LSTM-based Free-Text Keystroke Dynamics for Continuous Authentication

Benoit Azanguezet Q.
University of Dschang, Cameroon
PO Box 67 Dschang, Cameroon

Junie Toukem T.
University of Dschang, Cameroon
PO Box 67 Dschang, Cameroon

Elie Tagne Fute
University of Buea, University of Dschang, Cameroon
PO Box 20269 Yaounde, Cameroon

ABSTRACT
We live in a fully networked world, with the ability to access digital information systems anytime, anywhere, using a variety of technological devices. This raises concerns about the security of systems and the protection of users' personal information. The login/password pair is the most widely used authentication method in today's information systems. However, the system becomes vulnerable if a third party obtains this information. Keystroke dynamics can be used as an additional layer of security to continuously check whether the person using the system is legitimate. In this research, we propose a continuous authentication model that uses the temporal and textual characteristics of a user's keystroke dynamics based on a one-way LSTM. We have tested this approach on data collected from 8 users. The final result obtained by this model is promising, with an accuracy of 96.67%.

General Terms
Biometric authentication, behavioural characteristics, user's biometric information, distribution of latencies.

Keywords
Continuous authentication, keystroke dynamics, deep learning, time series, Identity verification.

1. INTRODUCTION
As information and communication technologies (ICT) continue to develop, users' dependence on these technologies increases significantly. Vast amounts of user data are generated, processed and stored on servers and cloud storage systems around the world. These storage systems need to implement secure user authentication mechanisms. The use of a login/password pair is the most commonly implemented form of authentication in current systems. However, a password can be forgotten, stolen or even cracked using brute force attacks, exposing the system to the risk of identity theft, which can lead to the disclosure and misuse of system information. What's more, with the login/password authentication method, the user's identity is only verified during the connection phase of their work session. This means that if a user leaves their computer without locking their session, a malicious person can use this open session to attack the system. To add an extra layer of security, biometric authentication has been introduced.

Biometrics was originally the science of "measuring living things". It can be seen as an attractive solution for user authentication [4]. It was developed to automatically verify a person's identity based on measurable human physiological or behavioural characteristics [5]. Physiological authentication focuses on the identification of specific physical characteristics such as fingerprints [8], face recognition [21] and iris recognition [13]. Behavioural modalities are related to the behaviour of a person, including signature dynamics [18], mouse dynamics [7] and keystroke dynamics [12, 22]. This last modality is deterministic and is subject to regular research, development and continuous improvement.

Keystroke biometric authentication, also known as keystroke dynamic, is a practical and user friendly authentication option. It is a behavioural biometric method that allows a person to be recognized by the way they type on a keyboard. Several parameters are required to identify a person based on their typing style, such as the time each key is pressed, the time it is released, the time between two consecutive keys or the number of fingers used [1, 3, 6, 17]. This authentication system is inexpensive and unlikely to interfere with the user experience as the monitoring is transparent. It has some instability due to transient factors such as emotion, stress, illness and many others. Authentication using keystroke dynamics makes it possible to better counter the risks of identity theft, particularly for service providers or teleworkers, where the threat is potentially greater (replacement of the authorized person is not visible to the organization), and to improve both the security of the information system and the comfort of the user (no need to remember passwords). Some studies on keystroke dynamics authentication have used classical machine learning classifiers such as SVMs, random forests and decision trees. This is the case of [10], where the authors used a one-class SVM for continuous authentication. On the other hand, the authors of [9, 24, 25] use neural networks to identify people. While these studies may be an improvement on previous work, there are a number of limitations. Firstly, the features used are essentially the time taken to press and release a character and the latency between two successive characters. However, in a real environment, it is unlikely that a person will behave in a stable way for all the characters they type. Furthermore, these works does not take into account the sensitivities that may exist in the way users type over time. In this paper, we propose an authentication model based on Long Short-Term Memory (LSTM), which captures and maintains long-term dependencies in users' input habits. In addition, we have used temporal and textual features to create a biometric model of each user. The rest of this paper is structured as follows: section II presents the current state of research on keystroke dynamics; III presents the basic concepts using in this work; IV presents the proposed authentication model; V presents the experiments performed and the results obtained; and VI concludes the paper.

2. REVIEW OF LITERATURE

Current research into authentication using keystroke dynamics can be divided into three groups: the use of short text, the use of controlled long text, and the use of free text. Current research on authentication using keystroke dynamics can be divided into three groups: the use of short text, the use of controlled long text, and the use of free text. The use of short text means that the user's behaviour is studied on the basis of a short text provided to the user. This same text is used during the login phase and the identity verification phase [19, 23, 26]. This work has produced very good results. Unfortunately, this approach does not guarantee continuous protection of the information system. With controlled long text authentication, the user is presented with a long text in advance and is free to enter only the words found in this text [16, 20]. In this case, the information used to create the profile is not necessarily the information used for identity verification. The use of predefined free text has the advantage of being easy to implement, but is not feasible in a real context, as a person cannot always depend on the same text. In the third group, no work constraints are imposed on users. They are free to type as much 2as they like. Implementing this approach is relatively difficult because what the user types is not known in advance. let alone how many keystrokes they make before stopping. It is therefore important to extract features that provide meaningful representations for each individual. The authors of [11] propose to extract features by dividing the keyboard into three disjoint regions: left (L), right (R) and space (S) keys. All digraphs are then classified into one of the following eight groups L-L, L-R, L-S, R-L, R-R, R-S, S-L and S-R. They then calculated the mean pressure time (DD-time) for each group and used it as a single characteristic value. They used statistical approaches based on distance to classify 35 people and obtained an average EER of 19.47%. [14] also proposes to extract the temporal characteristics between two consecutive keys by dividing the keyboard. The authors also propose to use summary statistical style features, namely the number of words typed per minute (WPM), the percentage of times a user makes a typing error, and the percentage of use of the 'CapsLock' key to produce capital letters in a given typing task. They used a SVM with an ant colony optimization technique on a total of 30 individuals and achieved a performance of 0.0183 in FAR and 0.444 in FRR. Other authors, such as [9, 24, 25] use deep learning algorithms to improve the performance of authentication models based on the dynamics of free text typing

3. BIOMETRIC SYSTEMS OPERATION

Biometrics allows each person to be uniquely identified by physical and behavioural characteristics. These characteristics can be used for identification or authentication. Identification involves comparing a given biometric pattern with all previously stored patterns, while authentication involves a single comparison with the pattern of the person requesting identity verification. In general, biometric systems operate in three modes: enrolment, identity verification and identification.

3.1 Enrolment

Enrolment is the first phase of any biometric system. This is the stage where users are enrolled into the system for the first time. During this phase, and specifically in the case of keystroke dynamics, users are asked to type on a keyboard and information such as the keys on the keyboard, the time taken to type, the time taken to release each key, etc. is captured in order to extract a numerical representation. This representation is then reduced using a well-defined extraction algorithm to reduce the amount of data that needs to be stored for easy verification and identification.

3.2 Identity verification

Identity verification or authentication is the process of verifying that the person using the system is who they say they are. Again, the system collects the user's biometric information, compares it with the corresponding biometric template stored in the database and returns only a binary decision (yes or no). This process can be formalized as follows: suppose that C_Us the vector defining the biometric characteristics of the user U extracted by the system, and $M_U$is his biometric template stored in the database, the system returns a Boolean value after calculating the function f defined by :

$$f(C_{u,}M_u) = \begin{cases} 1, if\ S(C_{u,}M_u) \geq \gamma \\ 0, \quad else \end{cases} \qquad (1)$$

where S is the similarity function that defines the correspondence between the two biometric vectors and γ is the decision threshold above which the two vectors are considered identical.

3.3 Identification

Identification consists of determining the identity of an unknown person from an identity database. In this case, the system can either assign the unknown individual the identity corresponding to the closest profile found in the database, or reject the individual. The process can be formalized as follows: assuming that the input vector Cu defines the biometric characteristics extracted by the system when a user U appears in front of it, identification amounts to determining the identity of I_t, $t \in \{0, 1, \cdots, N\}$, where $I_1, \cdots, I_N$are the identities of users previously enrolled in the system and I_0indicates an unknown identity. The identification function f can thus be defined by:

$$f(C_u) = \begin{cases} I_k \ \ if\ \ max_{1\leq k\leq N} S(C_{u,}M_k) \geq \gamma \\ I_0 \qquad\qquad else \end{cases} \qquad (2)$$

where $M_k$is the biometric model corresponding to the identity I_k, S is the similarity function and γ is the decision threshold.

This article only addresses the first two phases.

4. PROPOSED METHOD

4.1 Data collection

To carry out this work, software was developed and installed on the PCs of 8 people to collect their data over a period of 30 days. This software works transparently and continuously and collects the events generated by the participants when they use their keyboards. No constraints were imposed on the participants. For each action i performed by a user U, the information collected is stored in a vector A in the form :

$$A_i = \{U_{id}, type_i, key_i, time_i\}$$

Where $U_{id}$is the user ID, $type_i$is the type of action performed on the keyboard key, $key_i$is the relevant keyboard key encoded in utf-8 and $time_i$is the system time in milliseconds during which the action is performed. There are two possible types of action for a key: *keyDown*, which corresponds to pressing the key, and *keyUp*, which corresponds to releasing the key. The information collected is recorded in a file on the participant's computer.

4.2 Data processing

4.2.1 Data cleaning

Before starting data processing, a number of cleaning measures were defined to ensure the accuracy of the data to be used. To this end, inconsistencies in the files of the different participants were identified and cleaning criteria were defined..

- If there are several identical occurrences in the file, only the first one is taken into account ;
- All of the lines in the file that do not contain all the information are removed ;
- During processing, input times longer than 200 milliseconds are ignored because, as mentioned in [10], this is a pause in user behaviour.

4.2.2 Feature extraction

The data contained in the initial file is difficult to interpret because it does not provide direct information about how a user interacts with the keyboard. It is therefore necessary to analyze this file in order to obtain characteristic information for building the biometric model of each user. To enable the model to learn more about a user's behavior from this sequential data, two types of features were extracted: temporal features, which focus on the temporal and sequential aspects of events, and textual features, which refer to the textual data associated with these events. These two types of features are used together in the LSTM model to capture both the sequential and semantic aspects of user typing behaviour.

- **Temporal features**:Extract times from single characters, two consecutive characters, and three consecutive characters.For individual characters, the hold latency (HL), which corresponds to the time difference between presses and releases, was examined. For character pairs, the press time (PT), which corresponds to the sum of the HLs of the two characters, and the rise/fall time (UD), which corresponds to the difference between the release time of the first character and the press time of the second, were considered. For Tri-gram, the TP as the sum of the HLs of the three characters and the UD as the sum of the UDs of the pairs of consecutive characters that compose the Tri-gram were considered. Let a, b and c be three consecutive characters. The corresponding formulae are :

$$HL_a = time_{keyUp(a)} - time_{keyDown\ (a)}$$

$$TP_{di} = HL_a + HL_b$$

$$UD_{di} = time_{keyDown\ (b)} - time_{keyUp\ (a)}$$

$$TP_{tri} = HL_a + HL_b + HL_c$$

$$UD_{tri} = UD_{ab} - UD_{bc}$$

These characteristics are normalized using the *MinMaxScaler* function. The values are scaled to a range between 0 and 1.

- **Textual features:**Single characters were considered, pairs of consecutive characters and consecutive tri-grams. Each of these features is transformed into a vector by the embedding method and used as input to the LSTM model. These embeddings capture the semantic relationships between characters and are learned during model training.

Each entry *E* in the user's biometric model has the following representation:

$$E=\begin{bmatrix} C_1, C_2, C_3, HL_a, HL_b, HL_c, TP_{di(ab)}, TP_{di(bc)}, UD_{di(ab)}, \\ UD_{di(bc)}, TP_{tri(abc)}, UD_{tri(abc)} \end{bmatrix}$$

4.3 LSTM-based Architecture for dynamical keystroke

Long short-term memory (LSTM) is a popular Recurrent Neural Network (RNN) architecture introduced by Hochreiter and Schmidhuber [2] to solve the gradient vanishing problem. It is used to model long-range relationships in input data. The main advantage of LSTMs is the inclusion of a cell state, which simply acts as a memory string by storing information from previous states. An LSTM has three gates: an input gate, an output gate and a forget gate, as shown in Figure 1. The three gates act as a logical check within the model and are used to predict the output of the model. The forget gate is calculated using the equation 3, where σ is the sigmoid function, $h_{t-1}$is the output of the previous step, $x_t$is the input of the current step, $W_f$and $b_f$are the weight matrix and bias of the forget gate. At the input gate, the following calculations are performed using the equations 4 and 5. The results are then used to update the state of the cell using equation 6, where C_{t-1}represents the state of the previous cell. The final output for the current step is calculated using equations 7 and 8.

$$f_t = \sigma\left(W_f[h_{t-1}, x_t] + b_f\right) \quad (3)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \quad (4)$$

$$\hat{C}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_{t.}C_t - 1 + i_t\hat{C}_t \quad (6)$$

$$O_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = O_t.tanh(C_t) \quad (8)$$

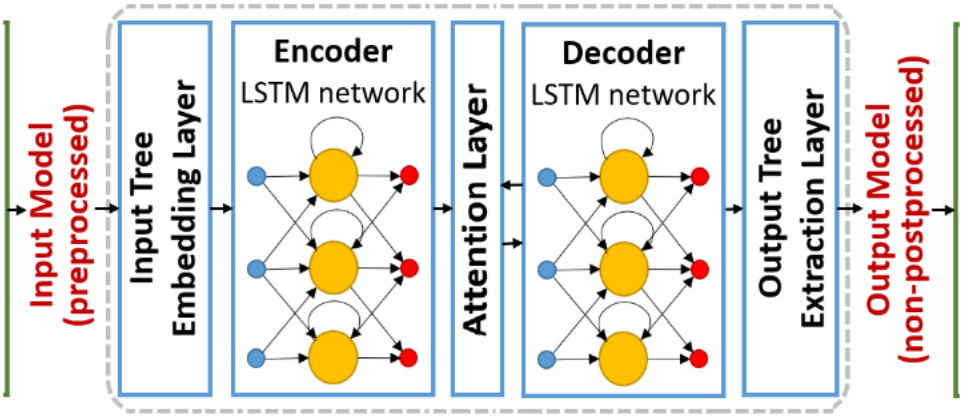

Figure 1: LSTM-based three layer architecture for dynamical keystroke

5. EXPERIMENTATION, RESULTS AND DISCUSSION

The various analyses and experiments carried out for the authentication model are presented in this section

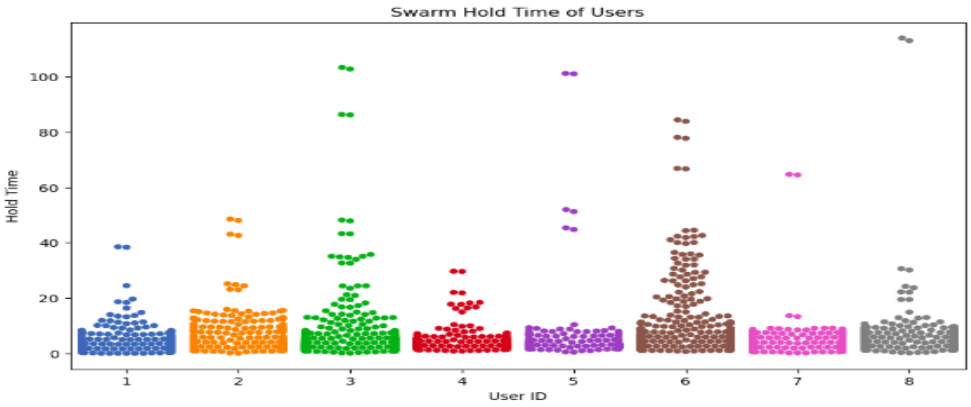

Figure 2: Time Variation for the input of two consecutive characters

5.1 Data exploration

The dataset used in the LSTM model contains the 12 features, including three textual features and nine temporal features, as described in section 4.2.2. Here the analysis is based on the typing data of 8 users. Figures 2 and 3 show the results of some analyses performed on the temporal features. Figure 2 shows the variation in the time taken to press two consecutive keys for the 8 users. This figure also shows that users 3 and 6 take much longer to press two consecutive keys than other users. On the other hand, user 7 has a relatively stable behaviour across all characters.

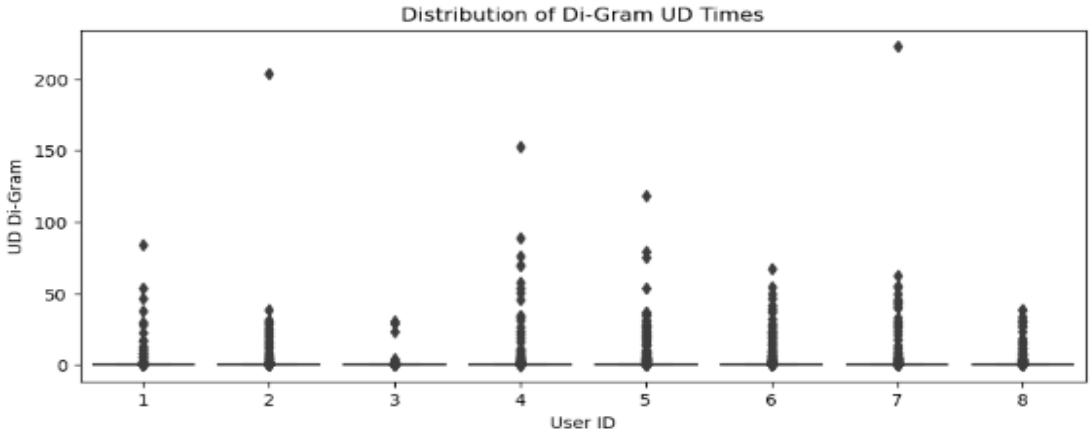

Figure 3: Two consecutive character UD time distribution

Figure 3 shows the distribution of latencies for two consecutive characters in relation to the median for all users. This figure also shows that user 3 takes less time between releasing a key and pressing the next key than the others.

5.2 Experimental classification

Continuous authentication using keystroke dynamics is a binary classification problem that involves continuously checking whether the person using the system is legitimate or an impostor. In this work, a specific model was created for each user. The legitimate data is that of the user and the impostor data is obtained by merging the data of the other 7 users. The data is divided in two, 80% for model training and 20% for testing. During the testing phase, the samples are classified using the sampling models obtained during the training phase. The proposed LSTM model achieves an accuracy of 98% on the training data and 96% on the validation data. Figure 4 shows the training curves of the model on the training and validation data.

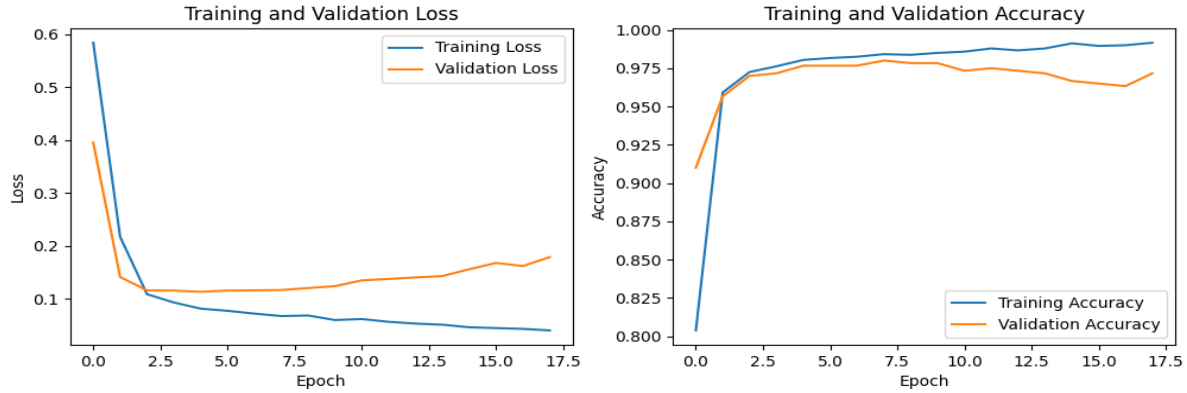

Figure 4: Learning and validation curves of the LSTM-based model

5.3 Discussion

The results obtained in Figure 4 show the importance of combining temporal and textual features for continuous authentication using keystroke dynamics. An accuracy of 96% on data never seen by the model proves that the extracted features are good indicators. Comparing the results of this work with those of previous works (references), a clear superiority emerges, which allows further exploration of this area of research. Table 1 details this comparison. In this table, the symbol "-" means that this feature is not used, "EER" represents the equal error rate and "ACC" the accuracy.

Table 1. Comparison of LSTM-based model with the existing works

	Model	Temporal features	Textual features	Results
Ahmed and Traore [9]	RNN	✓	-	EER=2.46 %.
Chang, Li, and Stamp [24]	CNN-GRU	✓	✓	ACC=92.3 %
Oak and Khare [15]	LDNA	✓	-	ACC=92%
Li, Chang, and Stamp [25]	CNN-RNN	✓	-	ACC=91.9 1%
LSTM-based approach	LSTM	✓	✓	ACC=96.6 7%

6. CONCLUSION

This paper proposes a continuous authentication system based on free-text keystroke dynamics. This system combines temporal and textual features. It uses a unidirectional Long Short-Term Memory (LSTM) model trained on data collected from eight volunteers. After extensive preprocessing and feature extraction, this model achieved an impressive accuracy of 96.67%, outperforming several existing methods in the field. This demonstrates that integrating temporal and textual data can significantly improve the reliability and accuracy of keystroke-based user authentication systems.

These promising results open several avenues for future work. - To ensure robustness across different demographic groups and typing styles, this model can be scaled by integrating a larger and more diverse dataset. - The integration of additional behavioral biometric data, such as mouse dynamics, eye tracking, or touchscreen gestures, could lead to a more comprehensive multimodal authentication system. - Deploying the model in real-time environments, such as online learning platforms or secure enterprise systems, would allow for further validation and refinement of the approach. - Finally, future research could explore adaptive learning mechanisms that update the model based on changes in user behavior, thus increasing long-term reliability and resistance to spoofing or impersonation attacks.

7. REFERENCES

[1] D. Umphress and G. Williams. “Identity verification through keyboard characteristics”. In: *International journal of man-machine studies* 23.3 (1985), pages 263–273.

[2] S. Hochreiter and J. Schmidhuber. “Long short-term memory”. In: *Neural computation* 9.8 (1997), pages 1735–1780.

[3] F. Monrose and A. Rubin. “Authentication via keystroke dynamics”. In: *Proceedings of the 4th ACM Conference on Computer and Communications Security*. 1997, pages 48–56.

[4] V. Matyas and Z. Riha. "Toward reliable user authentication through biometrics". In: *IEEE Security & Privacy* 1.3 (2003), pages 45–49.

[5] S. Prabhakar, S. Pankanti, and A. K. Jain. "Biometric recognition: Security and privacy concerns". In: *IEEE security & privacy* 1.2 (2003), pages 33–42.

[6] A. Peacock, X. Ke, and M. Wilkerson. "Typing patterns: A key to user identification". In: *IEEE Security & Privacy* 2.5 (2004), pages 40–47.

[7] K. Revett, H. Jahankhani, S. T. De Magalhaes, and H. M. Santos. "A survey of user authentication based on mouse dynamics". In: *Global E-Security: 4th International Conference, ICGeS 2008, London, UK, June 23-25, 2008. Proceedings*. Springer. 2008, pages 210– 219.

[8] N. Kaushal and P. Kaushal. "Human identification and fingerprints: a review". In: *J biomet biostat* 2.123 (2011), page 2.

[9] A. A. Ahmed and I. Traore. "Biometric Recognition Based on Free-Text Keystroke Dynamics". In: *IEEE Transactions on Cybernetics* 44.4 (2014), pages 458–472.

[10] P. Bours and S. Mondal. "Continuous authentication with keystroke dynamics". In: *Norwegian Information Security Laboratory NISlab* (2015), pages 41–58.

[11] P. Kang and S. Cho. "Keystroke dynamics-based user authentication using long and free text strings from various input devices". In: *Information Sciences* 308 (2015), pages 72– 93.

[12] Y. Zhong and Y. Deng. "A survey on keystroke dynamics biometrics: approaches, advances, and evaluations". In: *Recent Advances in User Authentication Using Keystroke Dynamics Biometrics* 1.1-22 (2015), page 2.

[13] M. De Marsico, A. Petrosino, and S. Ricciardi. "Iris recognition through machine learning techniques: A survey". In: *Pattern Recognition Letters* 82 (2016), pages 106–115.

[14] A. Alsultan, K. Warwick, and H. Wei. "Non-conventional keystroke dynamics for user authentication". In: *Pattern Recognition Letters* 89 (2017), pages 53–59.

[15] R. Oak and M. Khare. "A novel architecture for continuous authentication using behavioural biometrics". In: *2017 International Conference on Current Trends in Computer, Electrical, Electronics and Communication (CTCEEC)*. IEEE. 2017, pages 767–771.

[16] A. Alsultan, K. Warwick, and H. Wei. "Improving the performance of free-text keystroke dynamics authentication by fusion". In: *Applied Soft Computing* 70 (Sept. 2018), pages 1024– 1033.

[17] J. Kim, H. Kim, and P. Kang. "Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection". In: *Applied Soft Computing* 62 (2018), pages 1077–1087.

[18] J. Linden, R. Marquis, S. Bozza, and F. Taroni. "Dynamic signatures: A review of dynamic feature variation and forensic methodology". In: *Forensic science international* 291 (2018), pages 216–229.

[19] B. Bhana and S. Flowerday. "Passphrase and keystroke dynamics authentication: Usable security". In: *Computers & Security* 96 (2020), page 101925.

[20] A. T. Kiyani, A. Lasebae, and K. Ali. "Continuous user authentication based on deep neural networks". In: *2020 International Conference on UK-China Emerging Technologies (UCET)*. IEEE. 2020, pages 1–4.

[21] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri. "Face recognition systems: A survey". In: *Sensors* 20.2 (2020), page 342.

[22] N. Raul, R. Shankarmani, and P. Joshi. "A comprehensive review of keystroke dynamicsbased authentication mechanism". In: *International Conference on Innovative Computing and Communications: Proceedings of ICICC 2019, Volume 2*. Springer. 2020, pages 149– 162.

[23] S. Parkinson, S. Khan, A. Crampton, Q. Xu, W. Xie, N. Liu, and K. Dakin. "Password policy characteristics and keystroke biometric authentication". In: *IET Biometrics* 10 (2021), pages 163–178.

[24] H.-C. Chang, J. Li, and M. Stamp. "Machine Learning-Based Analysis of Free-Text Keystroke Dynamics". In: *Artificial Intelligence for Cybersecurity*. Springer, 2022, pages 331– 356.

[25] J. Li, H.-C. Chang, and M. Stamp. "Free-text keystroke dynamics for user authentication". In: *Artificial Intelligence for Cybersecurity*. Springer, 2022, pages 357–380.

[26] Y. B. W. Piugie, J. Di Manno, C. Rosenberger, and C. Charrier. "Keystroke dynamics based user authentication using deep learning neural networks". In: *2022 International Conference on Cyberworlds (CW)*. IEEE. 2022, pages 220–227.