iFace: A Deepfake Resilient Digital Identification Framework for Smart Cities

Alakananda Mitra
Dept. of Computer Science and Engineering
University of North Texas, USA.
Email: AlakanandaMitra@my.unt.edu

Saraju P. Mohanty
Dept. of Computer Science and Engineering
University of North Texas, USA.
Email: saraju.mohanty@unt.edu

Peter Corcoran
School of Engineering and Informatics
National University of Ireland, Galway, Ireland.
Email: peter.corcoran@nuigalway.ie

Elias Kougianos
Dept. of Electrical Engineering
University of North Texas, USA.
Email: elias.kougianos@unt.edu

***Abstract*—Digital ID is the gateway of "Smart City" for "Smart Citizens". It gives citizen access to all other stakeholders of smart cities like smart healthcare, smart transport, smart finance, smart energy, etc. effectively and easily. In this paper, we propose a bio-metric based digital ID which is implemented in IoT environment. It is a secured and robust system against deepfake attacks. A convolutional neural network (CNN) based feature extraction method has been employed to defeat deepfake attacks. The dlib face detector has been used in detecting face landmark points and in calculating distances in the iris and nose region to obtain unique facial features. A bio-key is generated from the combination of features from facial landmarks and various facial distances along with the username. An encoded key is stored in a cloud database during the registration process of the user. For accessing any facilities in a smart city, the user needs to be authenticated. The authentication process is performed at the edge. Small changes in an image due to unconstrained settings are corrected using the Reed Solomon algorithm. Once authenticated at a particular smart facility, the user is now eligible to use that facility.**

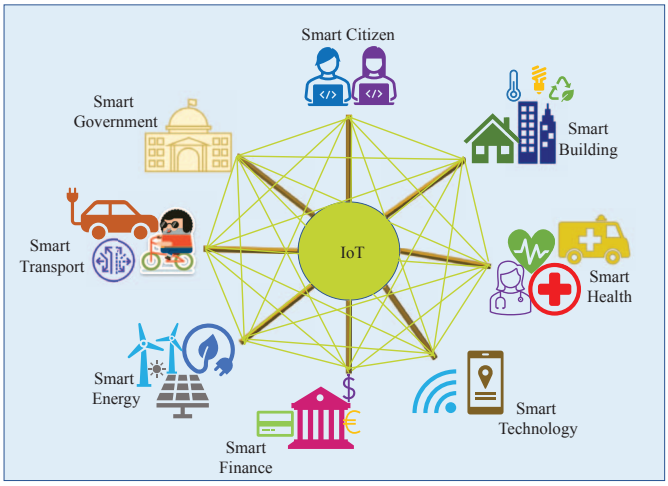


Fig. 1. Smart City Components Connected Through IoT.

I. Introduction

In the last two decades, due to growth in hardware and software design, information and communication technology (ICT) has taken a giant leap which increases the effectiveness of city life and operations, resulting into the "Smart City". Various stakeholders are interconnected through the Internet of Things (IoT), as shown in Fig. 1. By 2050, 70% of the total world population will be urban [1]. Citizens benefit from the various services a smart city provides. Hence, citizens need to be smart to avail all smart city facilities. The full potential of smart cities can only be unlocked when a digital ID will be issued to citizens to access all facilities or applications.

A bio-metrics digital ID is person-specific and unique as bio-metrics are connected to physiological or behavioral traits of a human being. So there is no need to keep any secret key to protect privacy and identity of a user, as the user himself is his secret key. So digital ID, based on bio-metrics, can be a smart choice to unlock the closed doors of a smart city.

In this paper, we propose a robust machine learning (ML) based digital ID system where from enrollment to authentication of a user can be performed remotely and securely at an edge device to work in an IoT environment of smart city. Our method is robust against deepfake attacks.

The rest of the paper has been organized into eight sections. Section II discusses the challenges of digital ID. Novel contributions of the work are presented in Section III. Existing works in this area are noted in Section IV. Our proposed work is described in detail in Section V. Implementation and results are stated in Section VI and Section VII. Section VIII concludes the paper with suggestions for future work.

II. Challenges of Digital Identification (ID)

The main challenge of digital ID is to authenticate a real person. To have better security and to protect the privacy of the user, a digital ID should be bio-metric based. The requirements for a bio-metric based digital ID are shown in Fig. 2.

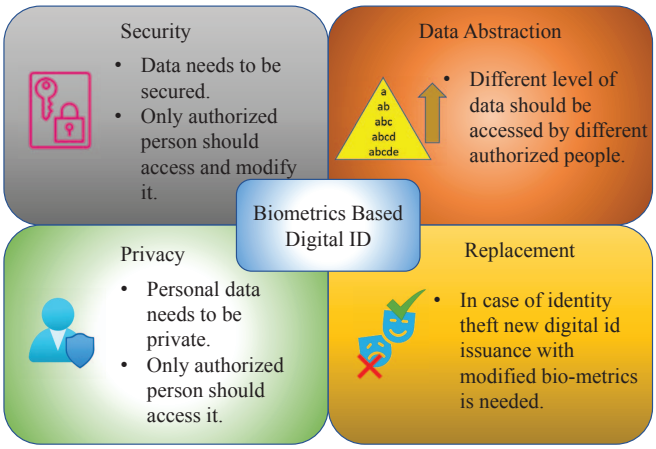


Fig. 2. Mandatory Requirements for Digital ID in a Smart City.

Facial bio-metric based digital IDs are prone to various types of attack such as face spoofing, deepfake, indirect attack, etc. Deepfake is a special type of face spoofing attack, based on deep neural networks. In [2]–[4] deepfake images and videos have been detected. In this paper, along with the end-to-end digital id system, we address the deepfake attack too.

III. Novel Contributions of the Current Paper

We propose a ML based facial bio-metric digital ID for smart cities. The novel features of this method are as following:

- The registration process is done at the edge, remotely and securely. The encoded bio key is stored at a cloud server.
- Authentication of the digital ID is also performed at the edge. As it is done near to the user, it is free of various indirect attacks.
- During authentication, a photo is taken at the edge and bio keys are generated with the help of the stored registration bio key. The system will not allow any tampering at this stage by checking for deepfake attack.
- Bio-metric data can vary at various photos taken at different times. It is not possible for people to keep their face in the same way all the time. Our method can accommodate a certain amount of modification of bio-metric data due to these changes.
- Our process will not be affected by unconstrained environments like different lighting conditions. Only a neutral frontal face (NFF) is needed for the authentication. Our model is robust for faces with light makeup.
- We use face features like facial landmark points and certain distances in the eye and nose regions as the bio-metric of our proposed digital ID. Identification of the face is much more relevant in a smart city setting. Fingerprints are not always accessible as our hands might not always be free. Behavioral traits are highly affected by the environment.

IV. Related Works

In this section, we discuss some relevant works which motivate us to propose our face-features based digital ID. We summarize papers which address cryptographic key generation from face features and also bio-metric authentication systems in the IoT . Works related to fingerprints, gait, handwriting, and speech based key generation methods have not been summarized.

An entropy based method has been explored in [5] to regenerate a cryptographic key. Features are extracted from the images using entropy based method and Reed-Solomon ECC has been performed to generate the bio-key. Look-up tables have been created to regenerate the original key. In our work, we also use look-up tables, but the use and scope are different. An eigenfunction based face recognition method has been mentioned in [6]. It does not generate any key but tracks the user's head. Finally, it recognizes the face by comparing the traits between the user and data stored. A 128 bit key has been generated from a principal component analysis feature vector using thresholding and distinguishable bits with a right sequence number are updated in a look-up table [7]. Finally, the Reed-Solomon algorithm has been used to create an error correcting code (ECC). Symmetric DES and the generated key are used to encrypt any message. In the decryption stage, the reverse procedure is done. A key is computed by connecting several multi-bit keys generated from various threshold value [8]. An optimal threshold value has been chosen to reduce the authentication error. The methods mentioned above generate bio key based on face features but have not been implemented in limited-resource IoT devices.

A detailed survey has been made for face verification and authentication for IoT mobile devices in [9]. Another low complexity deep learning based face recognition method has been implemented in an embedded device [10]. A secure biometrics based end-to-end IoT solution has been mentioned in [11]. To increase the security, pairing-based cryptography has been used. A face recognition system, implemented in FPGA for digital forensics application, has been presented in [12]. A deep learning based method has been described in [13] for an IoT-cloud setting with a tree-based cloud model for face verification. The edge part is optional for processing and filtering images. Our work fits in the same setting as this paper but with versatile scope, as our proposed method addresses the security part of the facial authentication system by computing the bio-key at the edge and by using an encoding key. During authentication, bio keys are compared in our method instead of images. This makes our method more robust.

V. Proposed Biometric Based Digital ID in Smart Cities Context

A. End-to-End System Level Architecture

The proposed biometric based digital ID system consists of a layered architecture which is distributed among edge and cloud computing platforms, starting from the end user to the cloud server, as shown in Fig.3. This four layered structure consists of the following:

1) Layer-1: It consists of the Smart Citizen with digital ID, various types of cameras from different smart city stake holders, and an input device to provide user ID. When digital ID verification of a person is necessary, these cameras take a photo and send it along with the username to Layer-2. Cameras can be smart phone cameras or any cameras installed as the end device. The user ID can be inserted using the keypads at the end device itself.
2) Layer-2: Edge Computing Platform works as Layer-2. The photo and username from layer 1 come to this layer. As both layer 1 and layer 2 are at the same location, no transmission of biometric data or username happens over open channels at this stage. This alleviates the necessity of encryption of the biometric data at this point. Most processing and computing steps are performed here. Bio keys are generated and encoded here.
3) Layer-3: It comprises of a cloud computing platform. It is connected to edge devices through various long range technologies like *4G*, *LTE*, etc. The data is encoded and sent to the cloud. Layer-3 is mainly used for storing large amount of bio-metric data and usernames.
4) Layer-4: Smart city stakeholders are the key components of layer 4. Once the bio key is authenticated in layer 2 with the information from layer 3, the digital ID is verified and smart city application is accessed through its API.

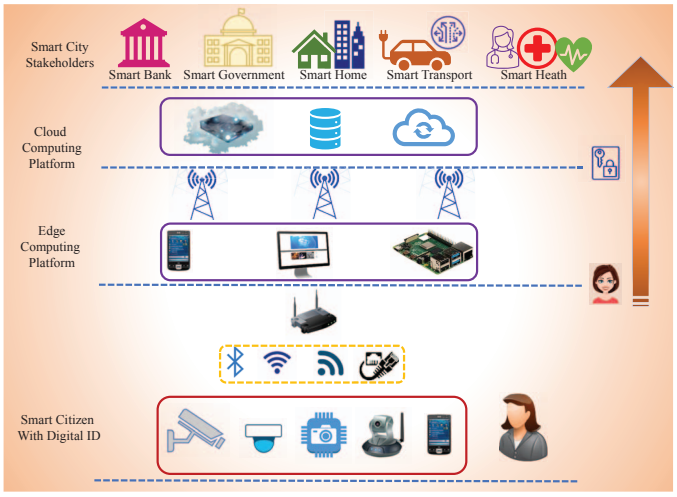


Fig. 3. End-to-End System Level Framework for Digital ID System in a Smart City.

B. Proposed Method

1) System overview: We propose a digital ID system for smart cities which consists of two phases:

- Enrollment or registration of new users.
- Authentication of existing users.

The process of enrollment of a new user is shown in Fig. 4. In the enrollment phase, a unique username is issued to the new user after verifying the existing government issued ID. Bio-metric facial features of the person are extracted from a neutral frontal face (NFF) image taken by the end device camera. Bio keys are then generated from the image, are encoded, and saved in the cloud server along with the username.

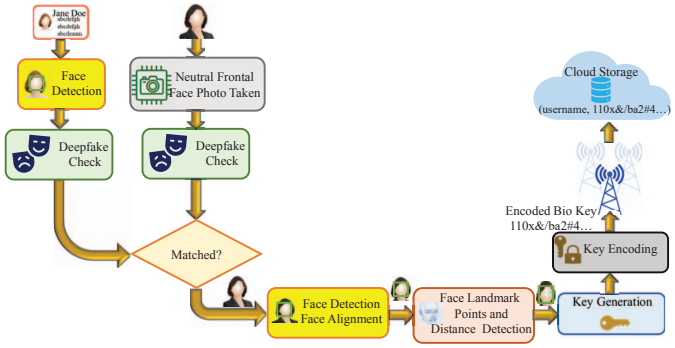


Fig. 4. Registration of a New User.

During the authentication phase shown in Fig. 5, the user provides his username and an NFF photo is shot by the end device. Once the photo is taken, bio keys are generated. To avoid any discrepancy in input data, NFF images are taken in both phases. But, as these photos are taken in unconstrained environments, it is highly unlikely that a person will have the exact same photo for all occurrences. If two photos are not exactly the same, they will generate different bio keys. Those bio keys are not completely different but little variations will be present. Our system can accommodate a certain level of such modifications while generating bio keys using Reed Solomon error correcting codes (ECC). If the system can generate bio key at registration from the bio key at authentication phase, it verifies the person through digital ID and gives access to the smart city facility that the user wants to access. If the two faces do not match, the user does not get access.

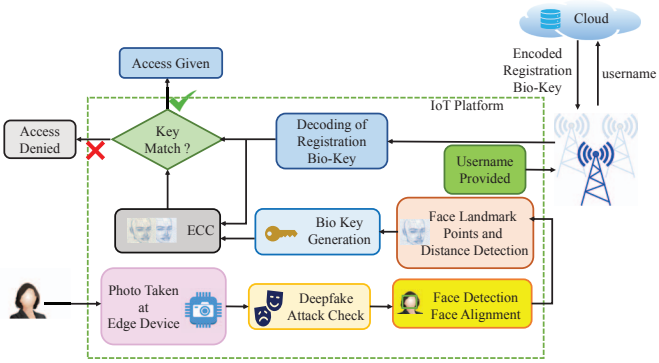


Fig. 5. Authentication of Existing User.

2) Deepfake Attack Detection: An NFF photo of the user is taken through the camera attached to the edge device. The taken photo is in the RGB color space. The photo is checked for deepfake detection attacks. We follow the procedure from

our previous work [14]. However here we use MobileNetV2 [15] as the feature extractor and a softmax layer as the classifier. We fine tune a pre-trained MobileNetV2. We train the last 40 layers in this 53 layer structure.

3) Biometric Features Extraction: Once the photo is passed through the deepfake check, it is ready for biometric feature extraction. The process of biometric face features extraction is shown in Fig.6 and Fig.7. First the face is detected from the NFF photo with the dlib [16] library using Histogram of Oriented Gradient (HOG) and linear Support Vector Machine (SVM) in Fig.6. The reasons behind choosing HOG based dlib library are the following:

- As we detect the face at the edge where resources are limited, HOG based dlib face detector is the best choice.
- It is the fastest and lightest model among face detectors and suitable for IoT environment.
- Frontal photo is considered for digital ID. We do not expect a person, when verifying with digital ID, will show a side view of the face. So, HOG based dlib works perfectly well in our used case scenario.
- This model works better with CPU. No GPU is required to detect faces. This characteristic makes it a good fit for using at an edge device.

Once the face is detected, it is aligned with OpenCV. It will limit some positional discrepancies of two photos taken at different times. 68 facial landmarks points related to jaw, both eyebrows and eyes, nose, and mouth are then detected using the HOG based dlib face detector [16]. The (x, y) coordinates of 68 facial landmark points make a 68×2 feature vector $\mathcal{F}1$ in Fig. 6.

Another feature vector $\mathcal{F}2$ of dimension 1×6 is formed with d_1, d_2, d_3, d_4, d_5, and d_6 as in Eq. 1:

$$\begin{aligned} d_j &= d_{1j} \\ d_{1j} &\in \mathcal{F}2^{(1\times 6)} \end{aligned} \tag{1}$$

d_j are calculated as in Fig. 7, by calculating the diagonal distance of both iris and the nose length. As the bio-metric data is unique for each individual, the feature vector is *sui generis* too.

4) Biometric Key Generation : In this section, a novel bio key generation technique is proposed. First, a binary key is generated from the feature vector $\mathcal{F}1$, as shown in Fig. 8. The process consists of three steps:

- First, the 68×2 dimensional feature vector is reshaped to 1×136 as $(x_1, y_1, x_2, y_2, ..., x_{68}, y_{68})$.
- Second, a unique threshold value d_m is calculated using Eq. 2, as shown in Fig. 7(b).

$$d_m = d_5 + d_6 \tag{2}$$

- Binarization of the feature vector $\mathcal{F}1_b$ is performed by comparing each element f_i of the feature vector $\mathcal{F}1$ to the threshold value d_m following Eq. 3:

$$f_{bi} = \begin{cases} 0, & \text{if } f_i < d_m \\ 1, & \text{if } f_i \geq d_m \end{cases} \tag{3}$$

Finally, $\mathcal{F}1_b$ and $\mathcal{F}2$ are concatenated to form the unique feature vector or final bio key $\mathcal{F}_{io}$ following Eq. 4:

$$\mathcal{F}_{io} = \mathcal{F}1_b + \mathcal{F}2 \tag{4}$$

5) Error Correction: In our proposed digital ID system, we use $\mathcal{F}_{io}$ as the biometric feature to authenticate a person. $\mathcal{F}_{io}$ is robust against lighting and is also unique to an individual. However, getting the same picture of a person at various times and by various cameras of a smart city is almost impossible. These variations in pictures can alter the bio key at a certain percent. To accommodate these variations and avoid false rejection ratio, we use Reed Solomon (RS) codes [17] to correct the errors. The face matching process is shown in Fig. 9(a) and Fig. 9(b).

In the registration phase, $\mathcal{F}_{io}$ is encoded with Reed Solomon codes and saved in a look-up table in the cloud server. The look-up table comprises of two columns - username U and encoded bio key $\mathfrak{F}_{\mathfrak{io}}$ as shown in Fig. 9(a).

During the authentication phase, as shown in Fig. 9(b), the user provides the username which finds the corresponding encoded bio key $\mathfrak{F}_{\mathfrak{io}}$ in the look-up table. Then we split $\mathfrak{F}_{\mathfrak{io}}$ in original input $\mathcal{F}_{io}$ and error correcting code ECC. ECC is then combined with $\mathcal{F}_{mod}$, collected at this stage from the authentication photo. It generates encoded authentication $\mathfrak{F}_{\mathfrak{mod}}$. If the photo at this stage differs from the photo taken at registration, it gets corrected $\mathcal{F}'_{io}$ with the Reed Solomon decoding module. If the decoding module is able to generate the original $\mathcal{F}_{io}$ at this point then the faces are matched and the user gets access to the specific facility of smart city where he used his digital ID.

An attacker can get the encoded bio key from the look-up table in cloud, but they will not be able to impersonate any person as the authentication process is being performed at the edge with the presence of user. That makes the system robust.

VI. Experimental Validation

A. Dataset

We use two different datasets for our work. DeepFakeDetection dataset part of Face Forensics ++ dataset [18] has been used for deepfake attack detection.

For evaluation of our proposed digital ID system, three different datasets have been used. The dataset details are summarized in Table I. Neutral frontal face (NFF) without any occlusion is required for our system. Various frontal face datasets are publicly available but very few datasets contain NFF. For this reason, we take 250 neutral faces or close to neutral faces from CelebA [19] dataset to form *Dataset-1*. The second dataset *Dataset-2* is a neutral face dataset from Kaggle [20]. However, both datasets contain only one NFF image for each individual. It does not fully evaluate our digital ID system. These datasets result in 0% False Acceptance Rate (FAR) and 0% False Rejection Rate (FRR) but this is not a fully correct evaluation of our method. A rightful user with a different image taken at authentication should be tested too. Finally, we test the performance of our digital ID system with a customized dataset *Dataset-3*

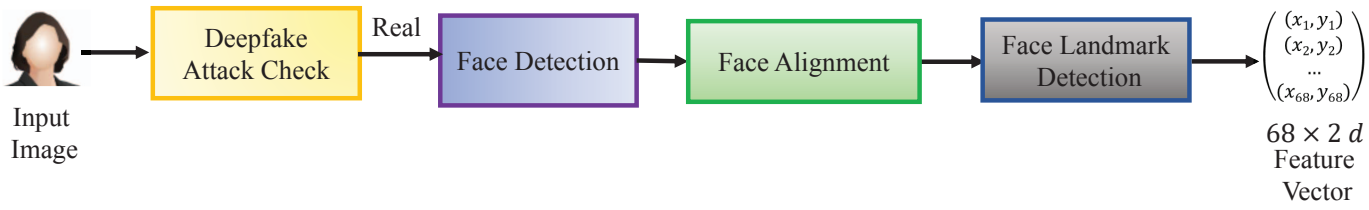
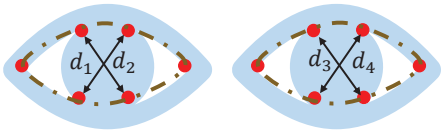
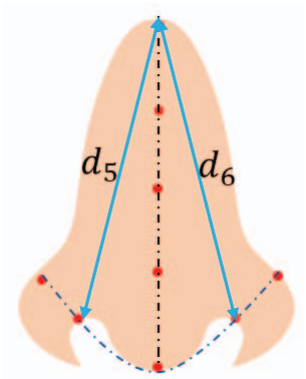


Fig. 6. Facial Landmark Points Detection Workflow.



(a) Diagonal Distance Calculation of Iris



(b) Mean Distance Calculation

Fig. 7. Facial Distance Calculation

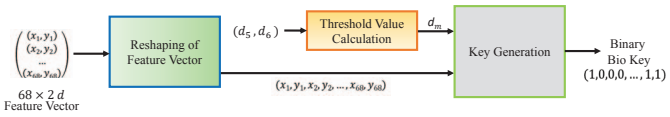
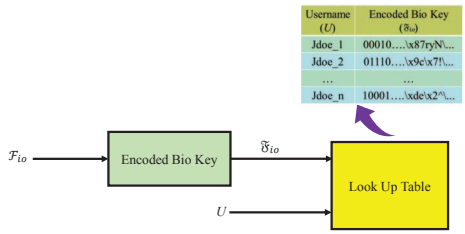


Fig. 8. Binary Key Generation from Face Landmark Points.

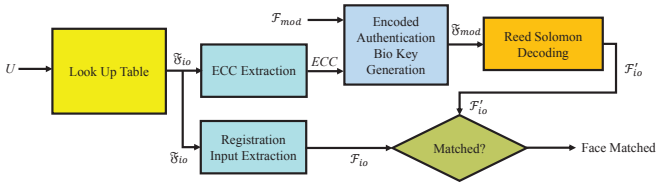
of total 60 images of 30 individuals, mostly celebrities and political figures collected from Google search. Two images are collected for each person - one image for enrollment and the other image for authentication. Most of the tested images are neutral but some are with a slight smile or with mouth open. For each registered user, the other 29 users' images have been used as impostor images resulting in 900 tested combinations during authentication.

TABLE I
DETAILS OF THE DATASET

Dataset Name	Source of the Dataset	No. of Images
Dataset-1	CelebA [19]	250
Dataset-2	Frontal Faces Neutral Expression 95 Landmarks [20]	240
Dataset-3	Internet	60



(a) Process of Saving Bio Key at Registration Phase



(b) Process of Face Matching at Authentication Phase

Fig. 9. Face Matching Workflow.

B. Implementation

We implement our proposed digital ID system in Python using a GeForce RTX 2060 laptop with a 6GB shared memory of total 16GB memory. We evaluate our system with the three datasets mentioned above. For deepfake detection we use [18] for both training and testing purpose. The message length during error correction of encoded message is 148. 4 bit RS codec has been used to avoid intruders.

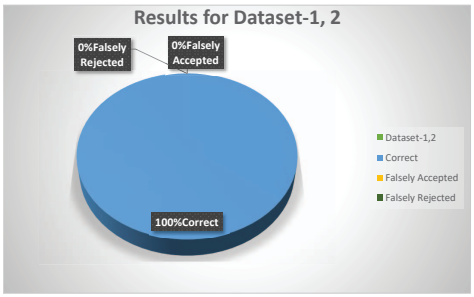
VII. Results

The results of the experiments with three different datasets are shown in Table. II and Fig. 10.

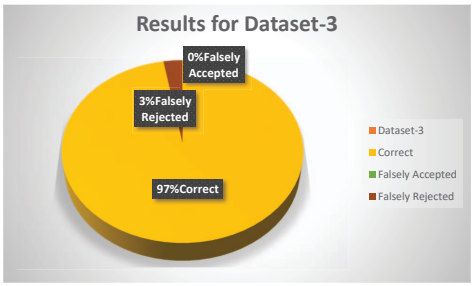
TABLE II
PERFORMANCE OF OUR PROPOSED MODEL

Dataset	No. of Testing Images	No. of Cases Authenticated		
		Correct	Falsely Accepted	Falsely Rejected
Dataset-1	1000	1000	0	0
Dataset-2	1000	1000	0	0
Dataset-3	900	875	0	25

For Datasets 1 and 2, as there is only one image per user, the same image has been used for registration end authentication. As a result, FRR is 0%. For a specific user, we declare an impostor set containing other users' images. Here 0% FAR



(a) For Dataset 1 and Dataset 2



(b) For Dataset 3

Fig. 10. Performance of Proposed Digital ID System with Different Datasets

means no one other than the user is given access for that specific user ID. So, no impostor can access the user data. These results are included in this paper to demonstrate that impostors cannot access the tested user's account.

For Dataset 3, we use different images for registration and authentication. Not all images fulfill the exact criterion of neutral frontal face (NFF). As a result we see 25 images are falsely rejected even if they are rightful people. This generates an FRR of 2.77%. The accuracy of the deepfake attack check in our case is 91%.

There are certain scenarios which have not been addressed in this work but will be considered in future work:

- If the person looks considerably different from the photo taken at registration, the system can not authenticate.
- Heavy eye make up like smokey eyes can generate a false rejection.
- Identical twins scenario has not been considered.

VIII. Conclusion and Future Work

In this paper, we propose a biometric based end-to-end digital ID system of a smart city. Our research is in preliminary stage but we have been able to achieve several things:

- Our system can detect certain deepfake attacks.
- It does not allow impostors to access users data.
- It is robust to various lighting conditions.

As a future work, presentation attack detection module can be added and deepfake detection module can be updated to accommodate deepfake attacks from different origins. An efficient digital ID system can be achieved which authenticates people with glasses, mask, hats and certain age related changes.

References

[1] S. P. Mohanty, U. Choppali, and E. Kougianos, "Everything you wanted to know about smart cities: The internet of things is the backbone," *IEEE Consumer Electronics Magazine*, vol. 5, no. 3, pp. 60–70, 2016.

[2] A. Mitra, S. P. Mohanty, P. Corcoran, and E. Kougianos, "A machine learning based approach for deepfake detection in social media through key video frame extraction," *SN Comput. Sci.*, vol. 2, no. 2, p. 98, 2021.

[3] A. Mitra, S. P. Mohanty, P. Corcoran, and E. Kougianos, "Easydeep: An iot friendly robust detection method for gan generated deepfake images in socialmedia," in *Proceedings of the 4th FIP International Internet of Things (IoT) Conference (IFIP-IoT)*, 2021, Accepted, In Press.

[4] A. Mitra, S. P. Mohanty, P. Corcoran, and E. Kougianos, "Detection of deep-morphed deepfake images to make robust automatic facial recognition systems," in *Proceedings of the 19th OITS International Conference on Information Technology (OCIT)*, 2021, Accepted, In Press.

[5] B. Chen and V. Chandran, "Biometric based cryptographic key generation from faces," in *9th Biennial Conference of the Australian Pattern Recognition Society on Digital Image Computing Techniques and Applications (DICTA 2007)*, 2007, pp. 394–401.

[6] M. A. Turk, A. Pentland *et al.*, "Face recognition using eigenfaces." in *IEEE Conf. Comp. Vision and Pattern Recognition*, vol. 91, 1991, pp. 586–591.

[7] L. Wu, X. Liu, S. Yuan, and P. Xiao, "A novel key generation cryptosystem based on face features," in *Proceedings of IEEE 10th International Conference on Signal Processing*, 2010, pp. 1675–1678.

[8] W. Zhang, Y.-J. Chang, and T. Chen, "Optimal thresholding for key generation based on biometrics," in *2004 International Conference on Image Processing, 2004. ICIP'04.*, vol. 5. IEEE, 2004, pp. 3451–3454.

[9] M. A. Ferrag, L. Maglaras, and A. Derhab, "Authentication and authorization for mobile iot devices using biofeatures: Recent advances and future trends," *Security and Communication Networks*, vol. 2019, 2019.

[10] S. H. Oh, G.-W. Kim, and K.-S. Lim, "Compact deep learned feature-based face recognition for visual internet of things," *Journal of Super Computing*, vol. 74, pp. 6729–6741, 2018.

[11] M. S. Hossain, G. Muhammad, S. M. M. Rahman, W. Abdul, A. Alelaiwi, and A. Alamri, "Toward end-to-end biomet rics-based security for iot infrastructure," *IEEE Wireless Communications*, vol. 23, no. 5, pp. 44–51, 2016.

[12] M. Pantopoulou and N. Sklavos, "An fpga-implemented parallel system of face recognition, for digital forensics applications," in *In Proceedings of IEEE 10th International Conference on Consumer Electronics (ICCE-Berlin)*, 2020, pp. 1–6.

[13] M. Masud, G. Muhammad, H. Alhumyani, S. S. Alshamrani, O. Cheikhrouhou, S. Ibrahim, and M. S. Hossain, "Deep learning-based intelligent face recognition in iot-cloud environment," *Computer Communications*, vol. 152, pp. 215–222, 2020.

[14] A. Mitra, S. P. Mohanty, P. Corcoran, and E. Kougianos, "A novel machine learning based method for deepfake video detection in social media," in *IEEE International Symposium on Smart Electronic Systems (iSES) (Formerly iNiS)*, 2020, pp. 91–96.

[15] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2018, pp. 4510–4520.

[16] D. E. King, "Dlib-ml: A machine learning toolkit," *Journal of Machine Learning Research*, vol. 10, pp. 1755–1758, 2009.

[17] I. S. Reed and G. Solomon, "Polynomial codes over certain finite fields," *Journal of the Society for Industrial and Applied Mathematics*, vol. 8, p. 300–304, 1960.

[18] A. Rössler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Nießner, "FaceForensics++: Learning to detect manipulated facial images," in *International Conference on Computer Vision (ICCV)*, 2019.

[19] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep learning face attributes in the wild," in *Proceedings of International Conference on Computer Vision*, December 2015.

[20] A. Morzhakov. Frontal faces neutral expression 95 landmarks. https://www.kaggle.com/antonmorzhakov/frontal-faces-neutral-expression-95-landmarks/code. Accessed on 01 July 2021.