2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Deep Learning-based User Authentication with Surface EMG Images of Hand Gestures

Qingqing Li[1], Zhirui Luo[1], Jun Zheng[1,*]

Abstract— **User authentication is an important security mechanism to prevent unauthorized accesses to systems or devices. In this paper, we propose a new user authentication method based on surface electromyogram (sEMG) images of hand gestures and deep anomaly detection. Multi-channel sEMG signals acquired during the user performing a hand gesture are converted into sEMG images which are used as the input of a deep anomaly detection model to classify the user as client or imposter. The performance of different sEMG image generation methods in three authentication test scenarios are investigated by using a public hand gesture sEMG dataset. Our experimental results demonstrate the viability of the proposed method for user authentication.**

I. Introduction

User authentication is the process to verify the identity of a user to prevent authorized accesses to the sensitive information stored on systems or devices. The identity of the user can be represented by three types of factors for authentication: knowledge factors, ownership factors, and inherence factors [1]. A single-factor method employs one of the three types of factors to authenticate a user. For example, accessing user accounts by passwords, patterns, and secure questions are commonly used knowledge factor-based user authentication methods. In addition, two or more factors can be combined as a multi-factor user authentication method.

Traditional single-factor user authentication methods such as PINs, passwords, and patterns are susceptible to various attacks such as guessing attacks, shoulder surfing attacks, smudge attacks, and other advanced attacks. For example, thermal attacks were proposed in [2] that use thermal cameras to reveal the PINs or patterns used for authentication. To achieve a more secure authentication, more sophisticated and novel methods using inherence factors such as bioelectrical signals have been proposed. Electroencephalography (EEG) and electrocardiography (ECG) are two commonly used bioelectrical signals for authentication. For example, an EEG-based authentication system called PassEEG was proposed in [3], which utilizes deep learning for user classification. ECG-based schemes have also developed for user authentication of Internet of Things (IoTs) [4], mobile devices and stand-alone wearable devices [5].

In addition to EEG and ECG, surface electromyogram (sEMG) is another popular type of bioelectrical signal produced by the electrical activities of muscles. sEMG has been widely used for applications like hand gesture recognition but only seen a limited application for user authentication. Unlike gesture recognition which can be treated as a supervised classification problem, user authentication can't be solved with supervised learning as there are unlimited potential imposters. Instead, user authentication can be treated as an anomaly detection or novelty detection problem, where the detector is trained with only normal instances collected from the client at the registration or enrollment stage. All new and unseen instances from the imposters will be identified by the detector as abnormal if they are significantly different from the normal instances.

In recent years, deep learning has been successfully applied in a variety of applications including sEMG-based hand gesture recognition [6]. In this paper, we propose to solve the sEMG-based user authentication problem with deep learning enabled anomaly detection, i.e. deep anomaly detection. Deep anomaly detection has demonstrated significantly better performance than traditional anomaly detection methods in many challenging applications due to its unique capabilities such as end-to-end optimization and tailored representation learning [7]. Specifically, we utilize multi-channel sEMG signals generated by hand gestures as the biometric trait for user authentication. The multi-channel sEMG signals are converted into images which are used as the input of the deep anomaly detection model to classify the user as client or imposter. The main contributions of this paper are: (1) we propose a new user authentication method based on sEMG images of hand gestures and deep anomaly detection; (2) we investigate different ways to generate images from multi-channel sEMG signals for user authentication; (3) we evaluate the authentication performance of the proposed method by using a public hand gesture sEMG dataset.

II. Methodology

In this section, we present the details of the proposed user authentication modality that uses sEMG images of hand gestures and deep anomaly detection.

A. Overview of the Proposed User Authentication Modality

The proposed user authentication modality is illustrated in Fig. 1, which consists of three steps: acquiring of multi-channel sEMG signals, sEMG image generation, and deep anomaly detection for user classification. During the authentication process, the user performs a hand gesture as the authentication code. In the meantime, multi-channel sEMG signals are acquired from the electrode grid placed on the forearm of the user. The multi-channel sEMG signals are then converted into a sequence of sEMG images, which are

*Corresponding author: jun.zheng@nmt.edu
[1]Qingqing Li, Zhirui Luo and Jun Zheng are with the Department of Computer Science and Engineering, New Mexico Institute of Mining and Technology, Socorro, NM, 87801 USA

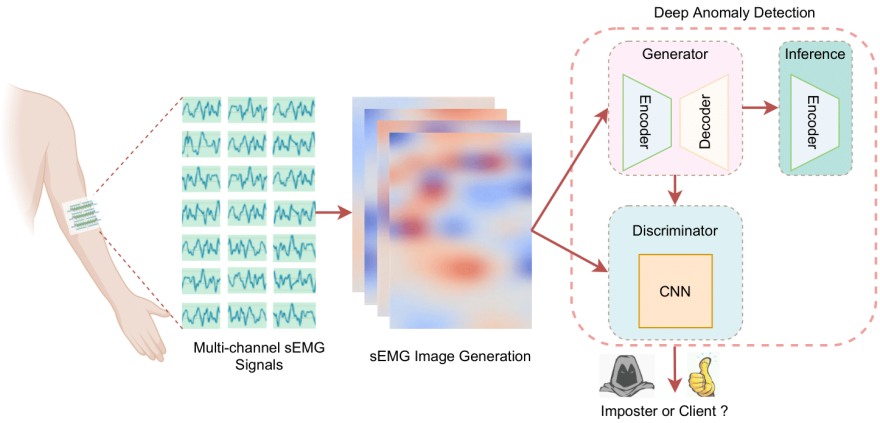

Fig. 1. Illustration of the proposed user authentication modality

fed into the deep anomaly detection model to classify the user as client or imposter. Note that the proposed modality can be used alone as a single-factor authentication method or combined with other factors to form a multi-factor authentication method. The modality can also use a sequence of gestures as the authentication code although we only consider a single gesture in this study.

B. sEMG Image Generation

Assume that the collected multi-channel sEMG data of a hand gesture has C channels where the sEMG signal in each channel has N samples (or frames). The channels are arranged as a two-dimensional array of size $W \times H$ based on the electrode grid where W and H are the numbers of rows and columns of the electrode grid, respectively. We apply three methods to generate images from the sEMG data: sEMG map [8], instantaneous sEMG image and difference sEMG image [6]. The generated sEMG images are then resized to fit the input size of the deep anomaly detection model if necessary.

1) sEMG Map: The sEMG map is generated by computing the time-domain feature value of a time windowed sEMG signal segment in a channel. The values obtained from all channels in the time window then form an image of size $W \times H$. A sequence of sEMG maps can be generated from the sEMG data by shifting the time window with a stride size of S. In this study, we consider three popular time-domain sEMG features: Mean Absolute Value (MAV), Root Mean Square (RMS), and Waveform Length (WL). The values of the three features in a time window are calculated as follows:

$$MAV = \frac{1}{M}\sum_{i=1}^{M}|X_i| \tag{1}$$

$$RMS = \sqrt{\frac{1}{M}\sum_{i=1}^{M}{X_i}^2} \tag{2}$$

$$WL = \sum_{i=1}^{M-1}|X_{i+1} - X_i| \tag{3}$$

where X_i is the ith sample of the sEMG signal segment and M is the size of the time window.

2) Instantaneous sEMG Image: Instantaneous sEMG image is directly generated from the samples of the channels in each time instance. The size of the image is determined by the electrode grid, i.e. $W \times H$. The multi-channel sEMG data of a hand gesture is converted into a sequence of N instantaneous sEMG images. Instantaneous sEMG image can be considered as an extreme case of sEMG map with a time window size of one.

3) Difference sEMG Image: Difference sEMG image is obtained as the difference between two consecutive instantaneous sEMG images which also has a size of $W \times H$. A sequence of $(N-1)$ difference sEMG images is generated from the multi-channel sEMG data of a hand gesture.

C. Deep Anomaly Detection

The deep anomaly detection model adopted in our study is GANomaly [9], an advanced model based on Generative Adversarial Networks (GANs). The architecture of GANomaly is shown in Fig. 2, which consists of three sub-networks, the generator $\boldsymbol{G}$, the inference $\boldsymbol{I}$, and the discriminator $\boldsymbol{D}$.

The generator $\boldsymbol{G}$ employs an encoder-decoder structure called adversarial auto-encoder (AAE). The encoder G_E takes an input image x and maps it into a lower dimension latent space representation $s = G_E(x)$, which can be considered as the best representative features of the input image x. The decoder G_D then reconstructs an image r_x from the latent space representation s, i.e. $r_x = G_D(s)$. The inference sub-network $\boldsymbol{I}$ is an encoder E that has the same architecture as G_E, which takes r_x as the input and maps it into a lower dimension latent space representation $r_s = E(r_x)$. The discriminator $\boldsymbol{D}$ is a CNN-based classifier which takes x and r_x from the generator $\boldsymbol{G}$ as inputs and classifies them as real or fake, respectively.

The objective function of GANomaly for model training combines three loss functions: adversarial loss, contextual loss, and encoder loss. The adversarial loss L_D is the L_2 distance between the feature representations of the original image x and the reconstructed image r_x generated by the

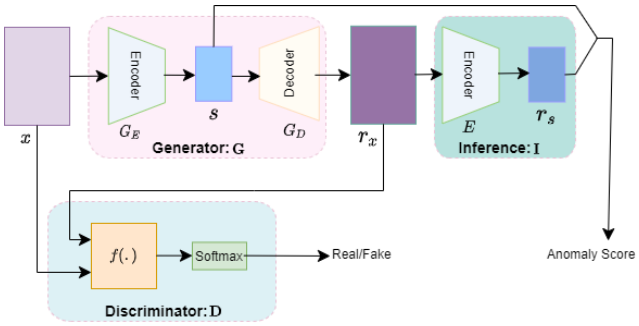

Fig. 2. The architecture of GANomaly

discriminator ***D*** which is defined as:

$$L_D = ||f(x) - f(r_x)||_2 \tag{4}$$

where $f(\cdot)$ is the output function of an intermediate layer of the discriminator ***D***. The adversarial loss is the feature matching loss for adversarial learning. The contextual loss L_G is defined as the L_1 distance between the original image x and the reconstructed image r_x:

$$L_G = ||x - r_x||_1 \tag{5}$$

The encoder loss L_E is defined as the L_2 distance between the latent space representation of the input image $s = G_E(x)$ and the latent space representation of the reconstructed image $r_s = E(r_x)$:

$$L_E = ||G_E(x) - E(r_x)||_2 \tag{6}$$

With the three loss functions, the objective function for model training is defined as:

$$L = w_D L_D + w_G L_G + w_E L_E \tag{7}$$

where w_D, w_G, and w_E are the weights of the three loss functions for adjusting their contributions to the objective function. In our experiments, the values of w_D, w_G, and w_E are set as 1, 50, and 1, respectively.

For a test image $\hat{x}$, the trained GANomaly model outputs an anomaly score $A(\hat{x})$ which is defined based on the encoder loss as shown in Equation (8). The value of the anomaly score indicates the abnormality of the test image.

$$A(\hat{x}) = ||G_E(\hat{x}) - E(G_D(G_E(\hat{x})))||_2 \tag{8}$$

D. Majority Voting

It has been shown in [6] that the accuracy of sEMG image-based hand gesture recognition can be significantly improved with a majority voting over the recognition results of multiple consecutive images. Similarly, the proposed modality can apply the majority voting strategy to improve the authentication performance where the predicted class label is the one representing the majority of the class labels predicted for the sEMG images in the voting window.

III. Performance Evaluation

A. Hand Gesture sEMG Dataset

We use the CapgMyo DB-a dataset from Zhejiang University [6] to evaluate the performance of the proposed authentication modality. The dataset contains high-density sEMG (HD-sEMG) signals collected from 18 subjects by an electrode grid of 16 rows and 8 columns placed on the forearm of subjects when they performed required hand gestures. A subject was required to perform 8 isometric and isotonic hand gestures and repeat each gesture for 10 times.

B. Experiments

To generate the sEMG maps for the three time-domain features, we set the slide window size and the stride size as $50ms$ and $10ms$, respectively, which means that the overlap between two consecutive windows is $40ms$. All sEMG images are resized from 16×8 to 16×16 by using the bi-cubic interpolation which are then fed into the GANomaly model.

We perform three test experiments for different authentication scenarios [10]: (1) normal test - the gesture registered for authentication is only known by the client; (2) leaked test - the imposter tries to gain access with the compromised authentication gesture; and (3) self test - the client forgets the authentication gesture and tries to use other gestures for authentication. For all three test experiments, half of the sEMG images generated from the authentication gesture of the client are used as the training dataset. Another half of the sEMG images of the authentication gesture of the client are combined with the imposter data to form the testing dataset. For the normal test experiment, the imposter data is the sEMG images generated from all gestures of other subjects. For the leaked test experiment, the imposter data is the sEMG images generated from the authentication gesture of other subjects. For the self-test experiment, the imposter data is the sEMG images generated from other gestures of the client.

The metrics used for performance evaluation are AUC (Area under the ROC Curve) and accuracy. ROC curve plots the performance of a classification model using true positive rate (TPR) and false positive rate (FPR) at different thresholds. The value of AUC indicates the classification capability of a model. To obtain the accuracy of a classification model, we use the Youden index which determines the classification threshold based on the ROC curve.

C. Results

Table I shows the results of the three test experiments by using different sEMG image generation methods without majority voting. The average AUC and accuracy of each sEMG image generation method for a test experiment are reported in Table I. It can be seen that MAV and RMS maps achieve significantly better performance than WL map, instantaneous and difference sEMG images in all three test scenarios. All methods have better performance in the normal and leaked tests than the self test showing that it's harder to differentiate the sEMG signals of different hand gestures

TABLE I

PERFORMANCE OF DIFFERENT SEMG IMAGE GENERATION METHODS IN TERMS OF AVERAGE AUC AND ACCURACY

Image Generation Method	Normal Test		Leaked Test		Self Test	
	AUC	Accuracy	AUC	Accuracy	AUC	Accuracy
MAV Map	0.988	0.961	0.975	0.949	0.918	0.845
RMS Map	0.986	0.960	0.978	0.952	0.921	0.875
WL Map	0.785	0.703	0.719	0.698	0.685	0.627
Instantaneous sEMG Image	0.850	0.787	0.820	0.754	0.723	0.673
Difference sEMG Image	0.813	0.761	0.786	0.732	0.712	0.655

performed by the client. On the other hand, a false positive in the self test scenario does little harm compared with other two test scenarios as it's still the client who gets authenticated.

The results of different sEMG image generation methods with majority voting in terms of average accuracy are shown in Fig. 3, where the x-axis is the size of the voting window represented as the number of sEMG frames in the voting window. It can be observed that majority voting can greatly improve the performance of instantaneous and difference sEMG images, especially in the early stage. The performance of MAV and RMS maps are only slightly improved with majority voting. When the voting window reaches around $80ms$, instantaneous and difference sEMG images achieve comparable results as MAV and RMS maps in the normal and leaked tests. MAV and RMS maps still outperform instantaneous and difference sEMG images in the self test even when majority voting is applied.

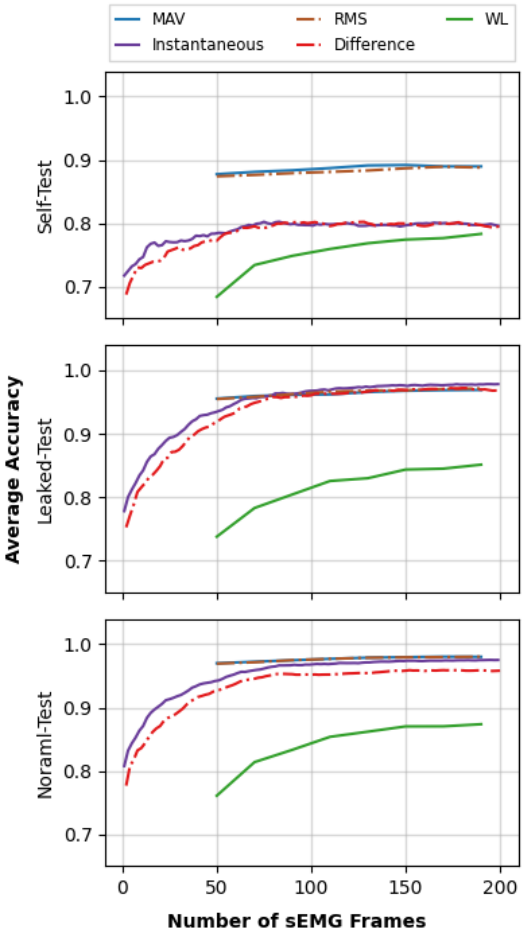

Fig. 3. Performance of different sEMG image generation methods with majority voting in terms of average accuracy

IV. CONCLUSIONS

In this paper, we propose a new user authentication method based on sEMG images of hand gestures and deep anomaly detection. Different sEMG image generation methods are investigated. The results show that the proposed method using the MAV and RMS sEMG maps can achieve a good authentication performance in three test scenarios. This indicates that the proposed method is a viable solution for user authentication.

ACKNOWLEDGMENT

This material is based upon work supported by the National Science Foundation EPSCoR Cooperative Agreement OIA-1757207.

REFERENCES

[1] C. Wang, Y. Wang, Y. Chen, H. Liu, and J. Liu, "User authentication on mobile devices: Approaches, threats and trends," *Computer Networks*, vol. 170, p. 107118, 2020.

[2] Y. Abdelrahman, M. Khamis, S. Schneegass, and F. Alt, "Stay cool! understanding thermal attacks on mobile-based user authentication," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, p. 3751–3763.

[3] G.-C. Yang, "Next-generation personal authentication scheme based on eeg signal and deep learning," *Journal of Information Processing Systems*, vol. 16, pp. 1034–1047, 2020.

[4] A. Barros, D. Rosário, P. Resque, and E. Cerqueira, "Heart of iot: Ecg as biometric sign for authentication and identification," in *2019 15th International Wireless Communications Mobile Computing Conference (IWCMC)*, 2019, pp. 307–312.

[5] J. S. Arteaga-Falconi, H. Al Osman, and A. El Saddik, "Ecg authentication for mobile devices," *IEEE Transactions on Instrumentation and Measurement*, vol. 65, no. 3, pp. 591–600, 2016.

[6] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface emg images," *Scientific Reports*, vol. 6, p. 36571, 2016.

[7] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, "Deep learning for anomaly detection: A review," *ACM Comput. Surv.*, vol. 54, no. 2, Mar. 2021.

[8] M. Rojas-Martínez, M. Mañanas, J. Alonso, and R. Merletti, "Identification of isometric contractions based on high density emg maps," *Journal of Electromyography and Kinesiology*, vol. 23, no. 1, pp. 33–42, 2013.

[9] S. Akcay, A. Atapour-Abarghouei, and T. P. Breckon, "Ganomaly: Semi-supervised anomaly detection via adversarial training," in *Computer Vision – ACCV 2018*, C. V. Jawahar, H. Li, G. Mori, and K. Schindler, Eds., 2019, pp. 622–637.

[10] J. He and N. Jiang, "Biometric from surface electromyogram (semg): Feasibility of user verification and identification based on gesture recognition," *Frontiers in Bioengineering and Biotechnology*, vol. 8, p. 58, 2020.