DeepClap: A Gesture-Based User Authentication Scheme with Deep Neural Networks

Attaullah Buriro
Free University of Bolzano-Bozen
Bolzano, Italy
attaullah.buriro@unibz.it

Kevin Fred Mwaita
Free University of Bolzano-Bozen
Bolzano, Italy
kmwaita@unibz.it

ABSTRACT

This paper presents DeepClap, a gesture-based smartwatch user authentication scheme. DeepClap utilizes a faint clapping action performed by the user while wearing the smartwatch in one hand to authenticate its user. Technically speaking, it extracts users' identity on collected arm-movement signatures using built-in accelerometers and gyroscope sensors and authenticates them afterward. Preliminary evaluation of DeepClap on a collected dataset of 20 users results in reaching an overall average accuracy of ≈ 97%, and True Accept Rate (TAR) of 94.27%, and a False Accept Rate (FAR) of 0.296% using Deep Neural Network (DNN). The user-friendly nature of DeepClap, which eliminates the need for users to remember secret codes or gestures, and is resilient to advanced Fast Gradient Sign Method (FGSM) attack, make it an attractive option for widespread adoption.

CCS CONCEPTS

• **Security and privacy** → **Security services**; Network security; • **Human-centered computing** → *Ubiquitous and mobile devices*; *Human computer interaction (HCI)*;

KEYWORDS

Smartphone Authentication, Behavioral biometrics, Sensors

ACM Reference format:
Attaullah Buriro and Kevin Fred Mwaita. 2023. DeepClap: A Gesture-Based User Authentication Scheme with Deep Neural Networks. In *Proceedings of Twenty Eighth ACM Symposium on Access Control Models and Technologies (SACMAT), Trento, TN, Italy, June 7–9, 2023 (SACMAT'23),* 4 pages.
https://doi.org/10.1145/3176258.3176318

1 INTRODUCTION

Smartwatch-based behavioral biometrics - the use of collected unique human behavioral patterns, such as swiping [15], typing rhythm [14], the way a person walks (gait) [1], or arm-movements [10], could be used to build Electronic Identification and Trust Services (eIDAS)[1] solutions to provide an additional layer of security to the authentication process. Exploiting behavioral biometrics in eIDAS solutions would, enhance the security of the developed system: as unlike physical biometrics, i.e., fingerprints or facial attributes which could be stolen or replicated, behavioral patterns are difficult to be mimiced or copied. Additionally, as behavioral biometrics are dependent on the user actions and habits, they become more suited to frictionless and unobtrusive user authentication [2].

Behavioral-biometric-based smartwatch user authentication schemes could be used in eIDAS supported online banking mechanisms to ensure a secure and convenient way of authentication of cross-border users. To this end, these schemes could be exploited (i) to provide a second factor of authentication[2], (ii) for contactless payments[3], (iii) for secure messaging, and finally (iv) for providing location-based user authentication[4], before authorization of the transaction.

In this paper, we propose a frictionless and user-friendly behavioral-biometric-based smartwatch user authentication scheme, namely DeepClap, as an additional modality, for secure and reliable user authentication. In the enrollment stage, DeepClap, while the user claps (a faint clap without a sound), collects sensory data from accelerometer and gyroscope sensors, to profile users' unique arm movements. Later, in the authentication stage, DeepClap collects the same movements and matches them with the stored ones to execute identity verification. DeepClap, by applying a Deep Neural Network (DNN) decides if the smartwatch is worn by a legit user or by an impostor. Access to the banking services is granted if the user is confirmed as a legit user otherwise it is denied. DeepClap is completely frictionless and usable as it does not require any cumbersome effort from the user. We frame the problem of user authentication for banking applications as an n-class classification problem where the classifier is trained on samples of several users, hosted on the server. The decision is made on the server and access is granted or denied on the smartwatch.

DeepClap is an effective authentication scheme as it is not only user-friendly (users are not required to remember any secret or manage a token) but accurate. We report an overall accuracy of ≈ 97% (TAR of 94.27% and FAR of 0.296%).
The main contributions of the paper are listed below:

- The proposal of a DNN-powered arm-movement-based smartwatch user authentication scheme. DeepClap authenticates the user based on differences (the smaller the better) in the arm movements footprints generated during short clap (2, 3, or 4 sec).

[1] https://digital-strategy.ec.europa.eu/en/policies/eidas-regulation

SACMAT'23, June 7–9, 2023, Trento, TN, Italy

ACM ISBN 978-1-4503-5632-9/18/03…$15.00
https://doi.org/10.1145/3176258.3176318

[2] A customer while logging into their banking account, can receive a notification on their smartwatch to confirm their identity
[3] The customer's on the smartwatch needs to be authenticated using biometric data before the payment is authorized
[4] The customer's smartwatch GPS sensor could be used to verify that the customer is in a specific location before authorizing the transaction

- The collection of data from an accelerometer and gyroscope sensors in three activities, *sitting*, *standing*, and *walking* from 20 users.
- Implementation and evaluation of DeepClap on real smartwatch .

2 RELATED WORK

With the development of smartwatches, various features have been integrated, such as sensors that can detect movements like wrist rotations, arm gestures, and finger actions, as well as physiological readings like heart-rate, blood oxygen levels, skin temperature, and conductance. These readings could be used to develop implicit user authentication schemes, which have become an intriguing area for research in behavioral biometrics. Although the potential of using tapping [14], swiping [15], and motion-based actions [10] [12] [4] [6] [11] [17] [9] for user authentication has been explored, behavioral biometric-based user authentication on smartwatches have remained a less explored area.

The studies based on motion-assisted behavioral biometrics, such as Buriro et al. [4][6], Liang et al. [11], and Huang et al. [9], are highly relevant to our research. Liang et al. [11] propose hand-punch behavior as a behavioral biometric modality to authenticate users. Their approach, exploited accelerometer data to build a one-class SVM classifier achieving an accuracy of 95.45% on the collected dataset of 20 users. In [4] authors suggest using motion-assisted finger-snapping gestures to authenticate the smartwatch users. This finger-snapping approach exploited the accelerometer and gyroscope sensors for user profiling. Using the Multilayer Perceptron (MLP) as the 1-class classifier, the authors reported a TAR of 82.34% at a FAR of 34.25% on 15 training samples only. Later, Buriro et al. [6] explore smartwatch-worn in-air-finger-writing as a behavioral modality to authenticate the users. Using an MLP 1-class classifier the authors reported a TAR of 80.52% at 21.65% FAR on just 15 training samples. Huang et al. [9] combine gyroscope-powered in-the-air signing gestures with Dynamic Time Warping (DTW) as a classifier, and reported 90.1% accuracy on an 11-volunteer dataset.

Table 1: Comparison of our scheme with the related work. Our comparison is limited to the work which involved sensory readings, the mode, the classifiers the number of users, and the obtained results.

Paper	Input Method	Sensors	Classifiers	Users	Results
[10]	Free-form arm-movement	Acceleromter Gyroscope	DTW	5	Accuracy = 84.6%
[11]	Hand-punch movement	Accelerometer	SVM	20	Accuracy = 95.45%
[8]	Hand-writing-based movements	Acceleromter Gyroscope	MLP SVM	21	EER = 6.56%
[6]	In-air-finger-based movements	Acceleromter Gyroscope	DTW MLP	11	TAR = 80.52% FAR = 21.65%
[4]	Finger-snaping-based movements	Acceleromter Gyroscope	MLP	11	TAR = 82.34% FAR = 34.12%
[5]	Hand-clapping-based movements	Accelerometer Gyroscope	KNN MLP RF	50	TAR = 93.3% FAR = 0.22% Accuracy = 96.54%
[16]	Motion-based activities (18)	Accelerometer Gyroscope	KNN DT RF	51	Accuracy = 96.65% (clapping)
This work	**Hand-clapping-based movements**	**Accelerometer Gyroscope**	**DNN**	20	**TAR = 94.27% FAR = 0.269% Accuracy = ≈ 97%**

The current paper builds upon the research presented in [5] and presents an extended version of the previous work with a number of key enhancements. One significant improvement is exploitation of Deep Neural Network as a classifier, which sets this work apart from other studies (as shown in Table 1). This approach has enabled more accurate and efficient user recognition from clapping gestures, improving the overall performance of the system. Another notable addition to the present study is the implementation and evaluation of the approach on a real Android smartwatch, reflecting real-world use cases of the technology. The results of this implementation demonstrate the practical effectiveness of our approach, as well as its superior accuracy when compared with existing methods. In addition, DeepClap is robust to adversarial attacks: We also evaluated the robustness of the proposed user authentication framework against adversarial examples. Overall, our DeepClap brings a significant advancement in the field of gesture-based user authentication, with notable improvements in terms of TAR, FAR, and accuracy. Furthermore, the integration of DNN as the classifier, robustness to adversarial attacks, and real-world implementation make DeepClap unique and highly effective, paving the way for further research in this area.

3 APPROACH

Our proposed method employs the concept of using arm micro-movements during clapping as a behavioral biometric modality. We employ n-class classifiers at server side to verify the wearer's identity and authenticate them. The flowchart of our proposed approach is illustrated in Figure 1.

DeepClap obtains arm micro-movements during the clapping gesture. Technically speaking, DeepClap firstly collects raw 3-dimensional sensory data (from accelerometer and gyroscope) and extracts 82 statistical features and sends to the centralized server. This server stores the received feature vector in its database as a template to match query samples for user authentication later.

4 EXPERIMENTAL ANALYSIS

In this section, we explain our experimental analysis approach.

4.1 Data Collection

We developed a customized Android application, namely *ClapAuth*, to collect user arm movements during the course of clapping. Our application can be installed on any Android smartwatch having latest Android version.

We recruited 20 participants for this study. All the participants were students or researchers: bachelor, masters or PhD. We collected clapping data in three body postures: *sitting*, *standing*, and *walking*. We collected 100s of clapping data per user in each body posture resulting in 300s of users arm-movements in all three body postures. In this way, we collected 6000s of clapping data from all 20 users.

DeepClap requires its users to signup. Here sign-up means the provision of 300s of data for user profiling. After the signup process and the users can test the application by clicking on the sign-in button and providing the test clapping action.

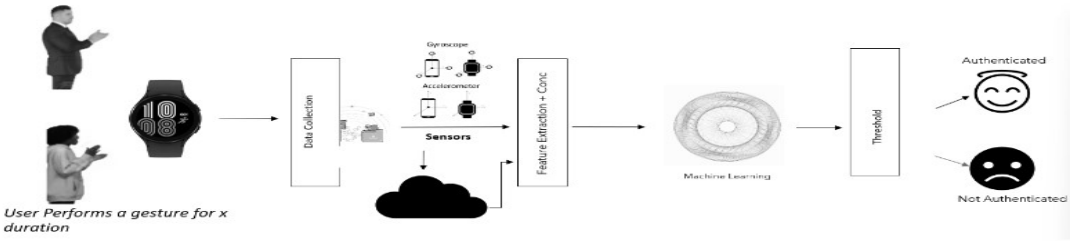


Figure 1: Block diagram of our approach

(a) Sign-in/up process **(b) data collector**

Figure 2: The data collection interface

4.2 Feature Extraction

The accelerometer and gyroscope are 3-dimensional sensors which means they generate readings in X, Y, and Z dimensions. Additionally, we computed the fourth dimension and named it as magnitude using the following mathematical formula:

$$m = \sqrt{sensor[x]^2 + sensor[y]^2 + sensor[z]^2} \quad (1)$$

In this way, we ended up having 4 dimensional data from both accelerometer and gyroscope sensors for all the clapping actions. We then extracted 11 statistical features from each dimension (41 from each of the sensors). The computed 11 features are: Min (4+4), Max (4+4), Mode (4+4), Median (4+4), Mean (4+4), Variance (4+4), Skewness (4+4), Kurtosis (4+4), Correlation (3+3), Abs (3+3), Cosine similarity (3+3). In this we computed 41 features from accelerometer and gyroscope sensors each. We then concatenate these 41 features from each stream and each of the sensors forming a final feature vector of 82 features.

4.3 Deep Neural Network

The proposed framework is an attempt of solving the problem of user authentication in client-server architecture such as banking, remote access, etc. In this scenario, the chosen ML classifier is trained on training samples of multiple users. We chose to develop a Deep Neural Networks (DNNs) based classification/verification scheme. To this aim, we searched the best number of layers (from 2 to 10), the best-required units in each layer (from 32 to 512 with a step size of 32), and the learning rate (0.01, 0.001, 0.0001). We leveraged Keras-tuner[5] for finding the best parameters of the DNN network using grid search. It is worth mentioning that we exploited training data (with a 33% validation size) for finding the best hyperparameters. We manage to find a 5-layers DNN architecture containing 288, 352, 512, 64 and 32, units, respectively. Additionally, the best learning rate was found to be 0.0001

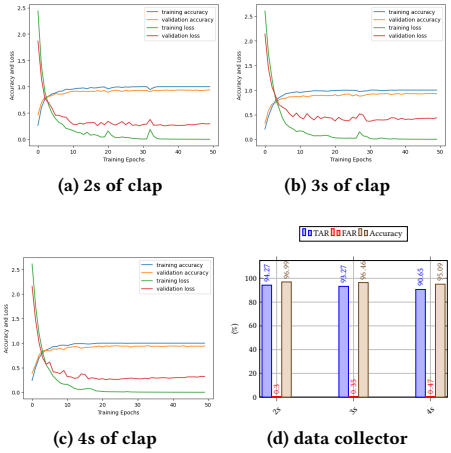
[5]https://keras.io/keras_tuner/

4.4 Classification Protocol

In this paper, we deal user authentication to a banking service as a multi-class classification problem where we train and test out classifier on the data provided by multiple users. We created 3 datasets based on the clap timings: 2s, 3s, and 4s. The assumption was to find out empirically how much duration of clapping action is adequate to authenticate the users. We use 66.67% of the data of each dataset for training and remaining 33.3% for testing.

4.5 Results

4.5.1 ***User Authentication.*** We summarise our classification results in Figure 3. Figure 3d illustrate the bar charts of our obtained results for different duration of claps. We achieved 94.27% TAR, 0.296% FAR and an accuracy of ≈ 97% on 2s of clapping data. Whereas for 3s and 4s duration, the accuracy drops a bit. Thus we can conclude that 2s of data has very low variance as compared to 3s and 4s durations, respectively. We also show learning curves for different durations in figures 3a (for 2s), 3b (for 3s) and 3c (for 4s). It is evident from these figures that classifier did not overfit.



(a) 2s of clap **(b) 3s of clap**

(c) 4s of clap **(d) data collector**

Figure 3: User authentication results

4.5.2 ***Adversarial Analysis***. Despite achieving remarkable performances, DNN-based systems have been shown to be susceptible to adversarial examples [18], which are modified versions of genuine samples intentionally distorted with adversarial noise to deceive DNNs and cause misclassifications. Evaluating pre-trained neural network-based systems under adversarial examples is vital for assessing robustness, revealing vulnerabilities, and minimizing dependence on the training data distribution. Therefore, our study also examined the robustness of the user authentication framework proposed in this research to adversarial examples.

Multiple techniques exist for generating adversarial examples, as described in Chakraborty et al. [7]. For our research, we utilized the (FGSM)[6], a white-box attack where the attacker has full understanding of the model's framework and components. The FGSM approach involves introducing a slight change to the input data through the use of a perturbation, calculated by obtaining the sign of the model's loss function gradient with respect to the input data. The main goal of this attack is to cause incorrect classification of the perturbed input by the model.

Various methods can be utilized to interpret FGSM attack outcomes, such as confidence scores, adversarial example visualizations, feature importance, and accuracy. Nevertheless, accuracy is the metric that is most frequently employed. If the accuracy of the classifier on the initial data split is meaningfully higher than that of the FGSM-created data, it indicates that the classifier is susceptible to adversarial attacks.

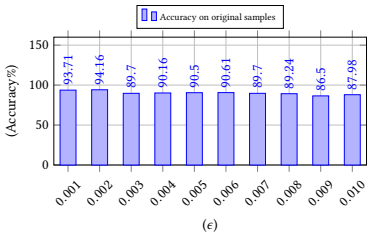


Figure 4: Results of FGSM using different epsilon values

In Figure 4, we present the results of generating FGSM attacks on our pre-trained 2s model. We report the results of FGSM attack on only one model because it worked well (yielded highest accuracy) and due to the space limitations. Our findings reveal that the pre-trained model is resilient to FGSM attack. Accuracy did not decrease much under different epsilon values. We varied the epsilon values for FGSM generation from 0.001 to 0.01, as this range is more suitable for tabular data. On the other hand, the values selected for image data range from 0.01 to 1, as reported in [13]. Our solution provides a secure platform for user authentication in the smartwatch user authentication sector by effectively countering adversarial attacks/examples.

5 CONCLUSIONS

This article introduces DeepClap, a secure, user-friendly, and resilient user authentication mechanism based on behavioral biometrics designed specifically for smartwatches. By using clap-based arm movements to register and authenticate the user, DeepClap takes advantage of the users' familiarity with the clapping action. Advanced Deep Neural Network was used to evaluate our proposed scheme, showing promising results with a high of ≈ 97% accuracy on 2s of clapping-based arm movements. To test the reliability of our model further, we subjected it to adversarial analysis using the Fast Gradient Sign Method (FGSM). Our results show that the model is robust against different levels of perturbation. In the future, we plan to develop a proof-of-concept application to apply our findings and evaluate our system's performance, security, and usability in different scenarios. Studies have shown that users' behavioral patterns vary under different conditions [3], and we plan to test our system accordingly.

[6]https://pytorch.org/tutorials/beginner/fgsm_tutorial.html

REFERENCES

[1] Neamah Al-Naffakh, Nathan Clarke, Fudong Li, and Paul Haskell-Dowland. 2017. Unobtrusive gait recognition using smartwatches. In *2017 International Conference of the Biometrics Special Interest Group (BIOSIG)*. IEEE, 1–5.

[2] Attaullah Buriro. 2017. *Behavioral biometrics for smartphone user authentication*. Ph.D. Dissertation. University of Trento.

[3] Attaullah Buriro, Bruno Crispo, Filippo Delfrari, and Konrad Wrona. 2016. Hold and sign: A novel behavioral biometrics for smartphone user authentication. In *2016 IEEE security and privacy workshops (SPW)*. IEEE, 276–285.

[4] Attaullah Buriro, Bruno Crispo, Mojtaba Eskandri, Sandeep Gupta, Athar Mahboob, and Rutger Van Acker. 2018. Snapauth: a gesture-based unobtrusive smartwatch user authentication scheme. In *International Workshop on Emerging Technologies for Authorization and Authentication*. Springer, 30–37.

[5] Attaullah Buriro and Francesco Ricci. 2023. ClapAuth: A Gesture-Based User-Friendly Authentication Scheme to Access a Secure Infrastructure. In *Emerging Technologies for Authorization and Authentication: 5th International Workshop, ETAA 2022, Copenhagen, Denmark, September 30, 2022, Revised Selected Papers*. Springer, 15–30.

[6] Attaullah Buriro, Rutger Van Acker, Bruno Crispo, and Athar Mahboob. 2018. Airsign: A gesture-based smartwatch user authentication. In *2018 International Carnahan Conference on Security Technology (ICCST)*. IEEE, 1–5.

[7] Anirban Chakraborty, Manaar Alam, Vishal Dey, Anupam Chattopadhyay, and Debdeep Mukhopadhyay. 2018. Adversarial attacks and defences: A survey. *arXiv preprint arXiv:1810.00069* (2018).

[8] Isaac Griswold-Steiner, Richard Matovu, and Abdul Serwadda. 2017. Handwriting watcher: A mechanism for smartwatch-driven handwriting authentication. In *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE, 216–224.

[9] Chenyu Huang, Zhice Yang, Huangxun Chen, and Qian Zhang. 2017. Signing in the air w/o constraints: robust gesture-based authentication for wrist wearables. In *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 1–6.

[10] Antwane Lewis, Yanyan Li, and Mengjun Xie. 2016. Real time motion-based authentication for smartwatch. In *2016 IEEE Conference on Communications and Network Security (CNS)*. IEEE, 380–381.

[11] Guan-Cheng Liang, Xiang-Yu Xu, and Jia-Di Yu. 2017. User-authentication on wearable devices based on punch gesture biometrics. In *ITM Web of Conferences*, Vol. 11. EDP Sciences, 01003.

[12] Chris Xiaoxuan Lu, Bowen Du, Xuan Kan, Hongkai Wen, Andrew Markham, and Niki Trigoni. 2017. VeriNet: User verification on smartwatches via behavior biometrics. In *Proceedings of the First ACM Workshop on Mobile Crowdsensing Systems and Applications*. 68–73.

[13] Arbena Musa, Kamer Vishi, and Blerim Rexha. 2021. Attack analysis of face recognition authentication systems using fast gradient sign method. *Applied artificial intelligence* 35, 15 (2021), 1346–1360.

[14] Toan Nguyen and Nasir Memon. 2018. Tap-based user authentication for smartwatches. *Computers & Security* 78 (2018), 174–186.

[15] Toan Van Nguyen, Napa Sae-Bae, and Nasir Memon. 2017. DRAW-A-PIN: Authentication using finger-drawn PIN on touch devices. *computers & security* 66 (2017), 115–128.

[16] Gary M Weiss, Kenichi Yoneda, and Thaier Hayajneh. 2019. Smartphone and smartwatch-based biometrics using activities of daily living. *IEEE Access* 7 (2019), 133190–133202.

[17] Xiaojing Yu, Zhijun Zhou, Mingxue Xu, Xuanke You, and Xiang-Yang Li. 2020. Thumbup: Identification and authentication by smartwatch using simple hand gestures. In *2020 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE Computer Society, 1–10.

[18] Jiliang Zhang and Chen Li. 2020. Adversarial Examples: Opportunities and Challenges. *IEEE Transactions on Neural Networks and Learning Systems* 31, 7 (2020), 2578–2593. https://doi.org/10.1109/TNNLS.2019.2933524