A New Deep Anomaly Detection-based Method for User Authentication Using Multi-Channel Surface EMG Signals of Hand Gestures

Qingqing Li, Zhirui Luo, and Jun Zheng, *Member, IEEE*

***Abstract*—User authentication plays an important role in securing systems and devices by preventing unauthorized accesses. Although surface Electromyogram (sEMG) has been widely applied for human machine interface (HMI) applications, it has only seen a very limited use for user authentication. In this paper, we investigate the use of multi-channel sEMG signals of hand gestures for user authentication. We propose a new deep anomaly detection-based user authentication method which employs sEMG images generated from multi-channel sEMG signals. The deep anomaly detection model classifies the user performing the hand gesture as client or imposter by using sEMG images as the input. Different sEMG image generation methods are studied in this paper. The performance of the proposed method is evaluated with a high-density hand gesture sEMG (HD-sEMG) dataset and a sparse-density hand gesture sEMG (SD-sEMG) dataset under three authentication test scenarios. Among the sEMG image generation methods, root mean square (RMS) map achieves significantly better performance than others. The proposed method with RMS map also greatly outperforms the reference method, especially when using SD-sEMG signals. The results demonstrate the validity of the proposed method with RMS map for user authentication.**

***Index Terms*—User authentication, deep anomaly detection, hand gesture, multi-channel surface electromyogram (sEMG) signal, sEMG image.**

I. Introduction

USER authentication is an important functionality of the identity management system (IDMS), which verifies the identity of a user to prevent unauthorized access to the sensitive information stored on systems or devices. There are three types of factors that can be used to represent the user's identity for the authentication purpose: knowledge factors, ownership factors, and inherence factors [1]. Passwords, PINs, and patterns are some typical examples of knowledge factors which are known by a user. Ownership factors are the elements possessed by a user, e.g. smart card and security token. Biometrics are the most popular type of inherence factors which are integral to a user. A user authentication method can employ one of the three types of authentication factors for a single-factor authentication or a combination of two or more factors for a multi-factor authentication as shown in Fig. 1. Knowledge factors have been widely used for user authentication but are susceptible to a number of attacks including brute-force attacks, guessing attacks, shoulder surfing attacks, and other more advanced attacks [2]. For example, smudge attacks were proposed in [3] to recover a mobile user's unlock pattern by observing the oily smudge left on the device screen. In [4], thermal cameras were used to launch attacks to reveal the PINs or patterns used for authentication. Naval et al. [5] developed an attack strategy to infer 4-digit PIN of a smartphone from mobile sensors' data. Ownership-based authentication methods require extra devices such as smart cards or security tokens to be carried by users which are inconvenient and the devices could be lost or stolen.

Inherent factors provide a number of advantages in usability and security over knowledge factors and ownership factors such as no need to memorize the authentication code and hard to stolen, which make them more and more popular for authentication in recent years. Biometrics based on physiological signals are a typical category of inherent factors that have attracted a lot of attention. The physiological signals commonly employed for user authentication are Electroencephalography (EEG) and Electrocardiography (ECG) [6]–[8]. For example, a two-factor user authentication scheme was proposed in [9] that combines EEG and signature data and employs a multimodal Siamese Neural Network (mSNN) for classification. In [10], Zhang et al. proposed DeepKey, a deep learning based multimodal biometric authentication system, which combines EEG and gait signals. A mobile user authentication scheme was developed in [11] that a user gains the access to the mobile device by touching two ECG electrodes of the device. Zhao et al. [12] designed a ECG-based authentication system, which generates ECG trajectory images using a generalized S-transformation to serve as the input to a convolutional neural network (CNN) model for classification.

On the other hand, surface Electromyogram (sEMG), another typical type of physiological signal widely applied for human machine interface (HMI) applications [13]–[20], has only seen a limited application for user authentication. In this paper, we propose a new deep learning based user authentication method using multi-channel sEMG signals of hand gestures. Deep learning has been successfully applied in a variety of areas including sEMG-based hand gesture recognition [21]. Unlike gesture recognition which can be treated as

This material is based upon work supported in part by the National Science Foundation EPSCoR Cooperative Agreement OIA-1757207, and in part by the Institute for Complex Additive Systems Analysis (ICASA) of New Mexico Institute of Mining and Technology. This paper is an extended version of our paper previously published in the proceedings of The 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC 2021), DOI: 10.1109/EMBC46164.2021.9630312. (*Corresponding author: Jun Zheng.*)

Q. Li, Z. Luo, and J. Zheng are with Department of Computer Science and Engineering, New Mexico Institute of Mining and Technology, Socorro, NM, 87801 USA (e-mail: jun.zheng@nmt.edu).

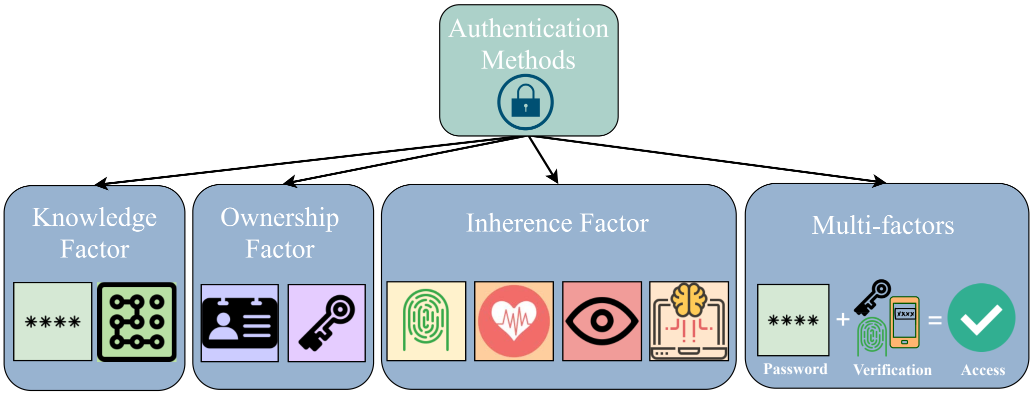

Fig. 1. User authentication factors and methods

a supervised classification problem, user authentication cannot be solved using 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. In this study, we solve the sEMG-based user authentication problem with 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 [22]. Specifically, we utilize multi-channel sEMG signals generated by hand gestures as the biometric trait for user authentication. Both sparse-density sEMG (SD-sEMG) and high-density sEMG (HD-sEMG) are investigated in our study. HD-sEMG signals are acquired with an electrode array of high channel count that provide richer information of electrical muscle activities than SD-sEMG signals acquired with a low-density electrode array. The multi-channel sEMG signals are converted as images to be used as the input of the deep anomaly detection model.

The main contributions of this paper are: (1) we propose a new user authentication method based on deep anomaly detection and sEMG images of hand gestures. To the best of our knowledge, this is the first work that employs a deep anomaly detection model with sEMG images for sEMG-based user authentication; (2) we investigate different techniques to generate images from multi-channel sEMG signals including instantaneous sEMG images, difference sEMG images, and sEMG maps for user authentication; (3) we conduct a performance evaluation with two publicly available hand gesture sEMG datasets to demonstrate that the proposed method is a viable solution for sEMG-based user authentication.

The rest of this paper is organized as follows. Section II introduces the related work on sEMG-based hand gesture recognition, traditional hand gesture based user authentication, and sEMG-based identity management. In Section III, we present the details of the proposed user authentication method based on sEMG images of hand gestures and deep anomaly detection. The experiments and results of the performance evaluation using two publicly available hand gesture sEMG datasets are shown in Section IV. Finally, conclusions of this paper are drawn in Section V.

II. Related Work

In this section, we review the existing work in three areas that are related to the proposed method: sEMG based hand gesture recognition, traditional hand gesture based user authentication, and sEMG based identity management.

A. *sEMG based Hand Gesture Recognition*

Hand gesture recognition has been widely used in HMI applications such as sign language [13], medical rehabilitation [17], virtual reality [18] and robot control [15]. Traditional hand gesture recognition can be categorized as static gesture recognition and dynamic gesture recognition [23]. The former recognizes gestures by using hand shape while the latter does the recognition through hand motion trajectory in space. Usually the gesture shape and location information is gathered using vision-based devices such as camera, depth sensitive sensor, or special wearable glove [24]. With the development of more applicable and user-friendly wearable devices for non-invasive acquisition of sEMG signals, sEMG based hand gesture recognition has become more and more popular. Three types of features extracted from sEMG signals can be used for hand gesture recognition: time domain features, frequency domain features, and time-frequency domain features [25].

A gesture recognition system using sEMG signals was proposed in [26] to classify nine gestures, which employs time domain features, root mean square ratio (RMSR) and autoregressive (AR) model, and a linear discriminant analysis (LDA) algorithm as the classifier. Neacsu et al. [27] proposed an sEMG-based hand gesture recognition system by using fully-connected neural networks and time domain features to classify the sEMG signals. Both systems use SD-sEMG signals for gesture recognition, 3 channels in [26] and 7 channels in [27].

In addition to traditional machine learning algorithms, deep learning has also been applied on the problem of sEMG-based hand gesture recognition. For example, CNNs have been widely adopted to extract the spatial relationship between multi-channel sEMG signals. Allard et al. [28] developed a CNN based scheme for hand gesture recognition in robotic arm guidance using sEMG based frequency domain features. In [21], Geng et al. proposed a CNN-based gesture recognition system using instantaneous sEMG images generated from the multi-channel sEMG signals as the input of the CNN model. Their experiment results show that the recognition accuracy can reach 89.3 % with a single sEMG image, which can be further improved to 99.0 % by using a simple majority voting over 40 consecutive sEMG images. A novel attention-based hybrid CNN and Recursive Neural Network (RNN) architecture was proposed in [29] for sEMG-based hand gesture recognition, where the multi-channel sEMG signals are generated as six different types of images to be used as the input of the hybrid CNN-RNN architecture. The hybrid CNN-RNN architecture aims to capture both spatial and temporal information of sEMG signals. Inspired by the work of [21], [29], we generate different types of images from multi-channel sEMG signals to serve as the input of the deep anomaly detection model for the purpose of user authentication.

B. Traditional Hand Gesture based User Authentication

Traditionally, hand gestures have been used as a popular way for the purpose of user authentication. The methods can be classified as two main categories: trajectory-based and vision-based [30].

1) Trajectory-based Methods: Methods in this category track key-points of hand obtained by specific devices or sensors during hand motion as trajectories, which are a kind of behavior traits. Features extracted from these trajectories are then used for the purpose of authentication. Khoh et al. [31] proposed an in-air hand gesture signature scheme which employs Microsoft Kinect sensor camera as the acquisition device. A sequence of images are captured when a subject performs the sign action in air. Features based on the temporal and motion information are extracted from the image sequence for user identification or verification. In [32], Lewis et al. proposed a finger movement-based authentication method which utilizes an optical motion capturing system employing optical markers placed on hand and optical tracking cameras. A group of goniometric (joint-related) and dermatologic (skin-related) features are extracted from the captured data to build the authentication model. A 3D gesture authentication scheme called GesID was proposed in [33] which utilizes an infrared depth camera, Leap Motion, as the acquisition device. The scheme builds the authentication model based on the sparse autoencoder (SAE) one-class classifier.

2) Vision-based Methods: Unlike trajectory-based methods, vision-based methods use features such as hand shape, hand geometry, skin texture, hand finger pattern, etc., directly extracted from the acquired images or videos for authentication. In [34], a hand biometric authentication method was proposed which employs images of stationary hand gestures captured by a low-cost video camera. Features are extracted from the original intensity image and the hand contour for classification. Wong and Kang [35] proposed a stationary hand gesture authentication scheme which utilizes edit distance on finger pointing direction interval (ED-FPDI) to model hand gesture behavior. In [30], a deep learning model called Dynamic-Hand-Gesture Authentication network (DHGA-net) was built for automatically extracting discriminant features from hand gesture video for authentication. The model consists of two components: a 3D convolution network to extract the spatial and temporal features from hand gesture videos and a Temporal-Identity-Extracting (TIE) module to handle the extracted spatio-temporal gesture features.

C. sEMG based Identity Management

Although sEMG has been widely applied for HMI applications, only limited works have been seen recently on using sEMG for identity management. Biometrics can be applied to two identity management functionalities: user authentication (or called user verification) and user identification (or called person identification) [36]. In [37], [38], sEMG signals generated from a list of hand gestures were used for building a mobile user authentication system. He and Jiang [39] studied the feasibility of using sEMG of wrist and hand gestures as a biometric trait for user verification and identification, where improved Discrete Fourier Transform (iDFT) features extracted from windowed raw sEMG signals were used for Mahalanobis distance based similarity matching. They further investigated the impact of feature sets and number of channels on the performance of sEMG based biometrics for user verification and identification [40]. Three feature sets, time-domain features, frequency division technique (FDT), and autoregressive (AR) features, and their combinations were studied. In [41], Li et al. proposed a two-factor user authentication scheme that utilizes sEMG-based biometrics to enhance the security of Android pattern unlock. 11 time domain features extracted from a single channel sEMG signal were used as the input of a one-class classifier to identify the user as client or imposter. A cancelable person identification scheme was proposed by Jiang et al. [42], where HD-sEMG signals under the isometric contractions of different finger muscles were used as biometric tokens. A feature vector combining three features, waveform length (WL), frequency median (FMD) and spatial synchronization (SS), from each channel, was fed into a KNN classifier for person identification. The work of [37], [38], [40], [42] didn't consider the spatial relationship between multi-channel sEMG signals. Jiang et al. [43] also proposed a neuromuscular password-based user authentication scheme using HD-sEMG based neuromuscular biometrics. The isometric contractions of different finger muscles are used as passwords for authentication. The scheme utilizes both time–frequency–space domain features at macroscopic level and motor neuron firing rate features at microscopic level to differentiate imposters and client. Unlike the aforementioned approaches, our work uses the deep anomaly detection model to automatically capture the spatial relationship between the multi-channel sEMG signals from the generated sEMG image for improved performance of user authentication.

III. Methodology

In this section, we first present an overview of the proposed user authentication method using sEMG images of hand gestures followed by the description of the methods used for

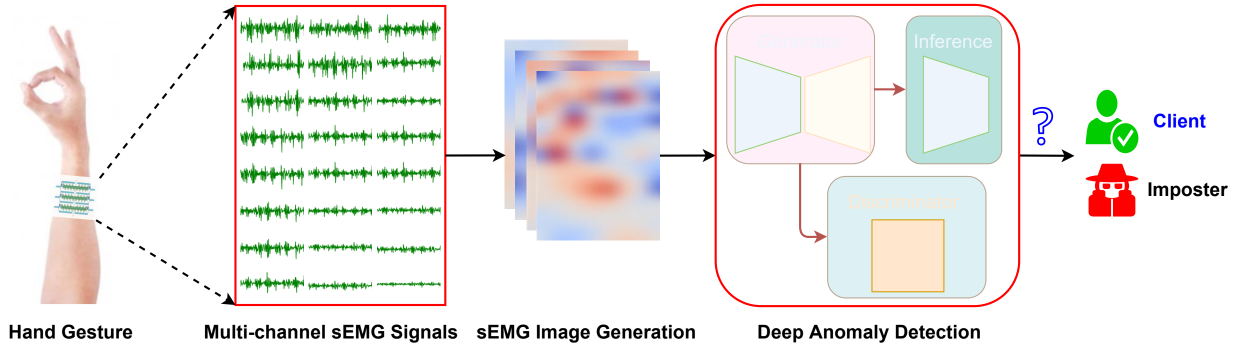

Fig. 2. An illustration of the proposed deep anomaly detection-based user authentication method using multi-channel sEMG signals of hand gestures

generating sEMG images. The deep anomaly detection model adopted for classifying the client and impostors is presented next. Finally, the majority voting strategy for improving the authentication performance is described.

A. Overview of the Proposed User Authentication Method

The proposed user authentication method consists of three steps as illustrated in Fig. 2. An electrode grid of size $W \times H$ is placed on the forearm of the user to be authenticated, where W and H are the numbers of rows and columns of the electrode grid respectively. In the first step of the authentication process, the user performs a hand gesture as the authentication code while the multi-channel sEMG signals are acquired from the electrode grid in the meantime. The second step is to convert the multi-channel sEMG signals into a sequence of sEMG images using a method described in Section III-B. The sEMG images will be fed into a deep anomaly detection model to classify the user as client or imposter in the final step. It should be noted that the proposed authentication method can be used individually as a single-factor method or combined with other authentication methods as a multi-factor method. Although we only consider a single gesture as the authentication code in this study, the proposed method can be easily extended to use a sequence of gestures as the authentication code.

B. sEMG Image Generation

According to Section III-A, the electrode grid placed on the forearm of the user has C electrodes which are arranged as a 2-D array of size $W \times H$. The multi-channel sEMG data of a hand gesture are acquired from the C channels where the sEMG signal of each channel has N samples (or frames). Three methods are applied for generating images from the multi-channel sEMG data: instantaneous sEMG image [21], difference sEMG image [21], and sEMG map [44]. Different time domain and frequency domain features are considered for generating sEMG maps. Depending on the input size of the adopted deep anomaly detection model, the generated sEMG images may need to be resized for using as the model input.

1) Instantaneous sEMG Image: An instantaneous sEMG image is directly generated by treating the sample values of the C channels in a time instance as pixels, which results in an image of the same size as the electrode grid, i.e. $W \times H$. By using this method, the multi-channel sEMG data of a hand gesture is converted into a sequence of N instantaneous sEMG images.

2) Difference sEMG Image: A difference sEMG image is obtained as the difference between two consecutive instantaneous sEMG images, whose size is also $W \times H$. A sequence of $(N-1)$ difference sEMG images is generated from the multi-channel sEMG data of a hand gesture by using this method.

3) sEMG Map: To generate an sEMG map, the value of a time domain or frequency domain feature is extracted from a time windowed sEMG signal segment in a channel first. In this study, we consider two time domain sEMG features: root mean square (RMS) and WL, and two frequency domain features: $Hjorth2$ and spectral entropy ($SpectralEn$), which are popular sEMG features widely used in previous studies [45]–[49]. The feature values obtained from all channels are used as pixels to form an image of size $W \times H$. A sequence of sEMG maps can be generated from the multi-channel sEMG data by shifting the time window with a stride size of S.

RMS and WL are two of the most popular time domain features for EMG signal processing where RMS indicates the envelope of the signal and WL implies the measure of complexity in each segment of EMG signal [48]. The values of the two features for a time windowed sEMG signal segment $\mathbf{X}$ are calculated as follows:

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

$$WL = \sum_{i=1}^{l-1} |X_{i+1} - X_i| \tag{2}$$

where X_i is the ith sample of $\mathbf{X}$ and l is the size of the time window.

$Hjorth2$ is calculated as the mean frequency or average dominant frequency of a time series signal as shown in Equation (3). $Hjorth2$ is also called Mobility derived from the statistical moments of the power spectrum [50].

$$Hjorth2 = \sqrt{\frac{var(\mathbf{X}')}{var(\mathbf{X})}} \tag{3}$$

where $\mathbf{X}'$ is the first-order derivative of sEMG segment $\mathbf{X}$, and $var(\mathbf{X})$ known as $Activity$ or $Hjorth1$ represents the signal power.

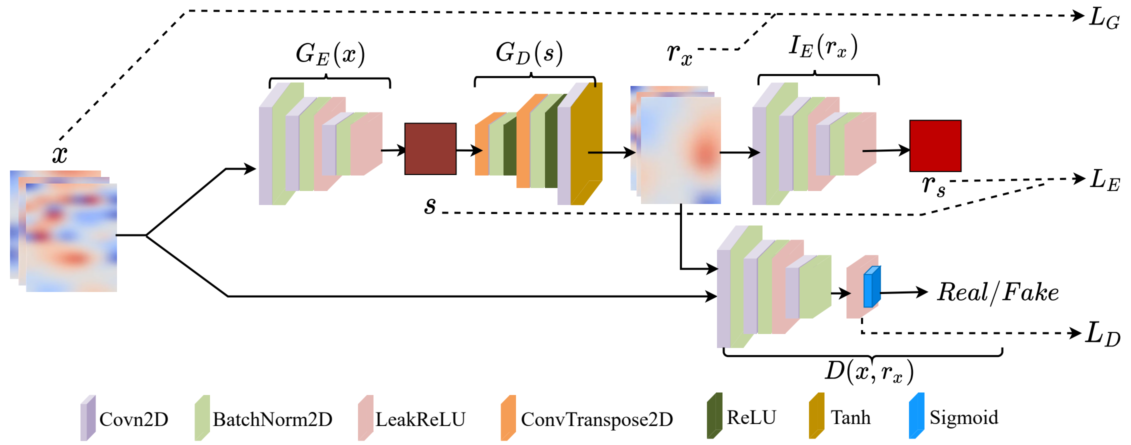

Fig. 3. The architecture of GANomaly

$SpectralEn$ measures the complexity of time series in the frequency domain. It's defined as the Shannon entropy of the power spectral density (PSD) of the sEMG signal segment $\mathbf{X}$ [51]:

$$SpectralEn = -\sum_{j=1}^{N} P_j \log_2 P_j \tag{4}$$

where P_j is the energy of jth frequency component of $\mathbf{X}$'s spectrum, and N is the number of components in the spectrum.

C. Deep Anomaly Detection Model

The deep anomaly detection model adopted in our study is GANomaly [52], which is an advanced model based on Generative Adversarial Networks (GANs). GANomaly has been successfully applied for applications such as geochemical anomaly detection [53], anomaly detection of aerospace data [54], machine fault detection [55], [56], and heart disease diagnosis [57]. Since the goal of this study is to investigate the validity of the proposed authentication method, we directly adopt the hyperparameters of the GANomaly model in [52] without fine-tuning.

As illustrated in Fig. 3, a GANomaly model is composed of three sub-networks: Generator $\boldsymbol{G}$, Inference $\boldsymbol{I}$, and Discriminator $\boldsymbol{D}$. Generator $\boldsymbol{G}$ employs an encoder-decoder structure called adversarial auto-encoder (AAE) contains encoder G_E and decoder G_D. The encoder G_E takes an input image x and forwards it into *2d-convolution* layers followed by the *LeakReLU* activation and *2d-BatchNorm* layers. The convolutional layers downscale image x and map 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 uses the convolutional transpose layers (*ConvTranspose2D*), with *ReLU* activation and *2d-BatchNorm* layers. The convolutional transpose layers upsample the latent space s to be the reconstruction of input image x, named as r_x, i.e. $r_x = G_D(s)$. Inference $\boldsymbol{I}$ is an encoder E with the same architecture as G_E but different parametrization. E takes r_x as the input and maps it into a lower dimension latent space representation $r_s = E(r_x)$, which has the same dimension as the first latent space s. For the decoder G_D in Generator, the *tanh* activation is used for the last convolutional layer to output the reconstruction r_x but not the *ReLU* activation, which makes a GAN architecture more stable [58]. Discriminator $\boldsymbol{D}$ is the third sub-network with a CNN-based classifier which takes x and r_x from $\boldsymbol{G}$ as inputs and classifies them as real or fake, respectively. The convolutional layers in the discriminator are fully connected. The *LeakyReLU* activation and *2d-BatchNorm* are used for all layers except for the last convolutional layer, which is flattened and then fed into a *Sigmoid* function layer for classification.

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 $\boldsymbol{D}$ which is defined as:

$$L_D = ||f(x) - f(r_x)||_2 \tag{5}$$

where $f(\cdot)$ is the output function of an intermediate layer of $\boldsymbol{D}$. 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:

$$L_G = ||x - r_x||_1 \tag{6}$$

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 $r_s = E(r_x)$:

$$L_E = ||G_E(x) - E(r_x)||_2 \tag{7}$$

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

$$L = w_D L_D + w_G L_G + w_E L_E \tag{8}$$

where w_D, w_G, and w_E are the weights of the three loss functions for adjusting their contributions to the objective function.

After receiving a test image $\hat{x}$ as input, the trained GANomaly model outputs an anomaly score $A(\hat{x})$ which is defined based on the encoder loss as shown in Equation (9).

The value of the anomaly score indicates the abnormality of the test image.

$$A(\hat{x}) = ||G_E(\hat{x}) - E(G_D(G_E(\hat{x})))||_2 \tag{9}$$

D. Majority Voting for Improving Authentication Performance

The simple majority voting strategy was used in [21] to improve the accuracy of hand gesture recognition with instantaneous and difference sEMG images. Similarly, the majority voting strategy can be applied in the proposed method to improve the authentication performance. When a user performs a hand gesture for authentication, a sequence of sEMG images are generated from the acquired multi-channel sEMG signals. The trained deep anomaly detection model classifies each image as one of the two classes: client or imposter. A voting time window is selected starting from the beginning of the signal acquisition. The final authentication result will be the class that has more than half of the votes from the images in the voting window. In practice, considering the authentication performance and time, one can select the voting time window size as the point when there is no significant performance improvement by further increasing the window size.

IV. Performance Evaluation

In this section, we first introduce the datasets used for performance evaluation of the proposed method followed by the description of experimental settings. Then we present the evaluation results which demonstrate the feasibility of the proposed method for user authentication.

A. Hand Gesture sEMG Datasets

Two publicly available hand gesture sEMG datasets are used in our study to evaluate the performance of the proposed user authentication method. The first one is CapgMyo DB-a [21], an HD-sEMG dataset, which contains sEMG data acquired from 18 subjects by using a 128-channel electrode grid. The 128 electrodes were organized as a 16×8 grid placed on the forearms of subjects when they performed required hand gestures. Each subject was asked to perform 8 isometric and isotonic hand gestures shown in Fig. 4 and repeat each gesture for 10 times.

The second dataset is Ninapro DB-1, a SD-sEMG dataset acquired from 27 subjects by using a 10-channel electrode grid [59], [60]. Ninapro DB-1 consists of sEMG data of four different exercises where exercise B includes the same 8 isometric and isotonic hand gestures as CapgMyo DB-a and 9 basic wrist movements. To be consistent in the evaluation, we only use the sEMG data of the 8 isometric and isotonic hand gestures of exercise B in our study. Similar to CapgMyo DB-a, each subject was asked to repeat a gesture for 10 times during the acquisition of Ninapro DB-1.

B. Experiments

We consider three authentication scenarios based on the relationship between the gesture used for authentication and the registered gesture to evaluate the proposed authentication method [39]: (1) *Normal Test* - the client uses the registered gesture for authentication while an imposter uses any possible gesture for authentication, i.e. the registered gesture is not compromised by the imposters; (2) *Leaked Test* - both the client and imposters use the registered gesture for authentication, i.e. the registered gesture is compromised by the imposters; and (3) *Self Test* - the client forgets the registered gesture and tries to use any possible gesture for authentication.

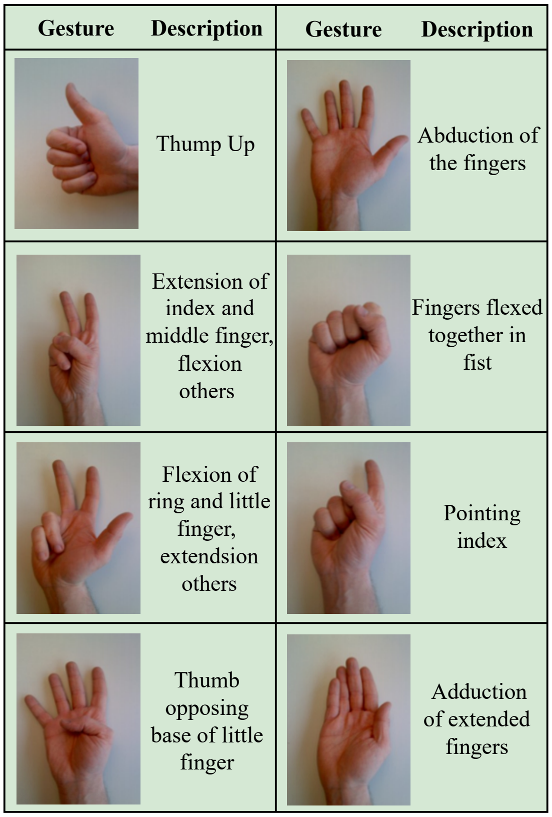
Gesture
Description
Gesture
Description
Thump Up
Abduction of the fingers
Extension of index and middle finger, flexion others
Fingers flexed together in fist
Flexion of ring and little finger, extendsion others
Pointing index
Thumb opposing base of little finger
Adduction of extended fingers

Fig. 4. 8 isometric and isotonic hand gestures of CapgMyo DB-a dataset

For each authentication scenario, the evaluation is repeated by treating each gesture in a dataset as the registered authentication gesture. Thus, the process repeats 144 and 216 times for CapgMyo DB-a and Ninapro DB-1, respectively. The subject who performed the registered gesture is treated as client while other subjects are treated as imposters. For all three authentication scenarios, the training dataset is formed by sEMG images generated from the sEMG data of randomly chosen 5 repetitions of the registered gesture. sEMG images generated from the sEMG data of other 5 repetitions of the registered gesture are combined with the imposter data to form the testing dataset. For the scenario of normal test, the imposter data are the sEMG images generated from all gestures of imposters. The imposter data of the leaked test are the sEMG images generated from the registered gestures of imposters. For the scenario of self test, the imposter data are the sEMG images from other gestures of the client.

When generating the sEMG maps from the sEMG data of a gesture, we set the sliding window size and stride size as 50 ms and 10 ms, respectively. Thus, there is an overlap of 40 ms between two consecutive windows. The sliding window and stride sizes were determined through our preliminary experiments empirically. All generated sEMG images are resized to 16 × 16 by using the bi-cubic interpolation, which is the input size of the GANomaly model.

We evaluate the performance of the proposed user authentication method by considering different sEMG image generation methods described in Section III-B. We also include a method proposed recently in [39] for comparison, which utilizes the iDFT features extracted from multi-channel sEMG signals for user verification and identification. The performance of different methods are analyzed by using the detection error tradeoff (DET) curve, which plots the relationship between false rejection rate (FRR) and false acceptance rate (FAR) under various detection thresholds. FRR represents how likely an authentication attempt is rejected by the method when the client performs the registered gesture. FAR represents how likely a gesture from an imposter is accepted by the method. We use two performance metrics obtained from the DET curve for evaluation: the area under the DET curve (AUC) and the equal error rate (EER). EER is obtained when FRR equals to FAR. For both metrics, a lower value indicates a better authentication performance. We use the accuracy as the performance metric to compare the performance of different methods with majority voting under different voting window sizes. The paired t-test with a level of significance $\alpha = 0.05$ is employed to evaluate the statistical significance of performance comparison.

C. Results

1) Results on HD-sEMG Dataset: We first present the results obtained from the HD-sEMG dataset, CapgMyo DB-a. Figs. 5 and 6 use the boxplots to summarize the performance of different methods in terms of AUC and EER, respectively. The results clearly show that RMS map achieves the best performance in all three test scenarios among the sEMG image generation methods. The proposed method with RMS map also has a significantly better performance than the iDFT-based method for all three test scenarios (all p-values $< .001$). The EERs (Mean ± Standard Deviation) obtained by the proposed method with RMS map for the normal, leaked, and self tests are 2.31±2.98 %, 3.07±3.70 %, and 13.39±9.76 %, respectively. Another observation from the results is that all methods have better performance in the normal and leaked tests than the self test. This shows that it is harder to differentiate multi-channel sEMG signals generated from hand gestures performed by the same user than signals generated from hand gestures performed by different users. However, a false positive in the self test scenario does much less harm compared with false positives in the other two test scenarios as it's still the client who gets authenticated.

In Fig. 7, we compare the performance of four different methods with majority voting using the accuracy as the performance metric. In the figure, the mean and standard deviation values of the results obtained by a method under different voting window sizes are plotted. We only report the results of RMS map in Fig. 7 because it has the best performance among the four sEMG map generation methods. The x-axis of Fig. 7 is the number of sEMG frames which is the size of the voting window. The results show that majority voting has significant impact on improving the performance of instantaneous image, difference image and the iDFT-based method, especially in the early stage, while it only slightly improves the performance of RMS map. When the voting window reaches around 70 ms, instantaneous and difference images achieve comparable performance as RMS map in the normal and leaked tests. RMS map still has a significantly better performance than other methods in the self test when majority voting is applied (p-values $< .001$ for all voting window sizes). The results also show that RMS map can achieve a good authentication performance in a small time window (50 ms), which implies its advantage in the authentication time.

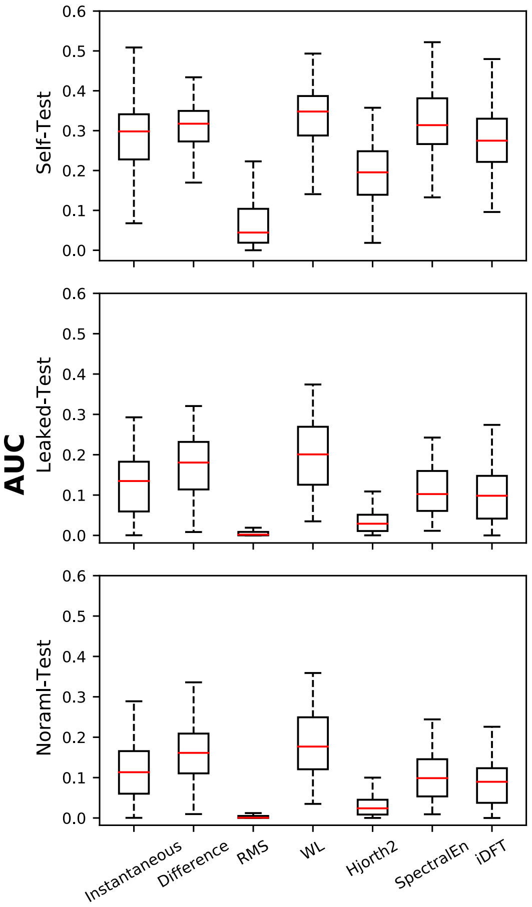

Fig. 5. Performance comparison of different methods without majority voting in terms of AUC for the CapgMyo-DB-a dataset

2) Results on SD-sEMG Dataset: We then present the results obtained from the SD-sEMG dataset, Ninapro DB-1. Performance of different methods in terms of AUC and EER are summarized in Figs. 8 and 9, respectively. The results are similar to those obtained from the CapgMyo DB-a dataset that RMS map significantly outperforms other methods in the three test scenarios (all p-values $< .001$). All methods also have worse performance in the self test compared with other two tests. The EERs obtained by the proposed method with

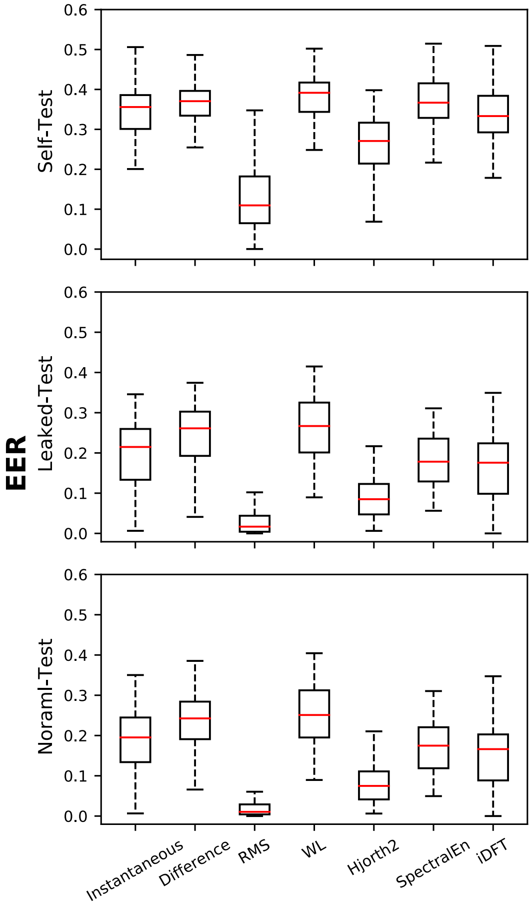

Fig. 6. Performance comparison of different methods without majority voting in terms of EER for the CapgMyo-DB-a dataset

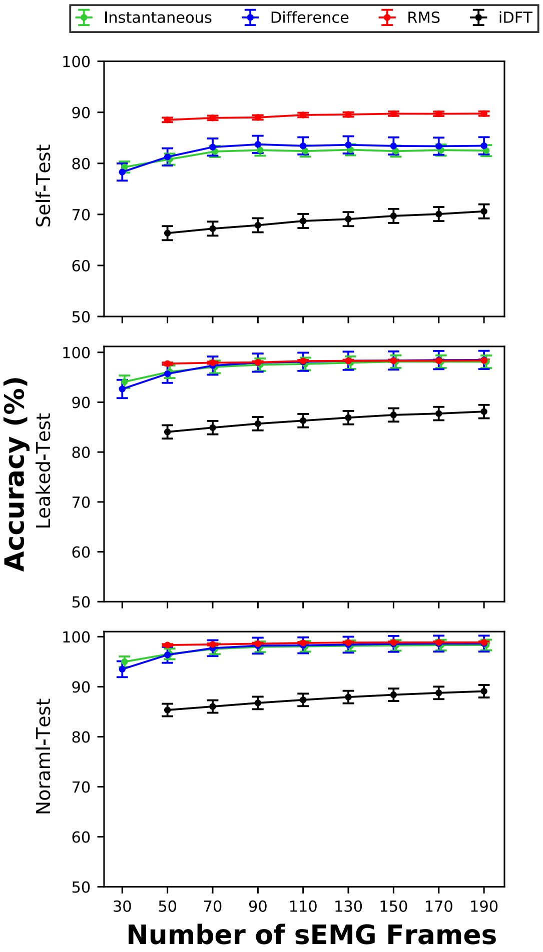

Fig. 7. Performance comparison of different methods with majority voting in terms of accuracy for the CapgMyo DB-a dataset

RMS map for the normal, leaked, and self tests are 6.18±6.99 %, 7.30±7.52 %, and 12.86±11.62 %, respectively. Compared with the results shown in Figs. 5 and 6, it can be observed that using SD-sEMG signals result in a larger variation in results than using HD-sEMG signals for all methods. This shows that HD-sEMG signals provide richer spatio-temporal information than SD-sEMG signals leading to a more stable authentication performance.

Same as Fig. 7, we compare the performance of four different methods with majority voting. The results are shown in Fig. 10. Unlike the results obtained from the CapgMyo DB-a dataset, major voting has a great impact on improving the performance of RMS map in all three test scenarios when using the SD-sEMG signals. The most significant impact is seen in the self test where the average accuracy of RMS map has an increment of 9.67 % from 50 ms to 170 ms. The results show that RMS map significantly outperforms other methods in all three test scenarios under different voting window sizes (all p-values $<$ .01). This demonstrates that the proposed method with RMS map is a viable solution for user authentication even when only SD-sEMG signals of hand gestures are acquired.

V. Conclusions

Multi-channel sEMG signals are widely employed for hand gesture recognition but have only seen limited applications for user authentication. In this paper, we propose a new use authentication method based on deep anomaly detection using multi-channel sEMG signals of hand gestures. Different methods to convert multi-channel sEMG signals as a sequence of images are investigated. The sEMG images are fed into a GANomaly-based deep anomaly detection model which classifies the user performing the hand gesture as client or imposter. We evaluate the performance of the proposed method under three test scenarios using two publicly available hand gesture sEMG datasets: an HD-sEMG dataset, CapgMyo DB-a, and a SD-sEMG dataset, Ninapro DB-1. The results show that RMS map has the best performance in all test scenarios among the investigated sEMG image generation methods. The proposed method with RMS map achieves much better performance than the reference method, especially when using SD-sEMG signals. It is found that the majority voting strategy can generally improve the authentication performance of the proposed method. Overall, the results demonstrate that the proposed method with RMS map is a viable solution for user authentication. In future, a fine-tuning of the hyperparameters

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TIM.2022.3164162, IEEE Transactions on Instrumentation and Measurement

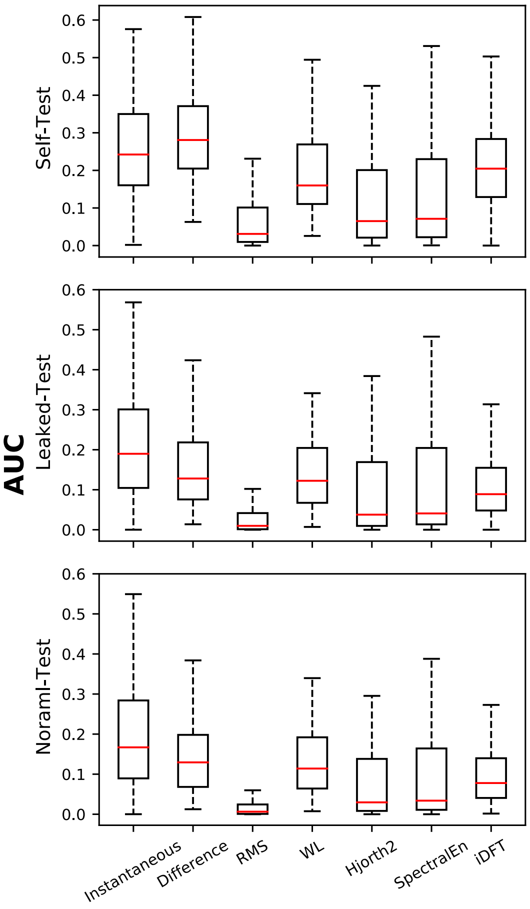

Fig. 8. Performance comparison of different methods without majority voting in terms of AUC for the Ninapro DB-1 dataset

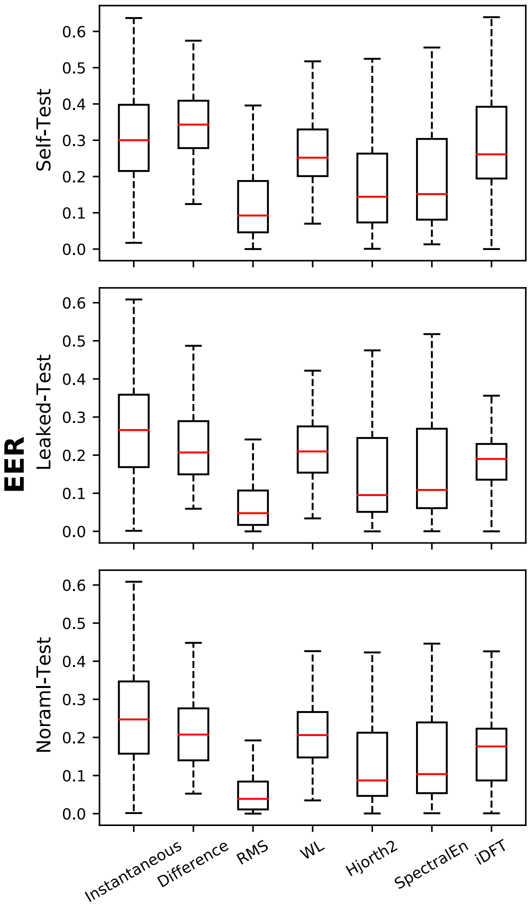

Fig. 9. Performance comparison of different methods without majority voting in terms of EER for the Ninapro DB-1 dataset

of the GANomaly model will be conducted to explore possible performance improvement. We will also research novel deep anomaly detection models that capture more discriminative spatio-temporal information in multi-channel sEMG signals, which could lead to improved authentication performance in the self test and when SD-sEMG signals are used.

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] X. Wang, Z. Yan, R. Zhang, and P. Zhang, "Attacks and defenses in user authentication systems: A survey," *Journal of Network and Computer Applications*, p. 103080, 2021.

[3] A. J. Aviv, K. L. Gibson, E. Mossop, M. Blaze, and J. M. Smith, "Smudge attacks on smartphone touch screens." *Woot*, vol. 10, pp. 1–7, 2010.

[4] 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.

[5] S. Naval, A. Pandey, S. Gupta, G. Singal, V. Vinoba, and N. Kumar, "PIN inference attack: A threat to mobile security and smartphone-controlled robots," *IEEE Sensors Journal*, 2021.

[6] Z. Rui and Z. Yan, "A survey on biometric authentication: Toward secure and privacy-preserving identification," *IEEE Access*, vol. 7, pp. 5994–6009, 2018.

[7] A. J. Bidgoly, H. J. Bidgoly, and Z. Arezoumand, "A survey on methods and challenges in EEG based authentication," *Computers & Security*, vol. 93, p. 101788, 2020.

[8] M. Hosseinzadeh, B. Vo, M. Y. Ghafour, and S. Naghipour, "Electrocardiogram signals-based user authentication systems using soft computing techniques," *Artificial Intelligence Review*, vol. 54, no. 1, pp. 667–709, 2021.

[9] D. D. Chakladar, P. Kumar, P. P. Roy, D. P. Dogra, E. Scheme, and V. Chang, "A multimodal-siamese neural network (mSNN) for person verification using signatures and EEG," *Information Fusion*, vol. 71, pp. 17–27, 2021.

[10] X. Zhang, L. Yao, C. Huang, T. Gu, Z. Yang, and Y. Liu, "Deepkey: A multimodal biometric authentication system via deep decoding gaits and brainwaves," *ACM Transactions on Intelligent Systems and Technology*, vol. 11, no. 4, pp. 1–24, 2020.

[11] 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.

[12] Z. Zhao, Y. Zhang, Y. Deng, and X. Zhang, "ECG authentication system design incorporating a convolutional neural network and generalized S-transformation," *Computers in Biology and Medicine*, vol. 102, pp. 168–179, 2018.

[13] R.-H. Liang and M. Ouhyoung, "A real-time continuous gesture recognition system for sign language," in *Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, 1998, pp. 558–567.

[14] C. Frigo and P. Crenna, "Multichannel sEMG in clinical gait analysis: a review and state-of-the-art," *Clinical Biomechanics*, vol. 24, no. 3, pp. 236–245, 2009.

[15] C. Zhu and W. Sheng, "Wearable sensor-based hand gesture and daily activity recognition for robot-assisted living," *IEEE Transactions on*

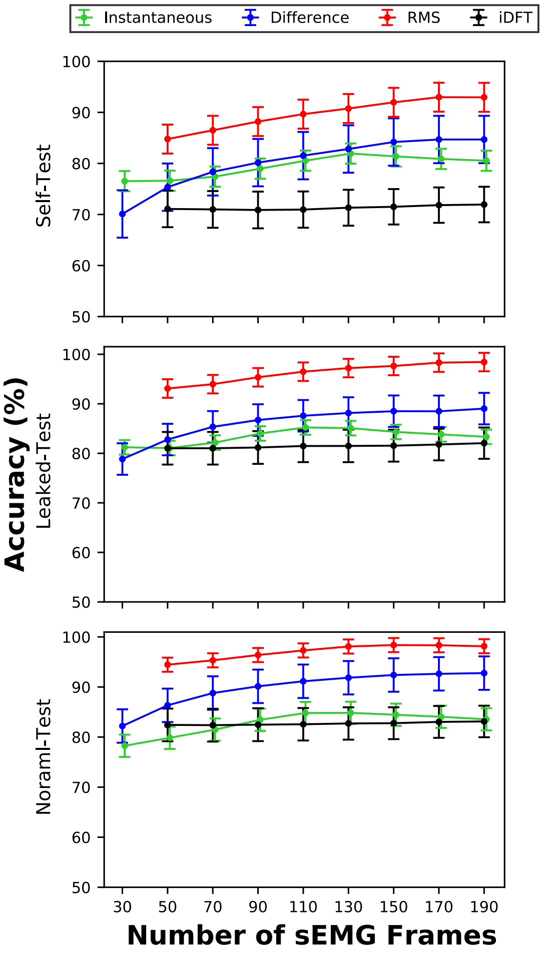

Fig. 10. Performance comparison of different methods with majority voting in terms of accuracy for the Ninapro DB-1 dataset

Systems, Man, and Cybernetics-Part A: Systems and Humans, vol. 41, no. 3, pp. 569–573, 2011.

[16] G.-C. Luh, Y.-H. Ma, C.-J. Yen, and H.-A. Lin, “Muscle-gesture robot hand control based on sEMG signals with wavelet transform features and neural network classifier,” in *2016 International Conference on Machine Learning and Cybernetics (ICMLC)*, vol. 2, 2016, pp. 627–632.

[17] W.-J. Li, C.-Y. Hsieh, L.-F. Lin, and W.-C. Chu, “Hand gesture recognition for post-stroke rehabilitation using leap motion,” in *2017 International Conference on Applied System Innovation (ICASI)*, 2017, pp. 386–388.

[18] K. M. Sagayam and D. J. Hemanth, “Hand posture and gesture recognition techniques for virtual reality applications: a survey,” *Virtual Reality*, vol. 21, no. 2, pp. 91–107, 2017.

[19] R. Meattini, S. Benatti, U. Scarcia, D. De Gregorio, L. Benini, and C. Melchiorri, “An sEMG-based human–robot interface for robotic hands using machine learning and synergies,” *IEEE Transactions on Components, Packaging and Manufacturing Technology*, vol. 8, no. 7, pp. 1149–1158, 2018.

[20] A. Prakash, B. Kumari, and S. Sharma, “A low-cost, wearable sEMG sensor for upper limb prosthetic application,” *Journal of Medical Engineering & Technology*, vol. 43, no. 4, pp. 235–247, 2019.

[21] 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.

[22] G. Pang, C. Shen, L. Cao, and A. V. D. Hengel, “Deep learning for anomaly detection: A review,” *ACM Computing Surveys*, vol. 54, no. 2, Mar. 2021.

[23] Y. Xu and Y. Dai, “Review of hand gesture recognition study and application,” *Contemporary Engineering Sciences*, vol. 10, no. 8, pp. 375–384, 2017.

[24] Y. Han, “A low-cost visual motion data glove as an input device to interpret human hand gestures,” *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 501–509, 2010.

[25] M. A. Oskoei and H. Hu, “Myoelectric control systems—a survey,” *Biomedical Signal Processing and Control*, vol. 2, no. 4, pp. 275–294, 2007.

[26] F. Duan, X. Ren, and Y. Yang, “A gesture recognition system based on time domain features and linear discriminant analysis,” *IEEE Transactions on Cognitive and Developmental Systems*, 2018.

[27] A. A. Neacsu, G. Cioroiu, A. Radoi, and C. Burileanu, “Automatic EMG-based hand gesture recognition system using time-domain descriptors and fully-connected neural networks,” in *2019 42nd International Conference on Telecommunications and Signal Processing (TSP)*, 2019, pp. 232–235.

[28] U. C. Allard, F. Nougarou, C. L. Fall, P. Giguère, C. Gosselin, F. Laviolette, and B. Gosselin, “A convolutional neural network for robotic arm guidance using sEMG based frequency-features,” in *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 2464–2470.

[29] Y. Hu, Y. Wong, W. Wei, Y. Du, M. Kankanhalli, and W. Geng, “A novel attention-based hybrid CNN-RNN architecture for sEMG-based gesture recognition,” *PloS One*, vol. 13, no. 10, p. e0206049, 2018.

[30] C. Liu, Y. Yang, X. Liu, L. Fang, and W. Kang, “Dynamic-hand-gesture authentication dataset and benchmark,” *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1550–1562, 2020.

[31] W. H. Khoh, Y. H. Pang, and A. B. J. Teoh, “In-air hand gesture signature recognition system based on 3-dimensional imagery,” *Multimedia Tools and Applications*, vol. 78, no. 6, pp. 6913–6937, 2019.

[32] B. Lewis, C. J. Nycz, G. S. Fischer, and K. K. Venkatasubramanian, “Authentication-based on biomechanics of finger movements captured using optical motion-capture,” in *International Symposium on Visual Computing*. Springer, 2018, pp. 167–179.

[33] X. Wang and J. Tanaka, “GesID: 3D gesture authentication based on depth camera and one-class classification,” *Sensors*, vol. 18, no. 10, p. 3265, 2018.

[34] S. Fong, Y. Zhuang, and I. Fister, “A biometric authentication model using hand gesture images,” *Biomedical Engineering Online*, vol. 12, no. 1, pp. 1–18, 2013.

[35] A. M. H. Wong and D.-K. Kang, “Stationary hand gesture authentication using edit distance on finger pointing direction interval,” *Scientific Programming*, vol. 2016, 2016.

[36] A. K. Jain, A. A. Ross, and K. Nandakumar, *Introduction to biometrics*. Springer Science & Business Media, 2011.

[37] H. Yamaba, A. Kurogi, S.-I. Kubota, T. Katayama, M. Park, and N. Okazaki, “Evaluation of feature values of surface electromyograms for user authentication on mobile devices,” *Artificial Life and Robotics*, vol. 22, no. 1, pp. 108–112, 2017.

[38] H. Yamaba, K. Aburada, T. Katayama, M. Park, and N. Okazaki, “Evaluation of user identification methods for realizing an authentication system using s-EMG,” in *International Conference on Network-Based Information Systems*, 2018, pp. 733–742.

[39] 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.

[40] A. Pradhan, J. He, and N. Jiang, “Performance optimization of surface electromyography based biometric sensing system for both verification and identification,” *IEEE Sensors Journal*, 2021.

[41] Q. Li, P. Dong, and J. Zheng, “Enhancing the security of pattern unlock with surface EMG-based biometrics,” *Applied Sciences*, vol. 10, no. 2, p. 541, 2020.

[42] X. Jiang, K. Xu, X. Liu, C. Dai, D. Clifton, E. A. Clancy, M. Akay, and W. Chen, “Cancelable HD-sEMG-based biometrics for cross-application discrepant personal identification,” *IEEE Journal of Biomedical and Health Informatics*, 2020.

[43] X. Jiang, K. Xu, X. Liu, C. Dai, D. A. Clifton, E. A. Clancy, M. Akay, and W. Chen, “Neuromuscular password-based user authentication,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2641–2652, 2020.

[44] 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.

[45] A. Bhattacharya, A. Sarkar, and P. Basak, “Time domain multi-feature extraction and classification of human hand movements using surface EMG,” in *2017 4th International Conference on Advanced Computing and Communication Systems (ICACCS)*. IEEE, 2017, pp. 1–5.

[46] I. Elamvazuthi, N. Duy, Z. Ali, S. Su, M. A. Khan, and S. Parasuraman, "Electromyography (EMG) based classification of neuromuscular disorders using multi-layer perceptron," *Procedia Computer Science*, vol. 76, pp. 223–228, 2015.

[47] R. M. Tello, T. Bastos-Filho, A. Frizera-Neto, S. Arjunan, and D. K. Kumar, "Feature extraction and classification of sEMG signals applied to a virtual hand prosthesis," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2013, pp. 1911–1914.

[48] D. Karabulut, F. Ortes, Y. Z. Arslan, and M. A. Adli, "Comparative evaluation of EMG signal features for myoelectric controlled human arm prosthetics," *Biocybernetics and Biomedical Engineering*, vol. 37, no. 2, pp. 326–335, 2017.

[49] X. Jiang, X. Liu, J. Fan, X. Ye, C. Dai, E. A. Clancy, D. Farina, and W. Chen, "Enhancing IoT security via cancelable HD-sEMG-based biometric authentication password, encoded by gesture," *IEEE Internet of Things Journal*, vol. 8, no. 22, pp. 16 535–16 547, 2021.

[50] W. Caesarendra, S. U. Lekson, K. A. Mustaqim, A. R. Winoto, and A. Widyotriatmo, "A classification method of hand EMG signals based on principal component analysis and artificial neural network," in *2016 International Conference on Instrumentation, Control and Automation (ICA)*. IEEE, 2016, pp. 22–27.

[51] X. Jiang, K. Xu, X. Liu, C. Dai, D. A. Clifton, E. A. Clancy, M. Akay, and W. Chen, "Neuromuscular password-based user authentication," *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2641–2652, 2020.

[52] 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.

[53] Z. Luo, R. Zuo, Y. Xiong, and X. Wang, "Detection of geochemical anomalies related to mineralization using the GANomaly network," *Applied Geochemistry*, vol. 131, p. 105043, 2021.

[54] J. Du, L. Guo, L. Song, H. Liang, and T. Chen, "Anomaly detection of aerospace facilities using GANomaly," in *Proceedings of the 2020 5th International Conference on Multimedia Systems and Signal Processing*, 2020, pp. 40–44.

[55] K. Yan, "Chiller fault detection and diagnosis with anomaly detective generative adversarial network," *Building and Environment*, p. 107982, 2021.

[56] Z. Wan, J. Ma, N. Qin, Z. Zhou, and D. Huang, "Fault detection of air-spring devices based on GANomaly and isolated forest algorithms," in *2021 IEEE 16th Conference on Industrial Electronics and Applications (ICIEA)*. IEEE, 2021, pp. 1328–1333.

[57] B. Zhou, S. Liu, B. Hooi, X. Cheng, and J. Ye, "Beatgan: Anomalous rhythm detection using adversarially generated time series." in *IJCAI*, 2019, pp. 4433–4439.

[58] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 2015.

[59] M. Atzori, A. Gijsberts, C. Castellini, B. Caputo, A.-G. M. Hager, S. Elsig, G. Giatsidis, F. Bassetto, and H. Müller, "Electromyography data for non-invasive naturally-controlled robotic hand prostheses," *Scientific Data*, vol. 1, no. 1, pp. 1–13, 2014.

[60] M. Atzori, A. Gijsberts, I. Kuzborskij, S. Elsig, A.-G. M. Hager, O. Deriaz, C. Castellini, H. Müller, and B. Caputo, "Characterization of a benchmark database for myoelectric movement classification," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 23, no. 1, pp. 73–83, 2014.

Qingqing Li was born in Hubei, China, in 1992. She received the B.S. degree in Software Engineering from Yangtze University, China, in 2016, and the M.S. degree in Computer Science from New Mexico Institute of Mining and Technology, USA, in 2020. Currently, she is pursuing the Ph.D. degree in Computer Science at New Mexico Institute of Mining and Technology. Her current research interests include mobile security, biometrics, machine learning, and deep learning.

Zhirui Luo was born in Changde, Hunan, China, in 1994. He received the B.E. degree in Software Engineering from Yangtze University, China, in 2018, and the B.S. and M.S. degrees in Computer Science from New Mexico Institute of Mining and Technology, USA, in 2018 and 2020, respectively. He is currently pursuing the Ph.D. degree in Computer Science at New Mexico Institute Mining and Technology. His research areas are machine learning, deep learning, and social computing.

Jun Zheng (Member, IEEE) received the B.S. and M.S. degrees in Electrical Engineering from Chongqing University, China, in 1993 and 1996, respectively. He received the M.S.E. degree in Biomedical Engineering from Wright State University, USA, in 2001, and the Ph.D. degree in Computer Engineering from University of Nevada, Las Vegas, USA, in 2005. From 2005 to 2008, he was an Assistant Professor with the Department of Computer Science, Queens College, The City University of New York, USA. He is currently a Professor in the Department of Computer Science and Engineering at New Mexico Institute of Mining and Technology, USA, where he is leading the Human-Centered Computing and Security Lab. His current research interests include mobile security, smart grid security and privacy, social computing, machine learning, and deep learning.