BioTouchPass2: Touchscreen Password Biometrics Using Time-Aligned Recurrent Neural Networks

Ruben Tolosana, Ruben Vera-Rodriguez, Julian Fierrez, *Member, IEEE*, and Javier Ortega-Garcia, *Fellow, IEEE*

Abstract—Passwords are still used on a daily basis for all kind of applications. However, they are not secure enough by themselves in many cases. This work enhances password scenarios through two-factor authentication asking the users to draw each character of the password instead of typing them as usual. The main contributions of this study are as follows: *i)* We present the novel MobileTouchDB public database, acquired in an unsupervised mobile scenario with no restrictions in terms of position, posture, and devices. This database contains more than 64K on-line character samples performed by 217 users, with 94 different smartphone models, and up to 6 acquisition sessions. *ii)* We perform a complete analysis of the proposed approach considering both traditional authentication systems such as Dynamic Time Warping (DTW) and novel approaches based on Recurrent Neural Networks (RNNs). In addition, we present a novel approach named Time-Aligned Recurrent Neural Networks (TA-RNNs). This approach combines the potential of DTW and RNNs to train more robust systems against attacks. A complete analysis of the proposed approach is carried out using both MobileTouchDB and e-BioDigitDB databases. Our proposed TA-RNN system outperforms the state of the art, achieving a final 2.38% Equal Error Rate, using just a 4-digit password and one training sample per character. These results encourage the deployment of our proposed approach in comparison with traditional typed-based password systems where the attack would have 100% success rate under the same impostor scenario.

Index Terms—Biometrics, passwords, handwriting, touch biometrics, TA-RNNs, RNN, DTW, MobileTouchDB, e-BioDigitDB.

I. INTRODUCTION

MOBILE devices have become an indispensable tool for most people nowadays [1]. The rapid and continuous deployment of mobile devices around the world has been motivated not only by the high technological evolution and new features incorporated but also to the new internet infrastructures like 5G that allows the communication and use of social media in real time, among many other factors. In this way, both public and private sectors are aware of the importance of mobile devices for the society and are trying to deploy their services through user-friendly mobile applications ensuring data protection and high security.

Passwords are still the most common way to authenticate users nowadays. They can range from Personal Identification Numbers (PIN) that require users to memorise them to One-Time Passwords (OTP) where the security system is in charge of selecting and providing to the user a different password each time it is required, e.g., sending messages to personal mobile devices or special tokens. We use passwords on a daily basis for all kinds of applications. However, are passwords secure enough? Apparently not, at least by themselves. Recent news put in evidence this fact, e.g., in January 2019 a total of 21 million passwords from all parts of the world were released together with their corresponding emails addresses [2]. This important problem is related not only to data breaches, but also to many other attack scenarios, as it has been pointed out in different studies [3], [4]. First, it is common to use passwords based on sequential digits (e.g., *"1 2 3 4 5 6"*), personal information such as birth dates, or simply words such as *"password"* or *"qwerty"* that are very easy to guess [5]. Second, passwords that are typed on mobile devices such as tablets or smartphones are susceptible to "smudge attacks", i.e., the deposition of finger grease traces on the touchscreen can be used by the impostors to guess lock patterns or passwords [6]. Finally, password-based authentication is also vulnerable to "shoulder surfing". This type of attack is produced when the impostor can observe directly or use external recording devices to collect the user information. This attack has attracted the attention of many researchers in recent years due to the increased deployment of handheld recording devices and public surveillance infrastructures [7], [8]. So, if we know that traditional passwords are not secure enough by themselves, but they continue to be present in our lives, how can we improve this authentication scenario?

Two-factor authentication approaches have been very successful in the last years. These approaches are based on the combination of two authentication stages. For example: *i)* the security platform sends the password to the personal email or mobile number of the claimed user, and *ii)* the claimed user introduces this password for the final verification. Behavioral biometric information has also been considered for two-factor authentication approaches. In [9], the authors proposed a two-factor verification system based on
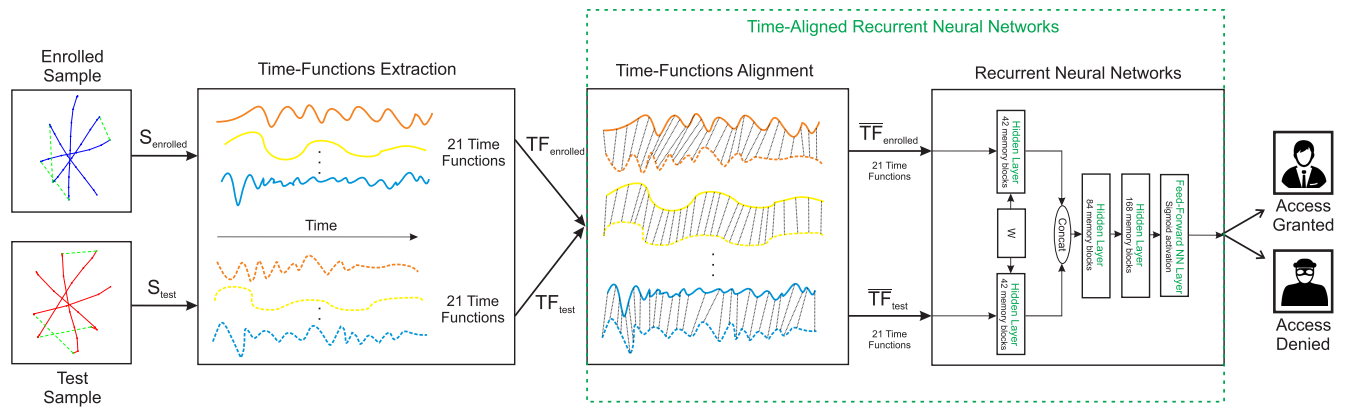
Fig. 1. Architecture of our proposed touchscreen biometric system based on time-aligned recurrent neural networks. S denotes one character sample, and TF and $\overline{TF}$ the original and pre-aligned 21 time functions, respectively. The recurrent neural networks block is enlarged in Fig. 2 for a better understanding.

timing-related features for dynamic lock patterns, achieving a final average Equal Error Rate (EER) of 10.39% for imitation attacks. A similar two-factor authentication approach was proposed in [10] for traditional Android unlock patterns but considering biometric dynamic features related to the position of the finger, pressure, finger size, and accelerometer sensor achieving a final 15.0% EER for imitation attacks. Two-factor authentication approaches have also been expanded to physiological biometric traits. In [11], Jenkins *et al.* proposed a system based on features extracted from periocular images acquired using an iPhone 5, achieving very good results for the task of identification.

In this article we propose two-factor authentication approaches based on the incorporation of touch biometrics to password authentication systems, asking the users to draw each character of the password on the touchscreen instead of typing them as usual. One example of use that motivates our proposed approach is on internet payments with credit cards. Banks usually send a password (typically between 6 and 8 characters) to the user. This password must be inserted by the user in the security platform in order to complete the payment. Our proposed approach enhances such scenario by including a second authentication factor based on touch interaction biometrics.

The main contributions of this study are related to the novel MobileTouchDB database, our proposed architecture, and the competitive results obtained with respect to related research:

- We present and describe the acquisition process of the new MobileTouchDB database. This database contains more than 64K on-line character samples performed by 217 users, using 94 different smartphone models. Mobile-TouchDB considers an unsupervised mobile scenario with a maximum of 6 captured sessions per subject and it is publicly available in GitHub.[1]
- The MobileTouchDB database opens the doors to many different applications: *i)* analyse the discriminative power of novel human touch interaction dynamics, *ii)* enhance traditional password authentication systems through the incorporation of touch biometric information as a sec-

ond level of user authentication, and *iii)* analyse the way we interact with mobile devices on a daily basis in order to enhance continuous authentication systems.
- We present a novel approach named Time-Aligned Recurrent Neural Networks (TA-RNNs).[2] Fig. 1 represents the general architecture of our proposed approach. It combines the potential of Dynamic Time Warping (DTW) and Recurrent Neural Networks (RNNs) to train more robust systems against attacks.
- We perform a complete analysis of our proposed approach using both MobileTouchDB and e-BioDigitDB public databases. Three experiments are considered: *i)* one-character analysis in order to evaluate the discriminative power of each character, *ii)* character combination analysis so as to measure the robustness of our proposed approach when increasing the length of the passwords from 1 to 9 characters, and *iii)* template update analysis.
- We compare our proposed TA-RNN system with both traditional and state-of-the-art authentication systems. Our proposed TA-RNN system outperforms the state of the art, achieving a final 2.38% EER, using just a 4-digit password and one training sample per character.
- We demonstrate the application of TA-RNNs for other time sequence tasks, i.e., on-line handwritten signature verification, outperforming in large margin the state of the art as well.

MobileTouchDB can be also useful for other research lines, e.g.: *i)* user-dependent effects [12], and development of user-dependent methods for handwriting recognition [13], *ii)* the neuromotor processes involved in writing over touchscreens [14], [15], *iii)* sensing factors in obtaining representative and clean touch interaction signals [16], [17], *iv)* human-device interaction factors involving touchscreen signals [18], [19], and development of improved interaction methods, and *v)* population statistics around touch interaction signals, and development of new methods aimed at recognising or serving particular population groups [20].

---

[1] https://github.com/BiDAlab/MobileTouchDB

[2] Spanish Patent Application (P202030060)

TABLE I

COMPARISON OF DIFFERENT TOUCHSCREEN PASSWORD BIOMETRIC APPROACHES FOR MOBILE SCENARIOS

Study	Classifiers	Verification Performance			# Participants (Database)
		Best Result	Password Length	Training Samples	
Kutzner *et al.* (2015) [21]	Bayes-Nets, KStar, kNN	FAR = 10.42% FRR = unknown	8	12	32 (Unavailable DB)
Nguyen *et al.* (2017) [22]	DTW	EER = 4.84%	4	5	20 (Unavailable DB)
Tolosana *et al.* (2019) [23]	DTW, RNNs	EER = 3.80%	7	3	93 (e-BioDigitDB)
Le Lan *et al.* (2019) [24]	RNNs	EER = 4.90%	4	4	93 (e-BioDigitDB)
Tolosana *et al.* (2019) [25]	DTW	EER = 5.90%	9	1	217 (MobileTouchDB)
Proposed Approach	**TA-RNNs**	**EER = 2.38%**	**4**	**1**	**217 (MobileTouchDB)**

The remainder of the paper is organised as follows. Sec. II summarises related works in touchscreen password biometrics. Sec. III describes our proposed TA-RNN biometric system. Sec. IV describes both e-BioDigitDB and the novel Mobile-TouchDB database considered in the experiments of this article. Sec. V and VI describe the experimental protocol and results achieved using our proposed approach, respectively. Finally, Sec. VIII draws the final conclusions and points out some future work lines.

II. RELATED WORKS

Touch biometrics are becoming a very attractive way to verify users on mobile devices [19], [26]. Table I summarises the most relevant approaches in the area of touchscreen password biometrics on mobile scenarios. For each study, we include information related with classifiers and databases considered. We also report in Table I the best verification performance results in addition to other important system settings such as the length of the password, and the number of training samples. Regarding the attack scenario, all studies consider the case in which impostors know the password of the user, i.e., an *imitation attack* [27].

In [21], Kutzner *et al.* asked the users to perform an 8-digit password on the screen of a tablet device. For each handwritten password, a total of 25 static and dynamic features were extracted and tested using different authentication algorithms such as Bayes-Nets, KStar, and kNN. Their proposed approach achieved a final 10.42% False Acceptance Rate (FAR). The False Rejection Rate (FRR) is not available. However, the authentication scenario considered in this preliminary approach restricted the deployment of the technology in real mobile applications as *i)* the authors considered a large number of training samples (i.e., 12), and *ii)* it seems to be only applicable to devices with large screens (such as tablets) as it would be very difficult for the users to perform such a long password (8 characters) on a screen of much smaller size. In [22], Nguyen *et al.* evaluated the use of handwritten touch biometrics for PIN-based authentication systems. Their proposed authentication approach overcame some of the drawbacks previously cited as they asked users to draw each character of a 4-digit PIN one by one. A final 4.84%

EER was achieved using 5 training samples, and a biometric system composed of 5 dynamic features and DTW algorithm. In [23], we released e-BioDigitDB, the first touchscreen password biometric public database, containing only numbers from 0 to 9. In addition, we reported a benchmark evaluation of biometric authentication using two different state-of-the-art approaches: *i)* DTW, and *ii)* RNNs. Despite the good results obtained using RNNs (outperforming DTW for 50% of the numerical characters), DTW still improved the RNN results with an average 0.5% EER absolute improvement. In [24], Le Lan *et al.* also evaluated RNN approaches using the e-BioDigitDB database. Their proposed approach achieved a final 4.90% EER using 4-digit passwords and 4 training samples per digit. Finally, a preliminary study of the work presented here was published in [25]. In that work we performed an initial analysis of MobileTouchDB considering only a DTW baseline system so as to provide an easily reproducible framework.

The study presented here extends the preliminary analysis carried out in [25]. The main improvements over [25] are:

- Sec. II has been included to survey and compare advantages and limitations of recent research on touchscreen password biometrics on mobile scenarios.
- Our preliminary touch biometric system in [25] based on DTW has been extended by incorporating RNNs and presenting a novel approach not published yet, named TA-RNNs (Sec. III). This approach combines the potential of DTW and RNNs with a Siamese architecture to train more robust systems against attacks.
- The results achieved in the present study outperform the results achieved in [25], and also the state of the art, achieving a final 2.38% EER. This improvement is achieved considering more user-friendly scenarios, i.e., using just a 4-digit password and one training sample per character.

III. PROPOSED APPROACH

This section describes our proposed Time-Aligned Recurrent Neural Networks for touchscreen password biometrics. A graphical representation is included in Fig. 1.

TABLE II
SET OF TIME FUNCTIONS CONSIDERED IN THIS WORK

#	Feature
1	X-coordinate: x_n
2	Y-coordinate: y_n
3	Path-tangent angle: θ_n
4	Path velocity magnitude: v_n
5	Log curvature radius: ρ_n
6	Total acceleration magnitude: a_n
7-12	First-order derivative of features 1-6: $\dot{x}_n, \dot{y}_n, \dot{\theta}_n, \dot{v}_n, \dot{\rho}_n, \dot{a}_n$
13-14	Second-order derivative of features 1-2: $\ddot{x}_n, \ddot{y}_n$
15	Ratio of the minimum over the maximum speed over a 5-samples window: v_n^r
16-17	Angle of consecutive samples and first-order derivative: $\alpha_n, \dot{\alpha}_n$
18	Sine of the angle of consecutive samples: s_n
19	Cosine of the angle of consecutive samples: c_n
20	Stroke length to width ratio over a 5-samples window: r_n^5
21	Stroke length to width ratio over a 7-samples window: r_n^7

A. Time-Functions Extraction

Our proposed touchscreen password biometric system is based on time functions. For each character sample acquired (i.e., $S_{enrolled}$ and S_{test} in Fig. 1), signals related to X and Y spatial coordinates are used to extract a set of 21 time functions (i.e., $TF_{enrolled}$ and TF_{test} in Fig. 1), similar to [23]. This set of 21 time functions is widely used in on-line signature verification [28], [29]. The complete set of time functions is described in Table II.

B. Time-Functions Alignment

One crucial point when comparing the similarity among time sequences is the proper alignment of them prior to calculating the similarity score through distance measurement functions (e.g., the Euclidean distance). DTW is one of the most popular algorithms in the literature, in particular for signature biometrics [30]. The goal of DTW is to find the optimal warping alignment path of a pair of time sequences A and B that minimises a given distance measure. The algorithm can be defined as follows. Let's define two sequences as:

$$A = a_1, a_2, \ldots, a_n, \ldots, a_N$$
$$B = b_1, b_2, \ldots, b_m, \ldots, b_M \quad (1)$$

and a distance measure among sequence samples as:

$$d(a_n, b_m) = (a_n - b_m)^2 \quad (2)$$

A warping path can be defined as:

$$C = c_1, c_2, \ldots, c_k, \ldots, c_K \quad (3)$$

where each c_k represents a correspondence (n, m) between samples of A and B. The initial condition of the algorithm is set to:

$$g_1 = g(1, 1) = d(a_1, b_1) \cdot w(1) \quad (4)$$

where g_k represents the accumulated distance after k steps and $w(k)$ is a weighting factor that must be defined. For each iteration, g_k is computed as:

$$g_k = g(n, m) = \min_{c_{k-1}} \left[g_{k-1} + d(c_k) \cdot w(k) \right] \quad (5)$$

until the N'th and M'th sample of both sequences respectively is reached. The resulting normalised distance is:

$$D(A, B) = \frac{g_K}{\sum_{k=1}^{K} w(k)} \quad (6)$$

where $\sum w(k)$ compensates the effect of the length of the sequences. The weighting factors $w(k)$ are defined in order to restrict which correspondences among samples of both sequences are allowed. In this work, only three transitions with the same value equal to 1 are allowed for the computation of g_k, which is the most common implementation found in the literature. Consequently, Eq. (5) becomes:

$$g_k = g(n, m) = \min \begin{bmatrix} g(n, m-1) + d(a_n, b_m) \\ g(n-1, m-1) + d(a_n, b_m) \\ g(n-1, m) + d(a_n, b_m) \end{bmatrix} \quad (7)$$

In this study we consider an improved version of the basic DTW algorithm named Sliding Window Dynamic Time Warping (SW-DTW), proposed in [31]. This updated version overcomes some of the incorrect alignments generated by the original DTW. This is produced due to the original DTW considers an element-to-element distance. SW-DTW modifies the original distance in Eq. (2) to consider the context by incorporating a weighted average of the neighbouring distances.

In our proposed approach, SW-DTW is applied in a first stage in order to convert the 21 original time functions (i.e., $TF_{enrolled}$ and TF_{test} in Fig. 1) into 21 pre-aligned time functions (i.e., $\overline{TF}_{enrolled}$ and $\overline{TF}_{test}$ in Fig. 1) before introducing them to the RNNs. This way our proposed RNN system is able to extract more meaningful features as time sequences have been previously normalised through the optimal warping path.

C. Recurrent Neural Networks

New trends based on the use of RNNs, which is a specific neural network architecture, are becoming more and more important nowadays for modelling sequential data with arbitrary length. In this study we adapt the original verification system proposed in [28] for on-line handwritten signature to touchscreen password biometrics. In [28] we proposed RNN systems based on a Siamese architecture. The main goal was to learn a dissimilarity metric from data minimising a discriminative cost function that drives the dissimilarity metric to be small for pairs of genuine samples from the same subject, and higher for pairs of samples coming from different subjects. In particular, we consider in this study a Bidirectional Long Short-Term Memory (BLSTM) network, which allows access to past, present, and future context, achieving much better results compared with the original LSTM.

Fig. 2 shows our proposed writer-independent BLSTM touch biometric system based on a Siamese architecture. For the input of the system, we feed the network with as much
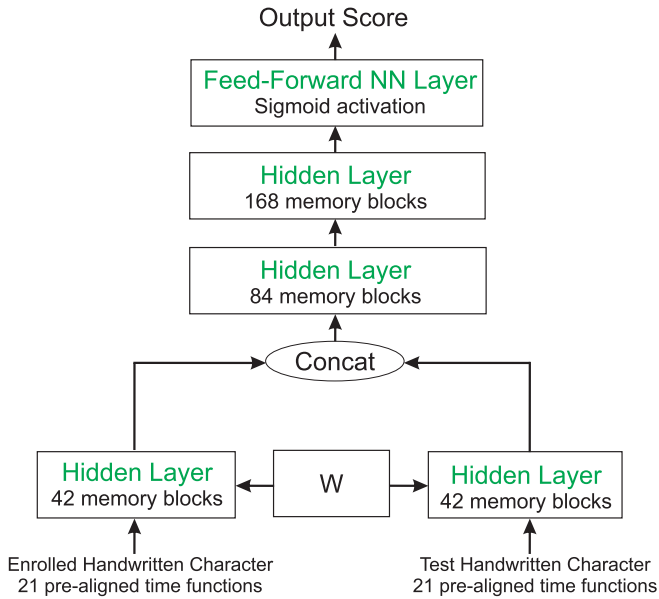
Fig. 2.   Proposed writer-independent BLSTM Siamese architecture.

information as possible, i.e., all 21 pre-aligned time functions per character previously normalised through SW-DTW. The first layer is composed of two BLSTM hidden layers with 42 memory blocks each, sharing the weights between them. The outputs of the first two parallel BLSTM hidden layers are concatenated and serve as input to the second layer, which corresponds to a BLSTM hidden layer with 84 memory blocks. Then, we add a third BLSTM hidden layer with 168 memory blocks. The activation functions considered in all BLSTM memory blocks are based on the standard approach [28]. Finally, a feed-forward neural network layer with a sigmoid activation is considered, providing an output score for each pair of characters. It is important to highlight that our approach is trained to distinguish between genuine and impostor patterns from all characters and users. Thus, we just train one writer-independent system for all characters and users through a development dataset.

IV. DATABASES

A. e-BioDigitDB

e-BioDigitDB database[3] is composed of 93 users and was originally captured in order to perform a preliminary study of handwritten passwords for touchscreen biometrics [23]. This database comprises on-line handwritten numerical characters from 0 to 9 acquired using a Samsung Galaxy Note 10.1 general purpose tablet. This device has a 10.1-inch LCD display with a resolution of 1280×800 pixels.

Regarding the acquisition protocol, subjects had to perform handwritten numerical characters one at a time, in a supervised and office-like scenario. Samples were collected in two sessions with a time gap of at least three weeks between them in order to consider inter-session variability, very important for behavioral biometric traits [32]. For each session, users had to perform a total of 4 numerical sequences from 0 to 9 using

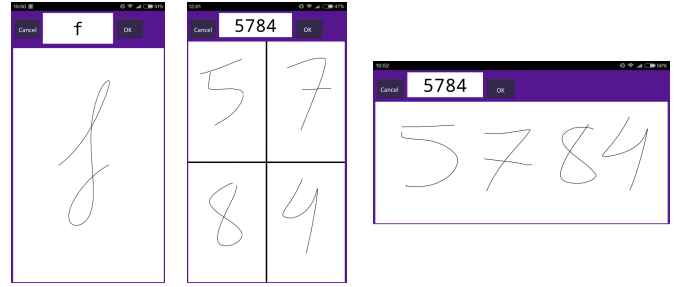[3]https://github.com/BiDAlab/eBioDigitDB

Fig. 3.   Different interfaces designed for the acquisition app. Both portrait and landscape orientations are considered in order to analyse different user experiences while drawing.

the finger as input. Therefore, there are a total of 8 samples per numerical digit and user.

B. MobileTouchDB

MobileTouchDB is a novel handwritten character mobile touch biometric database composed of more than 64K on-line character samples performed by 217 users. For the acquisition, we implemented an Android application. Fig. 3 represents the different interfaces designed for the acquisition. All interfaces are composed of: *i)* the character/password to draw (top, middle) and two buttons "OK" (top, right) and "Cancel" (top, left) to press after drawing if the sample was good or bad respectively. If the sample was not good, then it was repeated; and *ii)* a rectangular area to perform the character or password. In order to study an unsupervised mobile scenario, the acquisition app was uploaded to the Play Store. This way all participants could download and use the app on their own devices without any kind of supervision, simulating a practical scenario in which users can generate touchscreen information in any possible scenario, e.g., standing, sitting, walking, indoors, outdoors, etc. As a result, 94 different models from the following 16 brands were collected during the acquisition: *Alcatel, Blackberry, BQ, Coolpad, Doogee, Google, Huawei, LeTV, LG, Motorola, OnePlus, Samsung, Sony, UMIDIGI, Xiaomi, and ZTE*. The acquisition app was designed to capture the following time signals: *X* and *Y* spatial coordinates, the area covered by the finger, timestamp, accelerometer, and gyroscope. However, information related to the area covered by the finger, accelerometer, and gyroscope was not available in some cases depending on how old the acquisition device was.

The acquisition protocol considered in MobileTouchDB is depicted in Fig. 4. It comprises a total of 6 sessions (i.e., S1-S6) with different time gaps among them. It is important to highlight that in all sessions, the time gap refers to the minimum time between one user finishes a session and the following session is available. However, participants usually performed their corresponding sessions later on thanks to notifications sent automatically by the acquisition app to the users. Regarding the data acquired, each session comprises 8 different capturing blocks (i.e., from Block1 to Block8). Fig. 5 shows some examples of each of the eight acquisition blocks for two different users (indicated in blue and red colours). The green dashed lines indicate pen up trajectories between
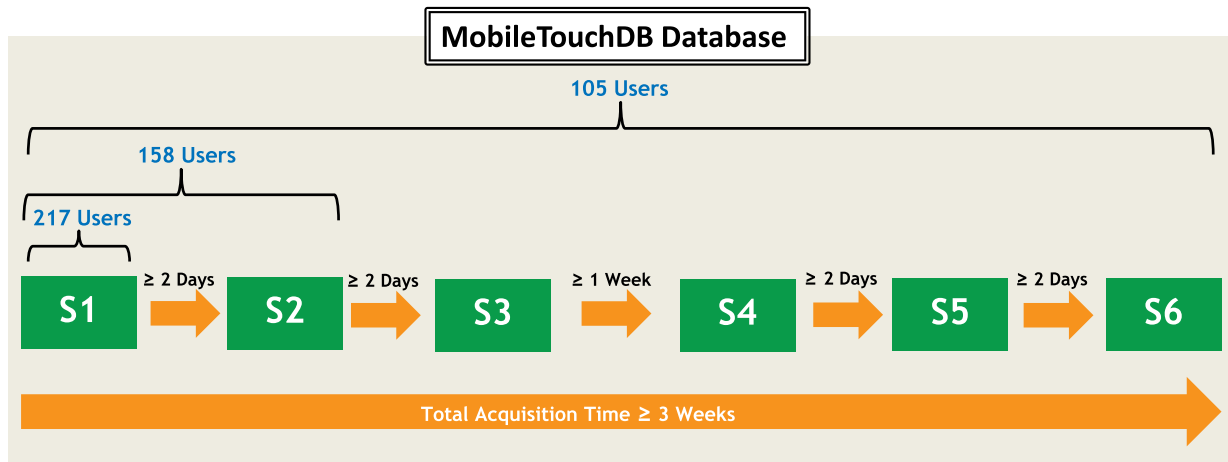
Fig. 4. Description of the design and number of available users of the new MobileTouchDB database.

strokes. In Block1, we asked users to draw all numbers (from 0 to 9). Block2 and Block3 comprise upper- and lower-case letters respectively, with a total of 27 letters each. Block4 is composed of 8 different symbols (i.e., *"?", "#", "*", "@", "%", "=", "ϵ", and "α"*). It is important to remark that inside each block, characters were randomised before asking users to draw them. This way, each user performs a different character sequence in each session. From Block1 to Block4, the acquisition interface was designed as portrait to provide a better user experience (see Fig. 3, left). After finishing the first 4 blocks focused on performing one single character at a time (one sample per character), we asked users to draw passwords composed of 4 numbers (always *"5 7 8 4"*) in different ways (6 samples in total). In Block5, users performed the password twice using a landscape orientation interface (see Fig. 3, right). We provided the users with a graphical visualization of the numbers while drawing them (i.e., visible mode). Then, in Block6, users had to repeat once the same task considered in Block5 but this time in an invisible mode, i.e., we did not provide to the users any visualization of the numbers while drawing them. The main motivation of this novel acquisition scenario is to protect us against shoulder surfing attacks, as commented in [22]. In Block 7, users had to draw each number of the password inside each of the four available boxes (two times), considering first a visible mode (see Fig. 3, middle). Finally, in Block8 users had to repeat once the same task considered in Block7 but this time in an invisible mode. In both Block7 and Block8 the acquisition interface was kept portrait to analyse the user experience in different settings.

Regarding the MobileTouchDB population statistics, 217 users completed the S1 acquisition session. S1 and S2 were completed by 159 users. Finally, a total of 105 users completed the six acquisition sessions. This participant reduction between S1 and S6 sessions is produced due to the challenging acquisition scenario considered in this study as it was completely unsupervised and comprised several acquisition sessions along time. Regarding the age distribution, 36.2% of the participants are younger than 22 years old, 31.9% are between 22 and 27 years old, and the remaining 31.9% are older than 27 years old. Regarding the gender, 63% of the participants

were males, and 37% females. 96% of the population was righthanded.

V. EXPERIMENTAL PROTOCOL

Three different experiments are considered: *i)* one-character analysis in order to evaluate the discriminative power of each character, *ii)* character combination analysis so as to measure the robustness of our proposed approach when increasing the length of the passwords from 1 to 9 characters, and *iii)* analysis of the system performance when samples performed in different acquisition sessions are considered for training. Due to the large amount of information acquired in MobileTouchDB, in this paper we focus on characters performed one at a time. Complete passwords acquired from Block5 to Block8 will be analysed in future studies.

Both e-BioDigitDB and MobileTouchDB are divided into development and evaluation datasets. For the e-BioDigitDB, we consider the same experimental protocol proposed in [23]. Thus, the first 50 users are considered for development whereas the remaining 43 users are left for the final evaluation of the systems. For MobileTouchDB, users are divided into development (175, ≈80%) and evaluation (42, ≈20%) datasets. Finally, the development dataset of both databases is also split into training (≈80%) and validation (≈20%) datasets in order to train the weights of the neural networks and select the best models. It is important to remark that for MobileTouchDB, all users included in the evaluation dataset (i.e., 42) contain the maximum number of acquisition sessions, i.e., 6. This way we can perform a fair evaluation of our trained models on different experimental conditions.

For the development of our proposed handwritten touch biometric systems (Sec. VI-A), Z genuine samples per character from the first session (i.e., 1 for MobileTouchDB and up to 4 for e-BioDigitDB) are used as enrolment samples, whereas the remaining genuine samples from the second session are used for testing. This way we consider the inter-session variability problem as genuine samples from different acquisition sessions are used as enrolment and testing samples, respectively. Finally, impostor scores are obtained by
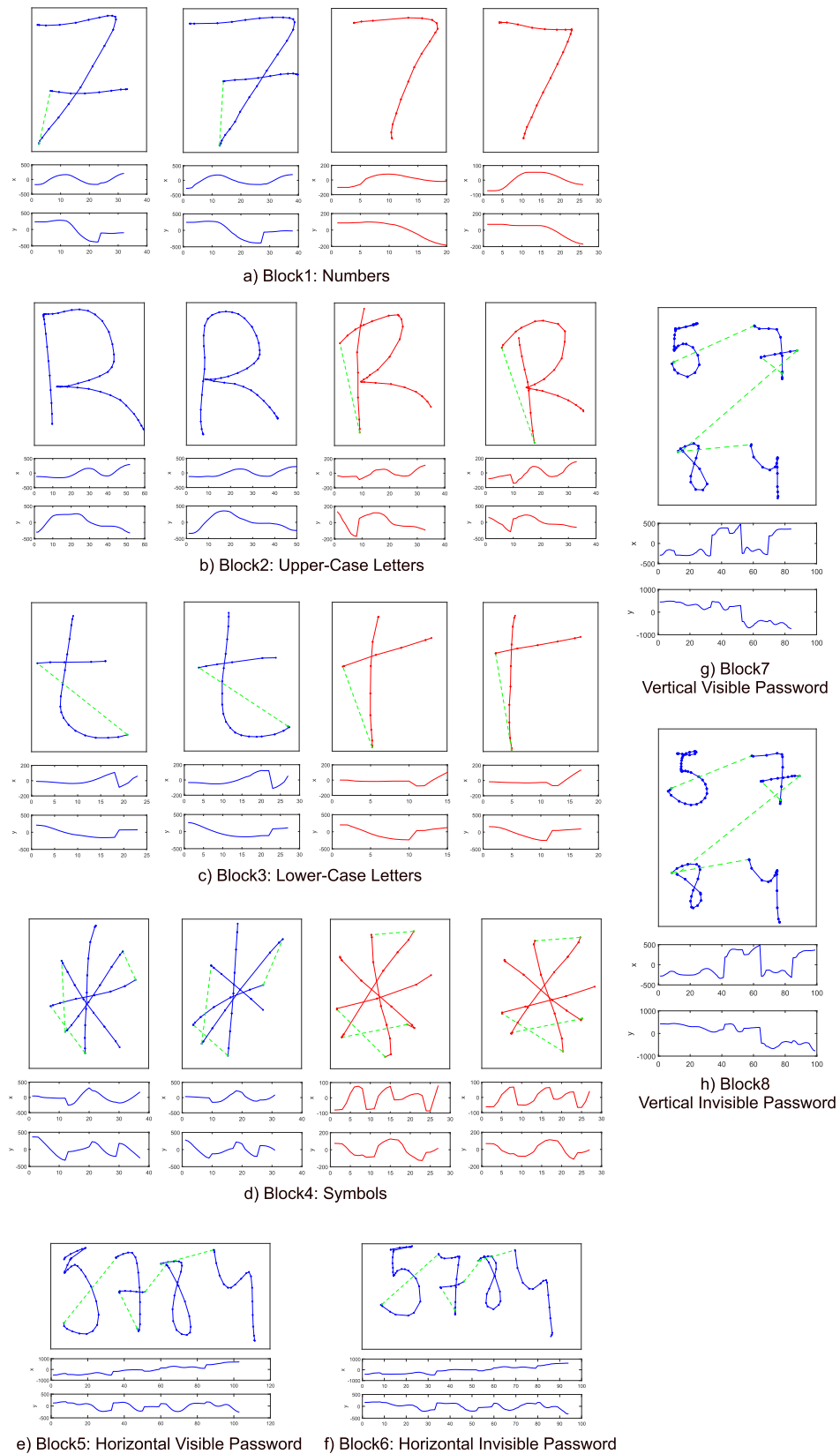
Fig. 5. Example of the data collected in MobileTouchDB database. Blue and red colours represent samples drawn by different users. The green dashed lines indicate pen up trajectories between strokes. Curves under each character represent *X* and *Y* trajectories over time.

TABLE III
Performance as EER(%) of Each Individual Upper- and Lower-Case Letter on the Evaluation Dataset of MobileTouchDB

	DTW	SW-DTW	RNNs	TA-RNNs		DTW	SW-DTW	RNNs	TA-RNNs
A	30.95	35,71	28,57	**23,81**	a	30,95	28,57	**16,67**	19,05
B	23.81	23,81	16,67	**11,90**	b	23,81	19,05	19,05	**16,67**
C	35.71	33,33	**23,81**	**23,81**	c	40,48	38,10	28,57	**26,19**
D	33.33	23,81	23,81	**21,43**	d	19,05	19,05	19,05	**14,29**
E	28,57	23,81	26,19	**21,43**	e	35,71	30,95	26,19	**21,43**
F	23,81	28,57	30,95	**19,05**	f	**19,05**	**19,05**	**19,05**	**19,05**
G	26,19	25,89	21,43	**16,67**	g	21,43	23,81	19,05	**16,67**
H	26,19	21,43	19,05	**14,29**	h	28,57	30,95	23,81	**16,67**
I	**23,81**	26,19	26,19	**23,81**	i	30,95	33,33	**19,05**	21,43
J	21,43	19,05	23,81	**16,67**	j	26,19	23,81	**19,05**	21,43
K	28,57	26,19	23,81	**16,67**	k	16,67	16,67	16,67	**14,29**
L	30,95	30,95	**23,81**	28,57	l	28,57	28,57	19,05	**16,67**
M	26,19	23,81	23,81	**19,05**	m	28,57	21,43	21,43	**16,67**
N	33,33	30,95	26,19	**23,81**	n	28,57	23,81	19,05	**16,67**
Ñ	30,95	28,57	21,43	**19,05**	ñ	23,81	23,81	**16,67**	**16,67**
O	26,19	26,19	**21,43**	**21,43**	o	26,19	23,81	26,19	26,19
P	28,57	28,57	**21,43**	23,81	p	33,33	33,33	**21,43**	**21,43**
Q	30,95	26,19	**23,81**	**23,81**	q	23,81	19,05	**16,67**	**16,67**
R	23,81	21,43	21,43	**16,67**	r	19,05	19,05	16,67	**11,90**
S	28,57	28,57	**21,43**	**21,43**	s	35,71	33,33	23,81	**21,43**
T	23,81	21,43	**16,67**	21,43	t	23,81	21,43	21,43	**19,05**
U	35,71	35,71	26,19	**23,81**	u	28,57	26,19	**19,05**	**19,05**
V	38,10	35,71	33,33	**30,95**	v	23,81	16,67	**16,67**	19,05
W	26,19	21,43	19,05	**14,29**	w	28,57	26,19	**19,05**	**19,05**
X	**19,05**	21,43	21,43	21,43	x	21,43	**19,05**	**19,05**	**19,05**
Y	33,33	30,95	26,19	**23,81**	y	**19,05**	**19,05**	21,43	**19,05**
Z	33,33	30,95	28,57	**23,81**	z	33,33	33,33	21,43	**19,05**
Average	28.57	27.06	23.72	**20.99**	Average	26.63	24.87	20.20	**18.70**

comparing the *Z* enrolment samples with one genuine sample of each of the remaining users of the same database (simulating this way the imitation attack in which the impostor knows the password).

For the final evaluation of our proposed touch biometric systems, different scenarios are generally considered regarding the number of available enrolment samples per user (i.e., *Z*vs1), in which the final score is performed as the average score of *Z* one-to-one comparisons. In addition, for the character combination analysis (Sec. VI-B), the final score is produced by fusing the different one by one character score comparisons using the sum of the scores [13].

VI. Experimental Results

A. One-Character Analysis

This section analyses the potential of each individual character for the task of user authentication. Four different touch biometric systems are considered in the analysis: *i)* the original DTW system considered in previous studies [23], [25], *ii)* the recent SW-DTW approach that overcomes the incorrect alignments generated by the original DTW [31], *iii)* the state-of-the-art RNN Siamese architecture proposed in [23], and finally *iv)* the TA-RNN Siamese architecture proposed in this

work that combines SW-DTW alignment with RNNs. For the development of the systems, we follow the experimental protocol described in Sec. V. Samples from all characters of the development datasets of both MobileTouchDB and e-BioDigitDB databases are considered together during training as we intend to distinguish between genuine and impostor handwritten samples regardless of the user and the character. This approach resulted in better generalisation results compared with the case of training one system per user and character. Therefore, both RNNs and TA-RNNs systems are trained considering two different cases: *i)* pairs of genuine characters drawn by the same user, and *ii)* pairs of genuine and impostor characters, one performed by the claimed user and the other one by the impostor.

Both RNN and TA-RNN systems have been implemented under Keras framework using Tensorflow as back-end, with a NVIDIA GeForce RTX 2080 Ti GPU. The weights of the BLSTM and feed-forward layers are initialised by random values drawn from the zero-mean Gaussian distribution with standard deviation 0.05. Adam optimiser is considered with default parameters (learning rate of 0.001) and a loss function based on binary cross-entropy

Tables III and IV show the results of each character over the evaluation dataset of MobileTouchDB, grouped according to

TABLE IV

PERFORMANCE AS EER(%) OF EACH INDIVIDUAL NUMBER AND SYMBOL ON THE EVALUATION DATASET OF MOBILETOUCHDB

	DTW	SW-DTW	RNNs	TA-RNNs
0	28,57	23,81	**21,43**	**21,43**
1	28,57	30,95	**21,43**	**21,43**
2	28,57	30,95	23,81	**16,67**
3	28,57	28,57	28,57	**21,43**
4	28,57	30,95	23,81	**21,43**
5	28,57	30,95	26,19	**19,05**
6	30,95	28,57	28,57	**23,81**
7	28,57	28,57	23,81	**19,05**
8	23,81	21,43	19,05	**14,29**
9	28,57	30,95	26,19	**23,81**
Average	28.33	28.57	24.29	**20.24**
?	**21,43**	**21,43**	**21,43**	28,57
#	26,19	21,43	26,19	**21,43**
*	30,95	23,81	**21,43**	26,19
@	23,81	19,05	21,43	**14,29**
%	23,81	23,81	23,81	**21,43**
=	40,48	38,10	**28,57**	30,95
ϵ	30,95	30,95	26,19	**23,81**
α	26,19	26,19	**19,05**	21,43
Average	27.98	25.60	**23.51**	**23.51**

TABLE V

PERFORMANCE AS EER(%) OF EACH INDIVIDUAL NUMBER ON THE EVALUATION DATASET OF E-BIODIGITDB

	DTW	SW-DTW	RNNs	TA-RNNs
0	34.90	34,59	31,25	**29,51**
1	32.30	30,96	**27,33**	28,78
2	32.80	31,40	29,36	**25,87**
3	35.00	31,69	30,67	**26,89**
4	23.50	**18,75**	25,58	21,22
5	24.40	22,09	22,67	**19,62**
6	36.90	34,16	34,88	**26,74**
7	22.50	**20,20**	25,58	23,84
8	26.00	24,27	22,38	**19,19**
9	29.60	26,89	25,29	**21,95**
Average	29.82	27.50	27.50	**24.36**

neural networks. The RNN system outperforms the results achieved using both DTW and SW-DTW, achieving an average 23.72% EER, an absolute improvement of 4.85% and 3.34% EERs compared with DTW and SW-DTW, respectively. Finally, our proposed TA-RNN system is also evaluated in the same conditions, outperforming the results achieved by the other biometric systems for most upper-case characters. Our TA-RNN system obtains an average 20.99% EER, outperforming in large margin the results achieved by DTW approaches, and also the state-of-the-art RNN system with an absolute improvement of 2.73% EER. It is interesting to remark the considerable improvements achieved by our proposed TA-RNN system in some upper-case letters such as "A", "B", and "Z", with absolute improvements higher than 4.5% EER compared with the second best approach. Finally, the proposed TA-RNN system achieves the best system performance result with an EER as low as 11.90% for the letter "B", achieving absolute improvements of 7.15%, 7.15%, and 4.77% EERs compared with DTW, SW-DTW, and RNNs, respectively. These results remark the potential of our proposed TA-RNN system, outperforming all previous results achieved in the literature.

We now compare the results achieved when drawing upper- and lower-case letters (Table III). In general, better results are obtained when drawing lower-case letters. For the DTW and SW-DTW systems, lower-case letters achieve average absolute improvements of 1.94% and 2.19% EERs in comparison with upper-case letters, respectively. Similar improvements are also produced for both RNNs and TA-RNNs, with average absolute improvements of 3.52% and 2.29% EERs, respectively. These results remark the higher discriminative power of lower-case letters compared with upper-case letters. We believe this is produced because most upper-case letters are based mostly on simple straight strokes, and not so much in curved strokes, providing therefore less variability among users. In addition, we usually write using lower-case letters, adapting our original writing model to more user-specific features compared with upper-case letters. One example that justifies our hypothesis is letter "r/R". Letter "r" provides the best result with a 11.90% EER for the TA-RNN system. However, the EER increases up to 16.67% when using letter "R". Similar observations apply to other letters such as "v/V" and "y/Y". Nevertheless, there are some cases where both upper- and lower-case letters obtain very similar results, such as letters "s/S" and "g/G" as they are mainly curved strokes in both lower- and upper-case letters. Finally, the same trends along the authentication systems are observed for lower-case letters. Our TA-RNN system obtains an average 18.70% EER, achieving average absolute improvements of 7.93%, 6.1%, and 1.5% EERs compared with DTW, SW-DTW, and RNNs, respectively.

Characters based on symbols and numbers are also considered in this analysis. Tables IV and V include the results achieved using the evaluation datasets of MobileTouchDB and e-BioDigitDB, respectively. For the case of drawing numbers, our proposed TA-RNN system outperforms the other systems for both MobileTouchDB and e-BioDigit databases. For the MobileTouchDB database (Table IV), the proposed TA-RNNs achieves an average 20.24% EER whereas for the

their corresponding acquisition block. The evaluation results achieved over the e-BioDigitDB database are also depicted in Table V. Finally, we also provide the average EER obtained for each authentication system and acquisition block for a better comprehension of the results.

We first compare the system performance results when drawing letters (Table III). Analysing upper-case letters, the original DTW system achieves an average 28.57% EER, being the letter "X" the one that provides the best system performance with a final 19.05% EER. The SW-DTW system outperforms the original DTW system, achieving an average 27.06% EER, which demonstrates the better alignment carried out using this improved DTW version. Both DTW approaches are also compared with state-of-the-art recurrent
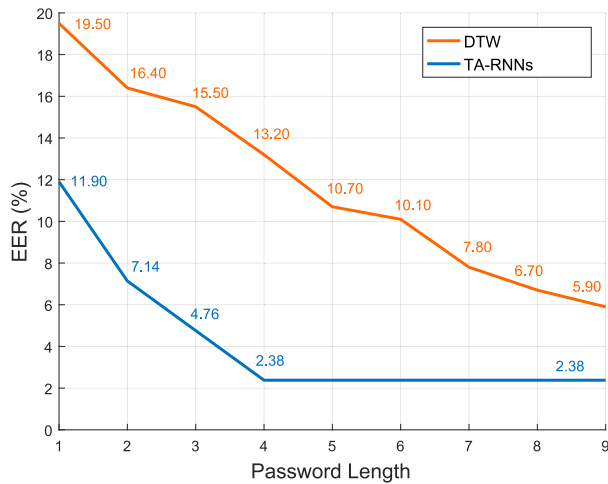
Fig. 6. System performance in terms of EER (%) when increasing the length of the password for the MobileTouchDB evaluation dataset.

e-BioDigitDB database this result increases up to 24.36% EER. We believe this is produced due to two main reasons: *i)* the time gap considered between train and test samples in the MobileTouchDB is about two days (i.e., s1 vs. s2) whereas in the e-BioDigitDB this time gap increases up to three weeks; and *ii)* the different acquisition scenario considered, office-like (e-BioDigitDB) and in the wild (MobileTouchDB). Finally, the symbol characters acquired in MobileTouchDB are also analysed in Table IV. In general, we can observe the same trends described before. However, it is interesting to remark the lower discriminative power of symbols compared with the other characters of MobileTouchDB when considering our proposed TA-RNN system.

B. Character Combination Analysis

This section evaluates the robustness of our proposed touchscreen password biometric approach when increasing the length of the password from 1 to 9 characters. The same touch biometric systems studied in Sec. VI-A are considered in this experiment. The final score of each password is produced by fusing the different one by one character score comparisons using the sum of the scores [13]. Fig. 6 shows the evolution of the system performance in terms of EER (%) when increasing the length of the password for the MobileTouchDB evaluation dataset. Passwords are created following the results extracted in the one-character analysis of Sec.VI-A, including the top ranked most discriminative characters in order, one at a time, e.g., the "B" and "r" characters are used for a 2-digit password for the TA-RNN system.

Analysing the results obtained in Fig. 6, our proposed TA-RNN system further outperforms the results achieved in [25] using the original DTW system. For example, for a 3-digit password, our proposed TA-RNN system achieves an absolute improvement of 10.74% EER compared with the DTW system. In general, we achieve an average absolute improvement of 7.52% EER when considering all length passwords (from 1 to 9 characters). Our proposed TA-RNN system achieves a final 2.38% EER using just a 4-digit password. However, for the DTW system, the best system performance result

is a 5.9% EER when considering a 9-digit password. The results achieved in this experiment prove the success of our proposed touchscreen password biometric approach and also the proposed TA-RNN system in terms of both usability (just a 4-digit password is required to achieve a 2.38% EER, not 9) and system performance (2.38% EER vs. 5.90% EER). Finally, it is important to note that in Fig. 6 no further improvement is observed by our proposed TA-RNN system when passwords comprise more than 4 characters. This might be produced due to passwords are created including the top ranked most discriminative characters, one at a time, avoiding repetitions. As a result, the discriminative power of the new characters included in the password is not enough to further improve the system performance.

C. Template Update Analysis

This section conducts some experiments in order to analyse the inter-session variability effect on the system performance. MobileTouchDB is considered in this experiment, with samples acquired in 6 different acquisition sessions with different time gaps between them, emulating practical scenarios. First, we focus on how the system performance can be improved along time when using samples from different acquisition sessions as training samples. Fig. 7 shows the average EER (%) of all 72 individual character comparisons of the evaluation dataset of MobileTouchDB. For each experiment, the first number indicates the training sessions (one sample per session), and the second one indicates the test session (i.e., the sixth session in all experiments). We first add training samples from sessions closer to the test. Then, we remove samples from older sessions in time.

First, we analyse the scenario where samples from sessions closer to the test are included for training. The results achieved in Fig. 7 show very similar trends for both DTW and the proposed TA-RNN systems. The system performance improves when we include samples from different acquisition sessions. For the DTW system, an average 28.80% EER is achieved when training samples are coming only from the first session (1 vs. 6). This system performance is further improved when increasing the number of training samples, e.g., an absolute improvement of 5.41% EER is achieved when using five training samples (1-5 vs. 6). Then, when we remove samples from sessions older to the test, the system performance gets worse. For the DTW system, an average 24.64% EER is achieved when using just one sample from the previous session to the test (5 vs. 6), an average 2.25% worsening compared with the case of using five training samples (1-5 vs. 6).

The same trend is observed when using our proposed TA-RNN system, but with a high system performance improvement compared with the DTW system. Our proposed system achieves an average absolute improvement of 8.52% EER. The best system performance achieved is a final 13.23% EER, which is 9.16% EER lower compared with the best result achieved by the DTW system. These results prove the robustness of our proposed TA-RNN approach with time, when samples form multiple acquisition sessions can be considered for training.
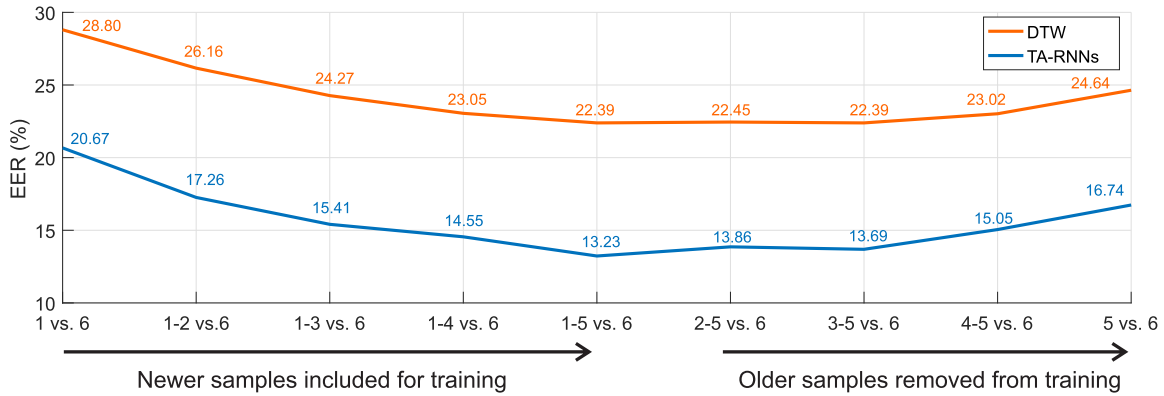
Fig. 7. Template update analysis: For each experiment, the first number indicates the training sessions, and the second one the test session (i.e., the sixth session in all experiments). We first add training samples from sessions closer to the test. Then, we remove samples from older sessions to the test.
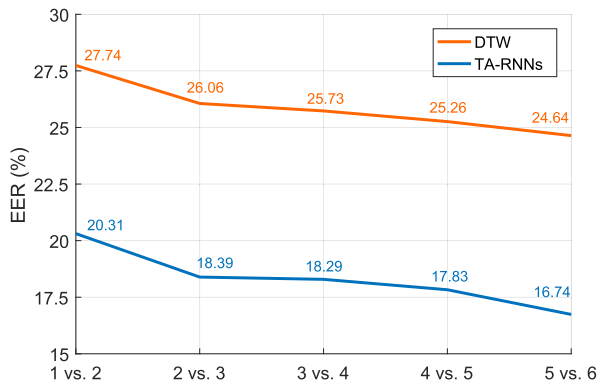

Fig. 8. Intra-user analysis on the evaluation dataset of MobileTouchDB.
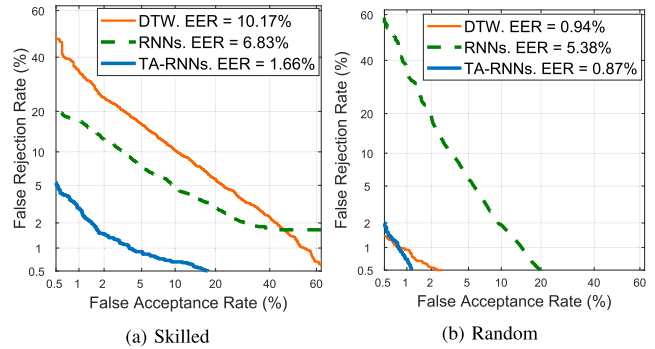

Fig. 9. System performance results of our proposed TA-RNN approach for the task of on-line signature verification.

TABLE VI
FURTHER APPLICATIONS OF OUR PROPOSED TA-RNN
SYSTEM: PERFORMANCE AS EER(%) FOR
SIGNATURE VERIFICATION

	DTW	RNNs	**TA-RNNs**
Skilled	10.17	6.83	**1.66**
Random	0.94	5.38	**0.87**

Finally, we also analyse the intra-user variability while drawing the characters along time. Fig. 8 shows the system performance in terms of EER(%) on the evaluation dataset of MobileTouchDB. Each experiment compares samples from one session with the same samples but for the next session in order to analyse the intra-user variability and the learning curve of the users. For both DTW and TA-RNN systems, the same trend is observed: the system performance improves with time. For the DTW system, an absolute improvement of 3.10% EER is achieved when we compare the 1 vs. 2 case with the 5 vs. 6 case whereas for the TA-RNN system, an absolute improvement of 3.57% EER is achieved for the same comparison. These results highlight how users tend to feel more comfortable with the system along time, reducing its variability while drawing the characters, and therefore, being possible to further improve the system performance with the number of acquisition sessions.

VII. TA-RNNs: FURTHER APPLICATIONS

Finally, our proposed TA-RNN system can be applied to many other different applications in which a direct comparison or verification of time sequence signals is performed, for example, in speech or gait [33], [34]. A related topic is signature verification. In this section we test our proposed TA-RNN approach for on-line signature verification considering the same experimental framework considered in [28], i.e., the first 300 users of BiosecurID are considered for development

whereas the remaining 100 users are considered for evaluation. It is important to remark that we consider just one training signature per user in this experiment (i.e., 1vs1 case) to show the real potential of TA-RNNs. Table VI shows the evaluation results in terms of EER (%) for both skilled and random forgery scenarios [27]. We also depict in Fig. 9 the results achieved in terms of Detection Error Tradeoff (DET) curves, for completeness. Our proposed TA-RNN system is compared with the state-of-the-art DTW and RNN configurations proposed for signature verification [35]. For skilled forgery scenarios, our proposed TA-RNN system achieves 1.66% EER, an absolute improvement of 8.51% and 5.17% EERs compared with the DTW and RNN systems respectively, which is a very important reduction of EER. Analysing random forgery scenarios, our proposed TA-RNN system achieves 0.87% EER, an absolute improvement of 4.51% EER compared with the RNN system proposed in [28]. It is important to remark that in [28], the RNN system was not able to outperform the DTW system for both skilled and random forgery scenarios. Nevertheless, our proposed TA-RNN system can learn better

feature representations of both type of impostors after the pre-warping of the time functions, outperforming the state of the art for both types of impostors.

VIII. CONCLUSIONS

This work enhances password scenarios through two-factor authentication approaches asking the users to draw each character of the password instead of typing them as usual. The main contributions of this study are as follows: *i)* We present the novel MobileTouchDB public database, acquired in an unsupervised mobile scenario with no restrictions in terms of position, posture, and devices. This database contains more than 64K on-line character samples performed by 217 users, with 94 different smartphone models, and up to 6 acquisition sessions. *ii)* We perform a complete analysis of the proposed approach considering both traditional authentication systems such as DTW and novel approaches based on RNNs. In addition, we present a novel approach named Time-Aligned Recurrent Neural Networks (TA-RNNs). This approach combines the potential of DTW and RNNs to train more robust systems against attacks.

In this study we have performed a complete analysis of the proposed approach using both MobileTouchDB and e-BioDigitDB. Our proposed TA-RNN system has outperformed the state of the art, achieving a final 2.38% EER, using just a 4-digit password and one training sample per character. These results encourage the deployment of our approach in comparison with traditional systems where the attack would have 100% success rate under the same impostor scenario.

In addition, we have also demonstrated the application of our proposed TA-RNNs for another time sequence recognition task, i.e., on-line handwritten signature verification. The proposed TA-RNN system has achieved 1.66% and 0.87% EERs for skilled and random forgery scenarios respectively, outperforming in large margin the state of the art.

Future work will be oriented to study the discriminative power of new features acquired in MobileTouchDB such as the area covered by the finger, accelerometer, and gyroscope in order to further improve the system performance [36]. We will also analyse the user experience in different acquisition settings through the analysis of the information acquired from Block5 to Block8 of the MobileTouchDB. Finally, we will evaluate the usability and performance improvement of our proposed TA-RNN approach for other behavioural biometric traits such as keystroke biometrics [37] and also on identification scenarios [38].

REFERENCES

[1] M. Salehan and A. Negahban, "Social networking on smartphones: When mobile phones become addictive," *Comput. Hum. Behav.*, vol. 29, no. 6, pp. 2632–2639, Nov. 2013.

[2] T. Hunt. (2019). *The 773 Million Record," Collection #1" Data Breach.* [Online]. Available: https://www.troyhunt.com/the-773-million-record-collection-1-data-reach/

[3] J. Bonneau, C. Herley, P. C. V. Oorschot, and F. Stajano, "The quest to replace passwords: A framework for comparative evaluation of Web authentication schemes," in *Proc. IEEE Symp. Secur. Privacy*, May 2012, pp. 553–567.

[4] J. Galbally, I. Coisel, and I. Sanchez, "A new multimodal approach for password strength estimation—Part I: Theory and algorithms," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 12, pp. 2829–2844, Dec. 2017.

[5] SplashData. *The Top 50 Worst Passwords of 2019.* Accessed: Jan. 2020. [Online]. Available: https://www.teamsid.com/1-50-worst-passwords-2019/

[6] A. Aviv, K. Gibson, E. Mossop, M. Blaze, and J. Smith, "Smudge attacks on smartphone touch screens," in *Proc. 4th USENIX Conf. Offensive Technol.*, 2010, pp. 1–7.

[7] D. Shukla, R. Kumar, A. Serwadda, and V. V. Phoha, "Beware, your hands reveal your secrets!" in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2014, pp. 904–917.

[8] Q. Yue, Z. Ling, X. Fu, B. Liu, W. Yu, and W. Zhao, "My Google glass sees your passwords!" in *Proc. Black Hat USA*, 2014, pp. 1–26.

[9] J. Angulo and E. Wastlund, "Exploring touch-screen biometrics for user identification on smart phones," *Privacy and Identity Management for Life*, J. Camenisch, B. Crispo, S. Fischer-Hübner, R. Leenes, and G. Russello, Eds. Berlin, Germany: Springer, 2011, pp. 130–143.

[10] P. Lacharme and C. Rosenberger, "Synchronous one time biometrics with pattern based authentication," in *Proc. 11th Int. Conf. Availability, Rel. Secur. (ARES)*, Aug. 2016, pp. 260–265.

[11] J. Jenkins, J. Shelton, and K. Roy, "One-time password for biometric systems: Disposable feature templates," in *Proc. SoutheastCon*, Mar. 2017, pp. 1–4.

[12] N. Yager and T. Dunstone, "The biometric menagerie," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 2, pp. 220–230, Feb. 2010.

[13] J. Fierrez, A. Morales, R. Vera-Rodriguez, and D. Camacho, "Multiple classifiers in biometrics. Part 2: Trends and challenges," *Inf. Fusion*, vol. 44, pp. 103–112, Nov. 2018.

[14] M. A. Ferrer, M. Diaz, C. Carmona-Duarte, and R. Plamondon, "IDeLog: Iterative dual spatial and kinematic extraction of sigma-lognormal parameters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 42, no. 1, pp. 114–125, Jan. 2020.

[15] R. Vera-Rodriguez *et al.*, "Modeling the Complexity of Signature and Touch-Screen Biometrics using the Lognormality Principle," in *The Lognormality Principle and its Applications*, R. Plamondon, A. Marcelli, and M. A. Ferrer, Eds. Singapore: World Scientific, 2019.

[16] R. Tolosana, R. Vera-Rodriguez, J. Ortega-Garcia, and J. Fierrez, "Preprocessing and feature selection for improved sensor interoperability in online biometric signature verification," *IEEE Access*, vol. 3, pp. 478–489, 2015.

[17] F. Alonso-Fernandez, J. Fierrez, and J. Ortega-Garcia, "Quality measures in biometric systems," *IEEE Security Privacy*, vol. 10, no. 9, pp. 52–62, Nov./Dec. 2012.

[18] M. Harbach, A. Luca, and S. Egelman, "The anatomy of smartphone unlocking: A field study of Android lock screens," in *Proc. Conf. Hum. Factors Comput. Syst.*, 2016, pp. 4806–4817.

[19] J. Fierrez, A. Pozo, M. Martinez-Diaz, J. Galbally, and A. Morales, "Benchmarking touchscreen biometrics for mobile authentication," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2720–2733, Nov. 2018.

[20] A. Acien, A. Morales, J. Fierrez, R. Vera-Rodriguez, and J. Hernandez-Ortega, "Active detection of age groups based on touch interaction," *IET Biometrics*, vol. 8, no. 1, pp. 101–108, Jan. 2019.

[21] T. Kutzner, F. Ye, I. Bonninger, C. Travieso, M. K. Dutta, and A. Singh, "User verification using safe handwritten passwords on smartphones," in *Proc. 8th Int. Conf. Contemp. Comput. (IC3)*, Aug. 2015, pp. 48–53.

[22] T. V. Nguyen, N. Sae-Bae, and N. Memon, "DRAW-A-PIN: Authentication using finger-drawn PIN on touch devices," *Comput. Secur.*, vol. 66, pp. 115–128, May 2017.

[23] R. Tolosana, R. Vera-Rodriguez, and J. Fierrez, "BioTouchPass: Handwritten passwords for touchscreen biometrics," *IEEE Trans. Mobile Comput.*, to be published.

[24] G. L. Lan and V. Frey, "Securing smartphone handwritten PIN codes with recurrent neural networks," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, 2019, pp. 2612–2616. May 2019.

[25] R. Tolosana, J. Gismero-Trujillo, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "MobileTouchDB: Mobile touch character database in the wild and biometric benchmark," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops*, 2019.

[26] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Benchmarking desktop and mobile handwriting across COTS devices: The e-BioSign biometric database," *PLoS ONE*, vol. 12, no. 5, May 2017, Art. no. e0176792.

[27] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Presentation attacks in signature biometrics: Types and introduction to attack detection," *Handbook Biometric Anti-Spoofing*, S. Marcel, M. S. Nixon, J. Fierrez, and N. Evans, Eds., 2nd ed. Cham, Switzerland: Springer, 2019.

[28] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Exploring recurrent neural networks for on-line handwritten signature biometrics," *IEEE Access*, vol. 6, pp. 5128–5138, 2018.

[29] M. Martinez-Diaz, J. Galbally, R. P. Krish, and J. Fierrez, "Mobile signature verification: Feature robustness and performance comparison," *IET Biometrics*, vol. 3, no. 4, pp. 267–277, Dec. 2014.

[30] M. Martinez-Diaz, J. Fierrez, and S. Hangai, "Signature matching," in *Encyclopedia Biometrics*, S. Z. Li and A. Jain, Eds. New York, NY, USA: Springer-Verlag, 2015, pp. 1382–1387.

[31] D. Folgado, M. Barandas, R. Matias, R. Martins, M. Carvalho, and H. Gamboa, "Time alignment measurement for time series," *Pattern Recognit.*, vol. 81, pp. 268–279, Sep. 2018.

[32] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, and J. Ortega-Garcia, "Reducing the template aging effect in on-line signature biometrics," *IET Biometrics*, vol. 8, no. 6, pp. 422–430, 2019.

[33] A. Graves, A.-R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proc. Int. Conf. Acoust., Speech Signal Process.*, May 2013, pp. 6645–6649.

[34] O. Costilla-Reyes, R. Vera-Rodriguez, P. Scully, and K. B. Ozanyan, "Analysis of spatio-temporal representations for robust footstep recognition with deep residual neural networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 41, no. 2, pp. 285–296, Feb. 2019.

[35] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales, and J. Ortega-Garcia, "Do you need more data? The DeepSignDB on-line handwritten signature biometric database," in *Proc. Int. Conf. Document Anal. Recognit. (ICDAR)*, Sep. 2019.

[36] D. Deb, A. Ross, A. K. Jain, K. Prakah-Asante, and K. V. Prasad, "Actions speak louder than (pass) words: Passive authentication of smartphone users via deep temporal features," in *Proc. Int. Conf. Biometrics*, 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8987433

[37] A. Morales *et al.*, "Keystroke biometrics ongoing competition," *IEEE Access*, vol. 4, pp. 7736–7746, 2016.

[38] X.-Y. Zhang, G.-S. Xie, C.-L. Liu, and Y. Bengio, "End-to-end online writer identification with recurrent neural network," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 2, pp. 285–292, Apr. 2017.

Ruben Vera-Rodriguez received the M.Sc. degree in telecommunications engineering from the Universidad de Sevilla, Spain, in 2006, and the Ph.D. degree in electrical and electronic engineering from Swansea University, U.K., in 2010. Since 2010, he has been with the Biometric Recognition Group, Universidad Autonoma de Madrid, Spain, where he has also been an Associate Professor since 2018. His research interests include signal and image processing, pattern recognition, and biometrics, with emphasis on signature, face, gait verification, and forensic applications of biometrics. He is actively involved in several National and European projects focused on biometrics. He has been the Program Chair for the IEEE 51st International Carnahan Conference on Security and Technology (ICCST) in 2017; and the 23rd Iberoamerican Congress on Pattern Recognition (CIARP 2018) in 2018.

Julian Fierrez (Member, IEEE) received the M.Sc. and Ph.D. degrees in telecommunications engineering from the Universidad Politecnica de Madrid, Spain, in 2001 and 2006, respectively. Since 2002, he has been with the Biometric Recognition Group, Universidad Politecnica de Madrid. Since 2004, he has also been with the Universidad Autonoma de Madrid, where he is currently an Associate Professor. From 2007 to 2009, he was a Visiting Researcher with Michigan State University, USA, under a Marie Curie Fellowship. His research interests include signal and image processing, pattern recognition, and biometrics, with an emphasis on multibiometrics, biometric evaluation, system security, forensics, and mobile applications of biometrics. He has been actively involved in multiple EU projects focused on biometrics (e.g., TABULA RASA and BEAT), and has attracted notable impact for his research. He was a recipient of a number of distinctions, including the EAB European Biometric Industry Award 2006, the EURASIP Best Ph.D. Award 2012, the Miguel Catalan Award to the Best Researcher under 40 in the Community of Madrid in the general area of science and technology, and the 2017 IAPR Young Biometrics Investigator Award. He is an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and the IEEE TRANSACTIONS ON IMAGE PROCESSING.

Ruben Tolosana received the M.Sc. degree in Telecommunication Engineering, and his Ph.D. degree in Computer and Telecommunication Engineering from Universidad Autonoma de Madrid, in 2014 and 2019, respectively. In April 2014, he joined the Biometrics and Data Pattern Analytics - BiDA Lab at the Universidad Autonoma de Madrid, where he is currently collaborating as a Post-Doctoral Researcher. Since then, Ruben has been granted with several awards, such as the FPU Research Fellowship from Spanish MECD (2015), and the European Biometrics Industry Award (2018). His research interests are mainly focused on signal and image processing, pattern recognition, and machine learning, particularly in the areas of face manipulation, human–computer interaction, and biometrics. He is author of several publications and also collaborates as a reviewer in many different high-impact conferences (e.g., ICDAR, IJCB, ICB, BTAS, EUSIPCO, etc.) and journals (e.g., IEEE TPAMI, TCYB, TIFS, TIP, ACM CSUR, etc.). Finally, he has participated in several National and European projects focused on the deployment of biometric security through the world.

Javier Ortega-Garcia (Fellow, IEEE) received the M.Sc. degree in electrical engineering and the Ph.D. degree *(cum laude)* in electrical engineering from the Universidad Politecnica de Madrid, Spain, in 1989 and 1996, respectively. He is currently a Full Professor with the Signal Processing Chair, Universidad Autonoma de Madrid, Spain, where he holds courses on biometric recognition and digital signal processing. He is the Founder and the Director of the BiDA-Lab, Biometrics and Data Pattern Analytics Group. He has authored over 300 international contributions, including book chapters, refereed journal, and conference papers. His research interests are focused on biometric pattern recognition (on-line signature verification, speaker recognition, human-device interaction) for security, e-health, and user profiling applications. He chaired Odyssey-04, The Speaker Recognition Workshop, ICB-2013, the 6th IAPR International Conference on Biometrics, ICCST2017, and the 51st IEEE International Carnahan Conference on Security Technology.