biometric recognition in automated border control: A survey," *ACM Comput. Surv.*, vol. 49, no. 2, pp. 24:1–24:39, Jun. 2016. [Online]. Available: http://doi.acm.org/10.1145/2933241

[3] D. Gorodnichy, S. Yanushkevich, and V. Shmerko, "Automated border control: Problem formalization," in *2014 IEEE Symposium on Computational Intelligence in Biometrics and Identity Management (CIBIM)*, Dec 2014, pp. 118–125.

[4] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Automated border control systems: Biometric challenges and research trends," in *Information Systems Security*, S. Jajoda and C. Mazumdar, Eds. Cham: Springer International Publishing, 2015, pp. 11–20.

[5] R. D. Labati, A. Genovese, E. Muoz, V. Piuri, F. Scotti, and G. Sforza, "Advanced design of automated border control gates: Biometric system techniques and research trends," in *2015 IEEE International Symposium on Systems Engineering (ISSE)*, Sep. 2015, pp. 412–419.

[6] R. D. Labati, A. Genovese, E. Muñoz, V. Piuri, F. Scotti, and G. Sforza, "Biometric recognition in automated border control: A survey," *ACM Comput. Surv.*, vol. 49, no. 2, pp. 24:1–24:39, Jun. 2016. [Online]. Available: http://doi.acm.org/10.1145/2933241

[7] A. Rattani, N. Reddy, and R. Derakhshani, "Multi-biometric convolutional neural networks for mobile user authentication," in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, 2018, pp. 1–6.

[8] J. Snchez del Ro, C. Conde, A. Tsitiridis, J. Ral Gmez, I. Martn de Diego, and E. Cabello, "Face-based recognition systems in the abc e-gates," in *2015 Annual IEEE Systems Conference (SysCon) Proceedings*, April 2015, pp. 340–346.

[9] M. Smith-Creasey, F. A. Albalooshi, and M. Rajarajan, "Continuous face authentication scheme for mobile devices with tracking and liveness detection," *Microprocessors and Microsystems*, vol. 63, pp. 147 – 157, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0141933117305239

[10] U. Mahbub, V. M. Patel, D. Chandra, B. Barbello, and R. Chellappa, "Partial face detection for continuous authentication," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sep. 2016, pp. 2991–2995.

[11] E. Ahmed, M. Jones, and T. K. Marks, "An improved deep learning architecture for person re-identification," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2015.

[12] D. Cheng, Y. Gong, S. Zhou, J. Wang, and N. Zheng, "Person re-identification by multi-channel parts-based cnn with improved triplet loss function," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[13] E. S. Jaha and M. S. Nixon, "Soft biometrics for subject identification using clothing attributes," in *IEEE International Joint Conference on Biometrics*, Sept 2014, pp. 1–6.

[14] J. D. Shutler, M. G. Grant, M. S. Nixon, and J. N. Carter, "On a large sequence-based human gait database," in *Applications and Science in Soft Computing*, A. Lotfi and J. M. Garibaldi, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 339–346.

[15] M. H. Nguyen, R. Puttagunta, Z. Li, and R. Derakhshani, "User re-identification using clothing information for smartphones," in *2018 IEEE International Symposium on Technologies for Homeland Security (HST)*, Oct 2018, pp. 1–6.

[16] A. Rattani and R. Derakhshani, "A survey of mobile face biometrics," *Computers Electrical Engineering*, vol. 72, pp. 39 – 52, 2018.

[17] A. Rattani, R. Derakhshani, and A. Ross, *Selfie Biometrics: Advances and Challenges*. Springer, 2019.

[18] D. E. King, "Dlib-ml: A machine learning toolkit," *J. Mach. Learn. Res.*, vol. 10, pp. 1755–1758, Dec. 2009. [Online]. Available: http://dl.acm.org/citation.cfm?id=1577069.1755843

[19] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," *CoRR*, vol. abs/1505.04597, 2015. [Online]. Available: http://arxiv.org/abs/1505.04597

[20] X. Wu, R. He, and Z. Sun, "A lightened CNN for deep face representation," *CoRR*, vol. abs/1511.02683, 2015. [Online]. Available: http://arxiv.org/abs/1511.02683

[21] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *CoRR*, vol. abs/1512.03385, 2015. [Online]. Available: http://arxiv.org/abs/1512.03385

[22] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *CoRR*, vol. abs/1409.1556, 2014. [Online]. Available: http://arxiv.org/abs/1409.1556

[23] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105. [Online]. Available: http://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf

[24] Y. Liu, A. Jourabloo, and X. Liu, "Learning deep models for face anti-spoofing: Binary or auxiliary supervision," *CoRR*, vol. abs/1803.11097, 2018. [Online]. Available: http://arxiv.org/abs/1803.11097

978-1-7281-5092-5/19/$31.00 ©2019 IEEE