Pattern Recognition Letters 000 (2018) 1–9

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

CNN-based anti-spoofing two-tier multi-factor authentication system

Muhammad Sajjad[a], Salman Khan[a], Tanveer Hussain[a], Khan Muhammad[b], Arun Kumar Sangaiah[c], Aniello Castiglione[d], Christian Esposito[d], Sung Wook Baik[b,*]

[a] Digital Image Processing Laboratory, Department of Computer Science, Islamia College Peshawar, Peshawar 25000, Pakistan
[b] Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul 143-747, Republic of Korea
[c] School of Computing Science and Engineering, VIT University, Vellore, Tamil Nadu 632014, India
[d] Department of Computer Science, University of Salerno, Via Giovanni Paolo II, Fisciano, (SA) 132 I-84084, Italy

ARTICLE INFO

Article history:
Available online xxx

Keywords:
Biometric recognition
Anti-spoofing
Information security
Convolutional neural networks

ABSTRACT

Many hybrid and multimodal biometric recognition techniques have been presented to provide secure and authentic systems, incorporating both soft and hard biometric schemes. This article proposes a new hybrid technique which ensures the authenticity of the user to the system, as well as monitors whether the user has passed the biometric system as a normal or spoofed one. The proposed scheme is two-fold: Tier I integrates fingerprint, palm vein print and face recognition to match with the corresponding databases, and Tier II uses fingerprint, palm vein print and face anti-spoofing convolutional neural networks (CNN) based models to detect spoofing. In first stage, the hash of a fingerprint is compared with the fingerprint database. After a successful match of the fingerprint, it is tested on a CNN-based model of the fingerprint to verify whether it is a spoof or real. A similar process is repeated for the palm and face, and based on collective evidence, the system permits the user to login the system. Experimental results over five benchmark datasets verified the effectiveness of the proposed system in providing efficient and robust verification, overcoming the limitations in normal authentication and spoofing practices.

1. Introduction

Information privacy and security have been playing an important role in human life for the past few decades. Information represents important aspects of our daily life [1] and personal authentication has therefore received considerable attention from researchers, as it is becoming an important issue. Different techniques have been used for securing personal and private information, including crystographic [2,3], steganographic [4–7] and biometric technologies [8–10], Biometric technology provides automated systems for securing information and enabling the authentication of individuals based on behavioural and biological information such as face, iris, voice, fingerprint and palm veins [11,12]. These methods are supposed to be more secure and accurate for authentication and identification, due to their high accuracy as compared to other non-biometric techniques. Biometric systems have been used in many applications such as video surveillance [13], biometric identification [14] and face indexing in multimedia contents [15].

There are many hybrid techniques [16,17] that combine multiple factors for extensive security of data and information. Spoof attacks [18] are the techniques used for outwitting a biometric system by presenting a false sample of the person to get authentication. In face spoofing, different types of high-resolution photographs, such as a 3D mask having resemblance with human skin, along with video reply attacks have been used [19]. On the other hand, in finger spoofing, the attacker uses moulds to outwit the biometric system. For palm vein spoofing, images of the palm have been printed on the standard printer which can easily bypass the biometric authentication [20]. To overcome these problems, different techniques have been used by researchers. For instance, using a texture-based approach, Maatta et al. [21] used multi-scale binary patterns for analysis of spoof attacks. But with the advancement of the technology, high quality images can gain illegitimate access from the biometric system. Another study conducted by Sun et al. [22] used eye blinking for spoof detection. But a video of high quality resolution can fool the system. Considering these limitations, we propose a hybrid scheme combining the fingerprint, palm and face recognition using hand-crafted features, along with their spoof detection using convolutional neural network (CNN) based high-level features.

The recent success of the deep convolutional neural networks in the field of image classification [23,24], as well as in object recog-

* Corresponding author.
E-mail addresses: muhammad.sajjad@icp.edu.pk (M. Sajjad), khan.muhammad@ieee.org (K. Muhammad), sbaik@sejong.ac.kr (S.W. Baik).

https://doi.org/10.1016/j.patrec.2018.02.015
0167-8655/

nition [25], has attracted researchers to utilize these multi-layer end-to-end learning architectures to perform a variety of tasks. CNNs consist of many convolutional layers, followed by fully connected layers, to produce a probability distribution for the training classes. The activation of the neurons from the connected layers is used for different applications, including scene recognition [26], action recognition [27], spoof detection [28] and image retrieval [29]. Conventional methods use hand-crafted features, followed by classifier training for solving anti-spoofing problems. For example, Marsico et al. [30] used a 3D projective invariant-based method for face anti-spoofing. Jiangwei et al. [31] presented a Fourier spectra based method, which is further based on the assumption that photos contain higher frequency components. On the other hand, Xu et al. [32] presented a CNN architecture for spoofing attacks by putting a long short-term memory (LSTM) layer over the fully connected layers for feature extraction. Many computer vision tasks, such as face recognition [33], gender [34] and video classification [35] use CNN due to its robust performance. Considering these motivations, CNN-based anti-spoofing models have been incorporated in the proposed framework.

The main objective of this paper is to develop a two-tier, efficient and robust framework, insuring the authenticity and security of the system. In the first level, we use hand-crafted feature matching for efficient fingerprint, palm and face identification. The fingerprint hash is searched in the database using perceptual hashing, the palm vein is recognized using SIFT (scale-invariant feature transform) [35], and ORB points (Oriented FAST and Rotated BRIEF) are used for face [36]. In the next level, CNN is trained to detect a spoof image in any of the biometrics (i.e. fingerprint, palm and face). Our CNN architecture was inspired by GoogLeNet [37] and is used to perform features extraction for anti-spoofing. Three pre-trained models were fine-tuned using anti-spoof fingerprint, anti-spoof palm and anti-spoof face datasets to function as a real-time feature extractor in the anti-spoof detection system. The proposed system uses these deep CNN features for spoof detection in any of the biometrics. Our main contributions are summarized below:

(1) We propose a two-tier novel framework for both recognition and anti-spoofing, unlike existing approaches which either focus on recognition or anti-spoofing.
(2) Our system uses both hand-crafted and learned high-level representations for recognition and anti-spoofing, respectively.
(3) We propose a hybrid scheme which combines the fingerprint, palm and face modalities for effective recognition and intelligent spoof detection.

The rest of the paper is organized as follows: Section 2 explains the proposed framework and discusses its two tiers and factors. Experimental results on different datasets are given in Section 3. Section 4 concludes the paper with a discussion on the strengths and shortcomings of this work, along with future directions.

2. Proposed methodology

This section explores the proposed method, describing the concept of making a system private, secure and preventing access by malicious users. The proposed system has two security tiers: T_1 and T_2. Each tier has three phases (T_{1a}, T_{1b}, T_{1c} and T_{2a}, T_{2b}, T_{2c}). The first security tier uses conventional feature matching techniques for recognition of biometrics such as fingerprint (FP), palm vein (PV) print and face (F) while the second tier uses CNN for detection of spoofing. Each phase has an output class which is either positive (Φ) or negative (Ψ) to determine if the output is acceptable, or not. The positive class refers to an output that is acceptable for further processing, and negative class output is not considered for next steps. In the very first step, the FP image is acquired through the reader and then through a perceptual hashing technique. The image is converted into a 64-bit unique hash, followed by its matching with database of hashes, Ω_{fp}. Hamming distance is used to calculate the difference between two fingerprint hashes, and the threshold selected is 0.8 after a large number of experiments for same two hashes. In case of an exact match, the fingerprint image of the individual is passed as input to T_{2a}, which has a CNN-based model, trained on benchmark database LiveDet 2013 [38] for FP anti-spoofing. The classifier in T_{2a} outputs the prediction as to whether the fingerprint under consideration is a spoof or belongs to the actual individual. In the next step, a palm vein image is obtained through a palm reader, from which SIFT features are extracted and are compared with features of the palm vein database, Ω_p. Upon a successful match, the palm image is classified as a spoof or original by the CNN classifier, trained on a standard dataset of Vera Spoofing Palm Vein [20]. Failure (spoof or not matched) at any step proclaims the individual as unauthorized. In the last step, T_{1c} and T_{2c} phases are involved. T_{1c} inputs the user image, detects and crops the face and recognizes it as known or anonymous from the database, Ω_f. In the case of a recognized person, the image is further passed to the CNN trained classifier to detect whether the image is spoof or real. If the face is marked as non-spoof, the individual is granted access to the system and is considered as an authorized person. The list of parameters used in our system are given in Table 1. The overall framework is given in Fig. 1.

Table 1
Description of the parameters used in our framework.

T_i	Tier I	FP	Fingerprint
T_{ij}	Tier I, Phase j	@FP	Fingerprint anti-spoofing
Ψ	Not matched, Spoof	PV	Palm vein
Ω_{fp}	Fingerprint database	@PV	Palm Vein Anti-Spoofing
Ω_p	Palm images database	F	Face recognition
Ω_f	Face recognition database	@F	Face anti-spoofing
Φ	Matched, Not spoof		

2.1. Tier I

The first tier contains three factors, and all of them use conventional feature extraction and matching techniques. The system has three locally collected databases: a fingerprint hash database, palm images database and authentic faces database. In the first factor for Tier I, a FP reader is used to take the fingerprint scan of the user. A perceptual hashing technique, with extra pace of histogram equalization, is used to obtain the hash of the image. The input scan image of the fingerprint is converted into grey scale. To reduce the computation, the image is resized to 256 by 256 pixels. In the next step, a threshold is set by calculating the mean value of the whole image, and this mean value is selected as threshold. Pixels having values greater than the threshold are set to one and others to zero. A 64-bit hash is obtained by converting the binary vector into the decimal system. Finally, the fingerprint hash is matched with the database of hashes. Hashing technique for the fingerprint matching is preferred because it is an efficient and robust technique, as compared to other fingerprint matching methods.

The second factor for Tier I is palm vein image matching with the database, during which the input image of the palm vein is acquired through a palm vein reader. As palm vein images use near-infrared (NIR) illumination, they usually appear dark and low contrast, which makes it unclear and difficult for palm- based recognition systems. To compare it with the database, SIFT features of the palm image are extracted and searched in the palm images database of users. The SIFT features are preferred for palm vein recognition systems because they are scale invariant, which makes them less constrained for palm vein recognition. If other features

Please cite this article as: M. Sajjad et al., CNN-based anti-spoofing two-tier multi-factor authentication system, Pattern Recognition Letters (2018), https://doi.org/10.1016/j.patrec.2018.02.015

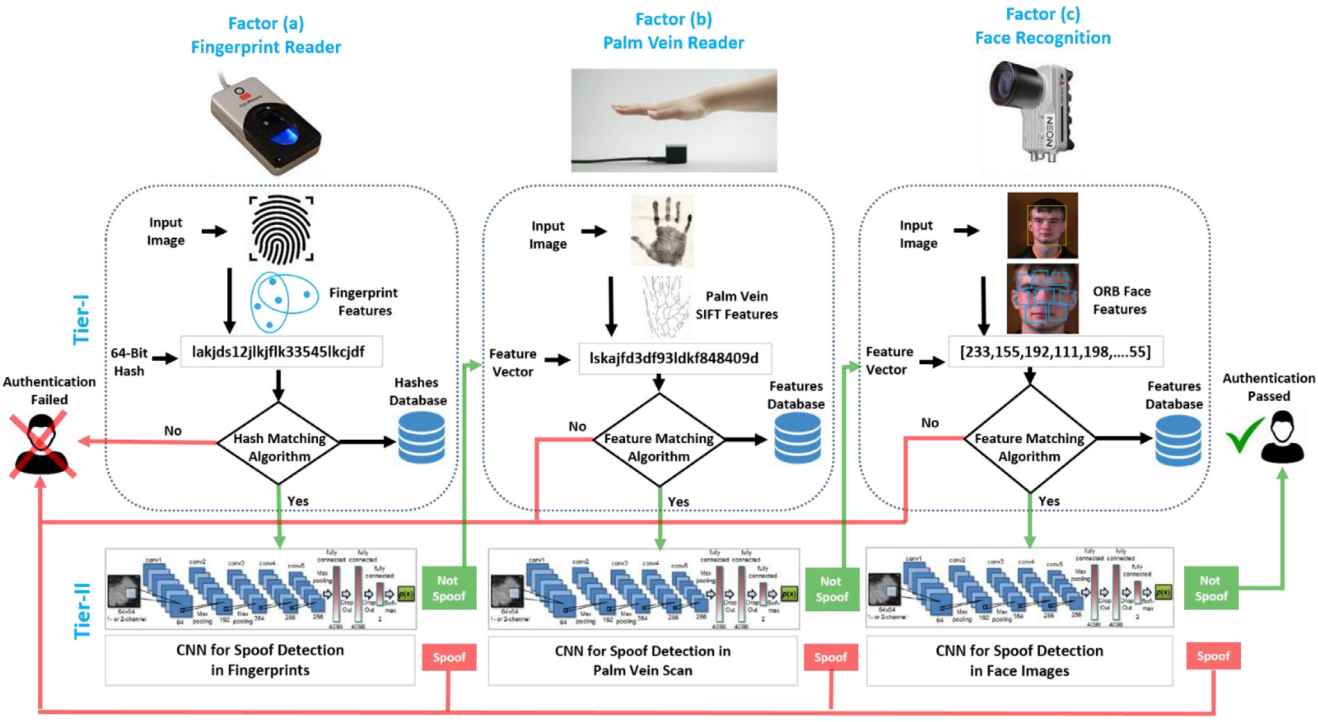

Fig. 1. Proposed framework for anti-spoofed authentication system.

are used, which are variant, the individual must place the palm with extra care for proper recognition. However, direct usage of the SIFT extraction technique makes the extraction of key points from the palm image very difficult, so contrast enhancement is used along with SIFT feature extraction.

The third and last Tier I factor includes detection of the face, feature extraction from the face and a comparison of the extracted features with the stored features in the database. Features extracted from the face are ORB points, which are drawn on the detected face image and matched with the same points of database. ORB features are more suitable and feasible for real-time face recognition systems with limited resources. That is why our framework uses these features.

2.2. Tier II

Tier II uses deep learning features so that the system may not be deceived by fake fingerprints, masks or any other replica biometrics. The first factor is related to detection of a fingerprint spoof. There are two methods for creating a fake fingertip: the cooperative and non-cooperative methods. In the cooperative method, the malicious user pushes the finger into a plastic material and creates a fake fingerprint mark as a mould, which is then filled with a material like gelatine or silicone to reproduce the same, but false, fingerprint characteristics [18]. In the very first phase of Tier 2, the image that was passed as authentic from the phase 1 of Tier I is given as input to the CNN classifier to classify the image as spoof or real. A deep learning model has been trained on datasets ATVS-FFp [39] and LivDet 2013 [38] through the GoogleNet model [37].

Vein recognition has emerged as a new biometric for accurate and fast people identification, and has received growing attention because of live-body, anti-interference identification and simple-acceptability. But there are too many malicious users who can breach these types of security techniques, thus intensive care is needed to secure the system. Deep neural networks are used for training processes on datasets of VERA spoofing palm-vein for spoof detection, which classifies the image of the palm as spoof or real.

Facial biometric spoofing techniques involve placing genuine photographs or dummies, and playing a video recording in front of the camera. The biometric system may be deceived by using a 3D physical model. This is known as a synthesis attack. To avoid all types of such attacks, we have trained the model on REPLAY-ATTACK [40] and CASIA [41] datasets using a convolution neural network, which observes and detects very minute points of the query image to detect the false face.

3. Experimental results

In this section, we present the experimental evaluation of the proposed scheme for the two-tier multi-factor authentication system. Experiments were performed on different anti-spoofing datasets, whose details are given in the subsequent sections.

3.1. Datasets

The proposed framework is evaluated using different datasets available to the research community. As spoofing is a highly technical task, even for a well-trained classifier, it is very difficult to detect the faulty user. Therefore, our method collects features from these datasets very carefully. Details of the datasets used in evaluations are provided in their respective sections.

(a) REPLAY-ATTACK [40]
For the anti-spoofing of a face, REPLAY-ATTACK was used, which is available for the research community from the IDIAP research institute. This database contains short videos clips of ~10 s in length, and the video format is .mov for both real-access and spoofing attack attempts of 50 different subjects. Videos are recorded in 320×240 resolution with a webcam from a 13 inch MacBook laptop. The video recordings are carried out under two different scenarios: (i) controlled, with a static background and artificial lighting, and

Table 2
Details of LivDet dataset.

Scanner	Model	Res (dpi)	Image size	Live samples	Fake samples
Biometrika	FX2000	569	312×372	2000	2000
Italdata	ET10	500	640×480	2000	2000
Crossmatch	L SCAN GUARDIAN	500	640×480	2500	2000
Swipe		96		2374	1979
	(Total Samples)			8874	7979

(ii) adverse, with natural illumination and a changing background.

(b) CASIA CBSR [41]
This database is publicly available from the Chinese Academy of Science (CASIA) for conducting research about security. The database contains 10 s video clips in avi format for both real and attack attempts. These short clips are recorded by 50 subjects using three different devices. The low-resolution types were recorded with an old 640×480 USB web camera. The normal resolution types were recorded with a 480×640 USB camera. The model of both cameras is specified in the database. The high-resolution video clips were recorded on a high-definition Sony NEX-5 with a resolution of 1920×1080 pixels. For attacks, three scenarios were considered. The first scenario was warped, which involved illegal attempts with a slight difference in the hard copies of high-resolution digital photographs of the genuine users. The second scenario was cutting, for which the same procedure for warped was followed, but only the eyes were cut out for the forging of eye blinking. The face of the attacker was put behind the photograph. In the third scenario, high resolution videos of the genuine users were replayed in front of an acquisition device with an iPad. Frames extracted from the videos of both real and fake access attempts can be found in the CASIA-FSD DB.

(c) ATVS-FFp [39]
The ATVS-FFp database consists of the index and mid fingers for the left and right hands of 17 subjects ($17 \times 4 = 68$ different fingers) and is publicly available for research purposes. Two spoofs are generated for each real finger using silicon with two methods, including the subject and not including the subject. Four fingerprint samples of real and spoofed were captured in one acquisition session with three sensors. The specification of the sensors is Biometrika FX2000, Precise SC100 and Yubee. The database consists of 816 real image samples (68 fingers $\times$ 4 samples $\times$ 3 sensors) and many spoofed images for each method.

(d) LivDet 2013 Fingerprint [38]
The LivDet 2013 is a fingerprint liveness database. It has been divided into four sub-sets, which consist of live and fake fingerprint samples, captured from four different devices. Images have been gathered by consensual methods and using different materials for non-natural copies of fingerprints (gelatine, silicone, play-doh, wood glue, ecoflex and body double). Table 2 contains information about the scanner, model, resolution in dpi, image size, and numbers of live and fake samples in the LivDet 2013 database.

(e) VERA Spoofing Palm-vein [20]
The VERA spoofing palm-vein database was generated at the Idiap Research Institute in Martigny, Switzerland. This database is based on 1000 images of 50 different clients from the Idiap Research Institute. Person identification is performed based on vein patterns formed by the blood vessel on the skin, which creates less external distortion and are more difficult to forge. An infrared beam is used by the sensor to capture the structure of the vein in the individual's hand, which is compared with an existing vein pattern to confirm the identity of the individual. The vein pattern in living humans is detectable and unique because it develops with birth and cannot be changed throughout the life. The infrared light is captured by the hemoglobin in the blood present in the vessel, forming a dark shadow of veins.

3.2. Experimental design

In the proposed system, experiments on these datasets are performed using a Python 3.5 running over a PC having a GPU of 3.20 GHz * 4 processors with 8.0 GB of RAM over Ubuntu version 16.04 and 64-bit operating system with a Caffe [42] framework. Several experiments were conducted to evaluate the performance of the proposed scheme in anti-spoofing applications. Details and outcomes of the experiments with different datasets are explained in the subsequent sections.

3.3. Tier I results

Tier I results are compared to other state-of-the-art methods in Table 3. The first column indicates factors, containing different matching of the biometrics, such as fingerprint matching, palm vein matching and face recognition. We used a local database of individuals for fingerprint, palm vein and face recognition for the identification process, and it contains 50 samples for each fingerprint, palm vein and face. The third column shows the accuracy (ACC) of the proposed method. Our preliminary results are dominating other state-of-the-art methods with accuracies of 100%, 99% and 98% for fingerprint, palm vein and face recognition, respectively.

3.4. Tier II results

Tier II is anti-spoofing phase that involves all the steps required for the identification of individuals as spoof or real. Table 4 shows the results of Tier II. The first column represents the factors used in Tier II containing fingerprint, palm vein and face anti-spoofing. The second column shortlists the datasets that were used in the training process for our method. For face anti-spoofing, we used Replay attack and CASIA MERA dataset. For fingerprint anti-spoofing, we used ATVS-FFp and LivDet 2013 datasets. The VERA spoofing palm-vein dataset was used for palm vein anti-spoofing. The remaining columns show the results of our method and a comparison with other methods. For face anti-spoofing, using the Replay attack dataset, the accuracy was 98.87% and for CASIA MERA, the accuracy was 99.55%. For fingerprint anti-spoofing, we obtained accuracies of 98% and 97.5% for the ATVS-FFp and LivDet 2013 datasets, respectively. Using the VERA spoofing palm-vein dataset for palm vein anti-spoofing, the accuracy was 99%.

In the visual results section, Figs 2 and 3 represent the anti-spoofing results for fingerprints for both the ATVS-FFp and LivDet 2013 datasets. The prediction made by the classifier in Fig 2(a)–(f) were all correctly classified as actual. The classifier in Fig. 3(a)–(f) predicted the fingerprints into their actual class, while in Fig. 3(c), the predicted result is negative. In the pictorial results section, Fig.4 represents the anti-spoofing results for the palm vein prints

Table 3
Accuracy of Tier I comparison with other methods.

Factor	Dataset	Our results (ACC)	Other state-of-the-art methods	
			ACC	Ref
Fingerprint matching	Local database	100%	98%	[43]
Palm vein matching	Local database	99%	96%	[43]
Face recognition	Local database	98%	96%	[44]

Table 4
Accuracy of Tier II comparison with other methods.

Factor	Dataset	Our results (ACC)	Other methods	
			ACC	Ref
Fingerprint anti-spoofing	ATVS-FFp LivDet 2013	99.75% 97.5%	97.45%	[45]
Palm vein anti-spoofing	Vera spoofing palmvein	99.5%	–	–
Face anti-spoofing	Replay attack CASIA MERA	98.87% 98.55%	98.75	[45]

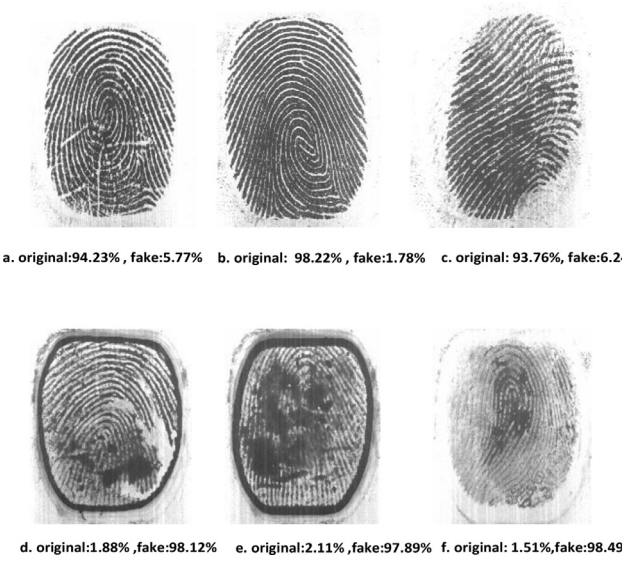

Fig. 2. Visual results with predictions given by our proposed model for ATVS-FFp F fingerprint anti-spoofing dataset.

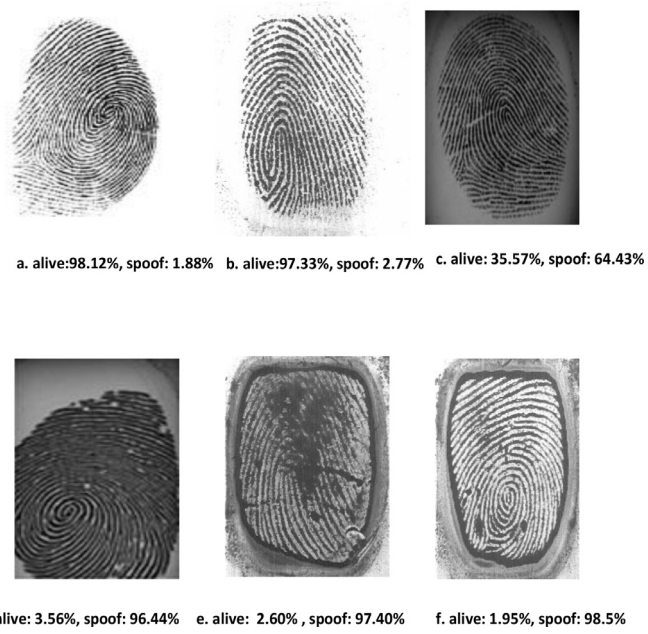

Fig. 3. Sample visual results with their predictions from LivDet 2013 Fingerprint anti-spoofing dataset.

from the VERA spoofing palm-vein dataset. The predicted outputs made by the classifier in Fig. 4(a)–(f) were correct as it classified them properly. Total results for the palm vein were classified correctly and there were no wrong classifications. In Figs 5 and 6, the anti-spoofing results for the face using both Replay attack and CASIA MERA datasets are shown. The prediction made by the classifier in Fig. 5(a)–(f) was accurately classified into its deserving class, while in Fig. 5(d), the predicted output was negative. Predictions made by the classifier in Fig 6(a)–(f) were accurate and it classified faces into their real class, while in Fig. 6(e), the predicted result was negative.

3.5. Complete results

(a) Overall accuracy of the proposed system
In order to evaluate the performance of the proposed system, we conducted experiments on various groups of individuals. There were 5 groups of entities, with each entity containing 10 individuals being tested on each step of the proposed system and then compared with the ground truth. Table 5 can be explained as follows. There were three phases of fingerprint, palm vein, and face, having both authentication and anti-spoofing steps. For the first group, for the fingerprint authentication phase denoted by the yellow colour, the ground truth indicates seven registered persons in the database and three non-registered persons. During fingerprint based experiments for authentication, our proposed system allowed seven users and did not give access to the three unregistered users, which is in accordance with the ground truth. In the ground truth of fingerprint anti-spoofing represented by yellow colour, the first group of individuals had five spoof and five not-spoof subjects. Our proposed system allowed two fake users to enter the system, showing the weak perspective of our system. The third and fourth column of Table 5 shows the ground truth of palm vein in the blue colour, having nine unregistered and one registered user and five spoofs and five not-spoof users. The third and fourth column shows results of our system, allowing eight users as not registered and two users as reg-

Table 5
Overall accuracy of the proposed system.

Groups	Ground Truth												Proposed System												Accuracy
	FP		@FP		PV		@PV		F		@F		FP		@FP		PV		@PV		F		@F		
	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	Ψ	Φ	
Group I	3	7	5	5	9	1	5	5	4	6	3	7	3	7	3	7	8	2	5	5	4	6	3	7	100%
Group II	6	4	4	6	3	7	4	6	3	7	6	4	6	4	5	5	3	7	4	6	3	7	6	4	100%
Group III	5	5	5	5	6	4	5	5	6	4	5	5	4	6	5	5	6	4	5	5	6	4	4	6	100%
Group IV	7	3	3	7	5	5	3	7	5	5	2	8	7	3	7	3	5	5	3	7	5	5	2	8	100%
Group V	4	6	8	2	8	2	8	2	8	2	8	2	4	6	4	6	8	2	7	8	8	2	8	2	100%

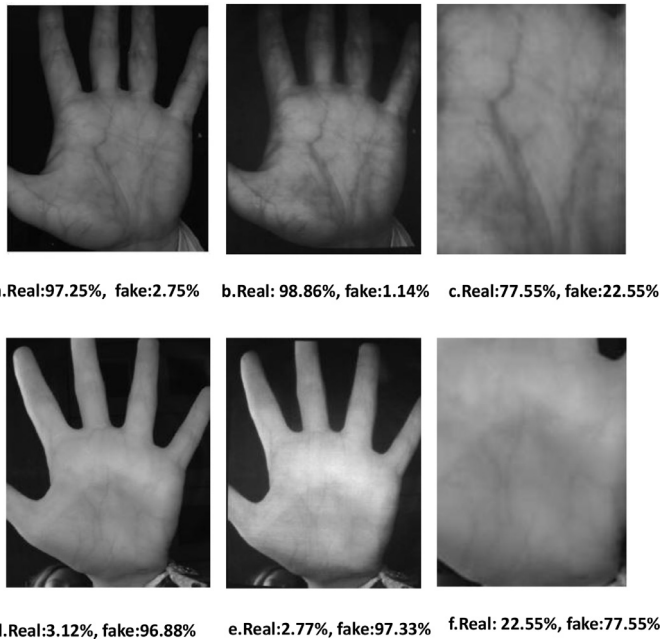

Fig. 4. Sample images from VERA spoofing palm-vein dataset with their predications provided by our method.

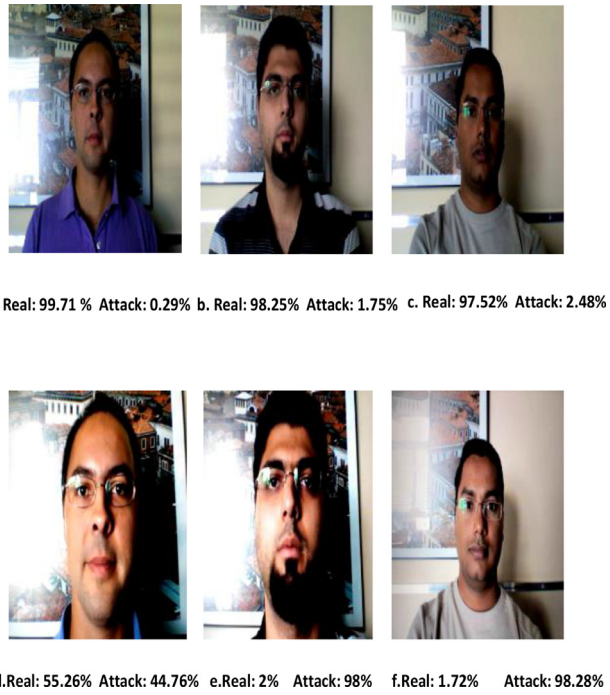

Fig. 5. Sample visual results with their predictions from REPLAY-ATTACK Face anti-spoofing dataset.

istered. In case of anti-spoofing, our system restricted five as spoof and allowed the other five as not-spoof. In the final step of security represented by green colour in Table 5, there is face recognition and face anti-spoofing with ground truth for four not-registered and six registered faces of individuals, along with three spoofs and seven not-spoof persons. In case of face matching, four faces were not-registered and six were recognized faces, while three faces sdfd considered as spoof and seven as non-spoofs, following the ground truth. From the above results for the first ten entities, it can be clearly observed that if a person by-passes a single step of security with any sort of fake technique, he can be trapped in the next steps giving 100% accuracy at the very last step. The same process was applied on the remaining four groups, giving 100% accuracy and trapping the malicious user, as shown in Table 5.

(b) Combined results based on all factors

The accuracy of different factors for the proposed system are represented in Fig. 7. There were 5 groups observed by the proposed system, with each group consisting of 10 individuals. The accuracy percentage increased when more factors were added. The y-axis represents the level of accuracy and x-axis represents the factors involved. The proposed factors include the following: (1) Fingerprint authentication (FP), (2) Fingerprint anti-spoofing (@FP), (3) Palm vein recognition (PV), (4) Palm vein anti-spoofing (@PV), (5) Face recognition (F) and (6) Face anti-spoofing (@F).

The accuracy increased as the individual was passed from one factor to the next. From the graph, for the accuracy of only fingerprint authentication, the system accuracy was only 52%. The accuracy increased from 52% to 76% with the

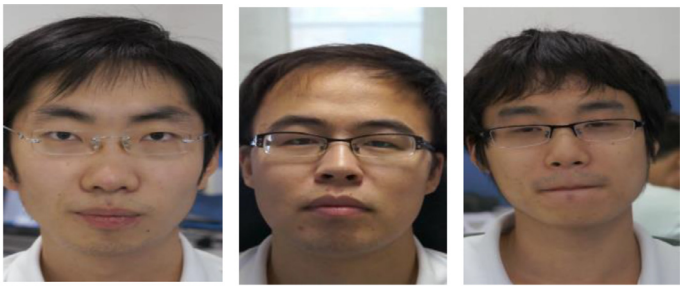

Fig. 6. Sample visual results with their predictions from CASIA CBSR Face anti-spoofing dataset.

additional factor of fingerprint anti-spoofing. By adding the palm vein authentication, the accuracy reached to 82%, and further increased to 88% by the addition of palm vein anti-spoofing. The accuracy reached 94% with the high secure step of face recognition, and finally to 100% with face anti-spoofing. The accuracy of the system continuously increased with forward steps. To avoid repetition, only group 1 is explained here, and accuracy for the rest of the groups can be viewed from the graph in Fig. 7.

(c) Computational time of proposed time

The computational cost of the proposed system is shown pictorially in Fig. 8. The computation cost of each step is given individually and overall. Fingerprint authentication step required 0.44 s and fingerprint anti spoofing classifier consumed 4.3 s. Palm vein recognition required 0.61 s and palm vein anti-spoof detection required 3.6 s. Face recognition and face anti-spoofing required 0.83 and 5 s, respectively. The total time required to run all the steps smoothly was 14.78 s.

4. Conclusion and future work

The literature for information security contains several applications, hardware and spoofing methods to compromise biometric authentication systems and pass without being detected. Dental mould and kid's clay are all needed to bypass any biometric system, which proves that fingerprint spoofing is too easy to implement. In addition, the practice of printed images of the palm vein and silicon sheets to deceive the palm vein recognition system is very common. Similarly, 3D masks, print attacks and replay attacks have bluffed the face recognition algorithms. In an attempt to avoid such attacks, we proposed a novel anti-spoofing framework containing several authentication and identification steps for ensuring the security of the system. The proposed technique is divided into two tiers: one for simple, but affective, conventional authentication and the other to detect fingerprint, palm vein or face spoof using CNN. Through the proposed security framework, there is no chance of replica access of unwanted people to the system. In case of passing a single authentication layer of the system in a false manner, the malicious user is trapped by another phase, as evident from the experimental results. Although the involvement of multiple authentication layers improves the security of the system, its computational complexity is affected.

In future work, we aim to balance the security level and running time of the system to make it more suitable for several resource-constrained applications. We also plan to investigate the performance and suitability of computationally efficient CNN architectures, such as SqueezeNet.

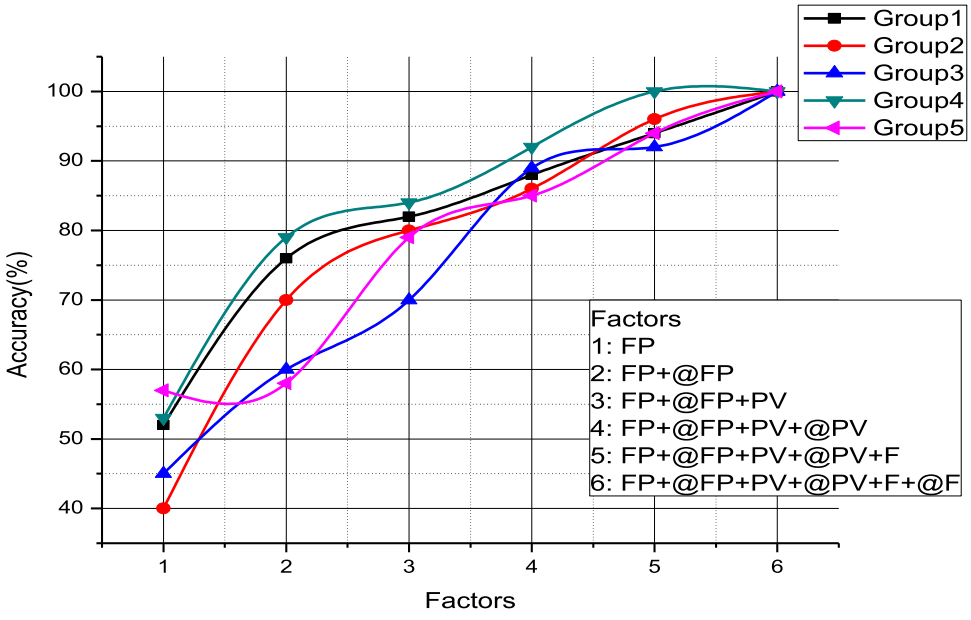

Fig. 7. Visualizing the effect of individual factors on accuracy.

Please cite this article as: M. Sajjad et al., CNN-based anti-spoofing two-tier multi-factor authentication system, Pattern Recognition Letters (2018), https://doi.org/10.1016/j.patrec.2018.02.015

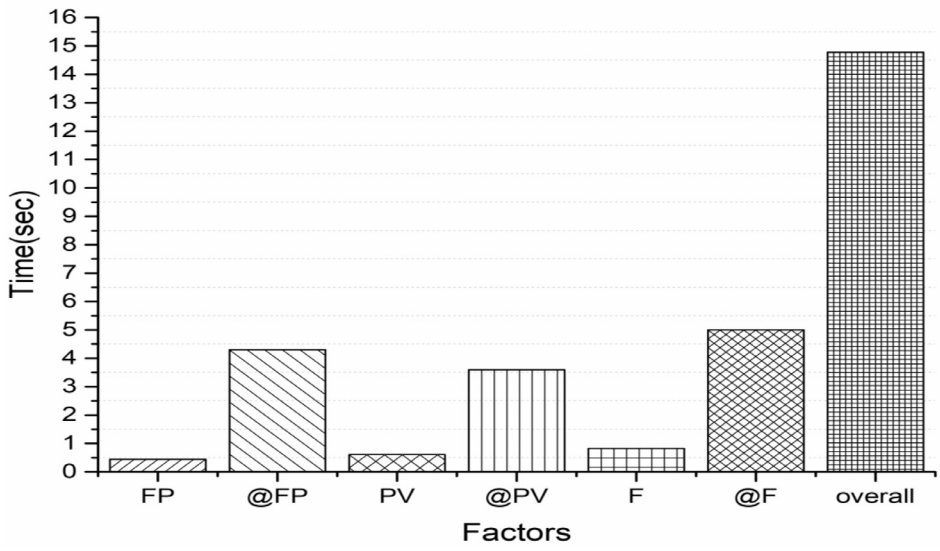

Fig. 8. Computational cost of the proposed system.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No.2016R1A2B4011712).

References

[1] Z. Akhtar, G.L. Foresti, Face spoof attack recognition using discriminative image patches, J. Electr. Comput. Eng. 2016 (2016).
[2] A.A.-A. Gutub, S. Arabia, Remodeling of elliptic curve cryptography scalar multiplication architecture using parallel jacobian coordinate system, Int. J. Comput. Sci. Secur. 4 (2010) 409.
[3] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H.H.G. Wang, S.W. Baik, Secure surveillance framework for IoT systems using probabilistic image encryption, IEEE Trans. Ind. Inf. (2018) 1–1, doi:10.1109/TII.2018.2791944.
[4] S.M. Al-Nofaie, M.M. Fattani, and A.A.-A. Gutub, "Merging two steganography techniques adjusted to improve arabic text data security."
[5] A.A.-A. Gutub, Pixel indicator technique for RGB image steganography, J. Emerg. Technol. Web Intell. 2 (2010) 56–64.
[6] K. Muhammad, J. Ahmad, H. Farman, Z. Jan, M. Sajjad, S.W. Baik, A secure method for color image steganography using gray-level modification and multi-level encryption, KSII Trans. Internet Inf. Syst. 9 (2015) 1938–1962.
[7] K. Muhammad, M. Sajjad, S.W. Baik, Dual-level security based cyclic18 steganographic method and its application for secure transmission of keyframes during wireless capsule endoscopy, J. Med. Syst. 40 (2016) 114.
[8] S. Liu, M. Silverman, A practical guide to biometric security technology, IT Profess. 3 (2001) 27–32.
[9] G. Lawton, Biometrics: a new era in security, Computer 31 (1998) 16–18.
[10] M. Sajjad, M. Nasir, K. Muhammad, S. Khan, Z. Jan, A.K. Sangaiah, et al., Raspberry Pi assisted face recognition framework for enhanced law-enforcement services in smart cities, Fut. Gen. Comput. Syst. (2017), doi:10.1016/j.future.2017.11.013.
[11] F. Wang, J. Han, Robust multimodal biometric authentication integrating iris, face and palmprint, Inf. Technol. Control 37 (2008).
[12] Y. Xu, A. Zhong, J. Yang, D. Zhang, Bimodal biometrics based on a representation and recognition approach, Opt. Eng. 50 (2011) 037202-037202-7.
[13] K.W. Bowyer, Face recognition technology: security versus privacy, IEEE Technol. Soc. Mag. 23 (2004) 9–19.
[14] A.K. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, IEEE Trans.Cir Syst. Video Technol. 14 (2004) 4–20.
[15] J.Y. Choi, W. De Neve, Y.M. Ro, K.N. Plataniotis, Automatic face annotation in personal photo collections using context-based unsupervised clustering and face information fusion, IEEE Trans. Circ. Syst. Video Technol. 20 (2010) 1292–1309.
[16] A. Gutub, N. Al-Juaid, E. Khan, Counting-based secret sharing technique for multimedia applications, Multimed. Tools Appl. (2017) 1–29.
[17] N.A. Al-Otaibi, A.A. Gutub, 2-leyer security system for hiding sensitive text data on personal computers, Lect. Note Inform. Theory 2 (2014).
[18] T. Ring, Spoofing: are the hackers beating biometrics? Biom. Technol. Today 2015 (2015) 5–9.
[19] Z. Boulkenafet, J. Komulainen, A. Hadid, Face spoofing detection using colour texture analysis, IEEE Trans. Inf. Forensics Secur. 11 (2016) 1818–1830.
[20] P. Tome, S. Marcel, On the vulnerability of palm vein recognition to spoofing attacks, in: Biometrics (ICB), 2015 International Conference on, 2015, pp. 319–325.
[21] J. Määttä, A. Hadid, M. Pietikäinen, Face spoofing detection from single images using texture and local shape analysis, IET Biom. 1 (2012) 3–10.
[22] L. Sun, G. Pan, Z. Wu, S. Lao, Blinking-based live face detection using conditional random fields, Adv. Biom. (2007) 252–260.
[23] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, in: Advances in Neural Information Processing Systems, 2012, pp. 1097–1105.
[24] K. Muhammad, J. Ahmad, S.W. Baik, Early fire detection using convolutional neural networks during surveillance for effective disaster management, Neurocomputing (2017), doi:10.1016/j.neucom.2017.04.083.
[25] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.
[26] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, et al., Decaf: A deep convolutional activation feature for generic visual recognition, in: International Conference on Machine Learning, 2014, pp. 647–655.
[27] A. Ullah, J. Ahmad, K. Muhammad, M. Sajjad, S.W. Baik, Action recognition in video sequences using deep Bi-directional LSTM with CNN features, IEEE Access 6 (2017) 1155–1166.
[28] M. Asim, Z. Ming, M.Y. Javed, CNN based spatio-temporal feature extraction for face anti-spoofing, in: Image, Vision and Computing (ICIVC), 2017 2nd International Conference on, 2017, pp. 234–238.
[29] A. Babenko, A. Slesarev, A. Chigorin, V. Lempitsky, Neural codes for image retrieval, in: European Conference on Computer Vision, 2014, pp. 584–599.
[30] M. De Marsico, M. Nappi, D. Riccio, J.-L. Dugelay, Moving face spoofing detection via 3D projective invariants, in: Biometrics (ICB), 2012 5th IAPR International Conference on, 2012, pp. 73–78.
[31] J. Li, Y. Wang, T. Tan, A.K. Jain, Live face detection based on the analysis of fourier spectra, in: Biometric Technology for Human Identification, 2004, pp. 296–304.
[32] Z. Xu, S. Li, W. Deng, Learning temporal features using LSTM-CNN architecture for face anti-spoofing, in: Pattern Recognition (ACPR), 2015 3rd IAPR Asian Conference on, 2015, pp. 141–145.
[33] F. Schroff, D. Kalenichenko, J. Philbin, Facenet: a unified embedding for face recognition and clustering, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 815–823.
[34] E. Fazl-Ersi, M.E. Mousa-Pasandi, R. Laganiere, M. Awad, Age and gender recognition using informative features of various types, in: Image Processing (icip), 2014 IEEE International Conference on, 2014, pp. 5891–5895.
[35] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2004) 91–110.
[36] E. Rublee, V. Rabaud, K. Konolige, G. Bradski, ORB: an efficient alternative to SIFT or SURF, in: Computer Vision (ICCV), 2011 IEEE International Conference on, 2011, pp. 2564–2571.
[37] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, et al., Going deeper with convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1–9.
[38] L. Ghiani, D. Yambay, V. Mura, S. Tocco, G.L. Marcialis, F. Roli, et al., Livdet 2013 fingerprint liveness detection competition 2013, in: Biometrics (ICB), 2013 International Conference on, 2013, pp. 1–6.
[39] J. Galbally, F. Alonso-Fernandez, J. Fierrez, J. Ortega-Garcia, A high performance fingerprint liveness detection method based on quality related features, Fut. Gen. Comput. Syst. 28 (2012) 311–321.

[40] I. Chingovska, A. Anjos, S. Marcel, On the effectiveness of local binary patterns in face anti-spoofing, in: Biometrics Special Interest Group (BIOSIG), 2012 BIOSIG-Proceedings of the International Conference of the, 2012, pp. 1–7.
[41] Z. Zhang, J. Yan, S. Liu, Z. Lei, D. Yi, S.Z. Li, A face antispoofing database with diverse attacks, in: Biometrics (ICB), 2012 5th IAPR International Conference on, 2012, pp. 26–31.
[42] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, et al., Caffe: Convolutional architecture for fast feature embedding, in: Proceedings of the 22nd ACM International Conference on Multimedia, 2014, pp. 675–678.
[43] F. Chen, X. Huang, J. Zhou, Hierarchical minutiae matching for fingerprint and palmprint identification, IEEE Trans. Image Process. 22 (2013) 4964–4971.
[44] Y. Duan, J. Lu, J. Feng, J. Zhou, Context-aware local binary feature learning for face recognition, IEEE Trans. Pattern Anal. Mach. Intell. (2017).
[45] D. Menotti, G. Chiachia, A. Pinto, W.R. Schwartz, H. Pedrini, A.X. Falcao, et al., Deep representations for iris, face, and fingerprint spoofing detection, IEEE Trans. Inf. Forens. Secur. 10 (2015) 864–879.

Please cite this article as: M. Sajjad et al., CNN-based anti-spoofing two-tier multi-factor authentication system, Pattern Recognition Letters (2018), https://doi.org/10.1016/j.patrec.2018.02.015