# Cancellable Template Design for Privacy-Preserving EEG Biometric Authentication Systems

Min Wang, *Member, IEEE*, Song Wang, and Jiankun Hu, *Senior Member, IEEE*

***Abstract*— As a promising candidate to complement traditional biometric modalities, brain biometrics using electroencephalography (EEG) data has received a widespread attention in recent years. However, compared with existing biometrics such as fingerprints and face recognition, research on EEG biometrics is still in its infant stage. Most of the studies focus on either designing signal elicitation protocols from the perspective of neuroscience or developing feature extraction and classification algorithms from the viewpoint of machine learning. These studies have laid the ground for the feasibility of using EEG as a biometric verification modality, but they have also raised security and privacy concerns as EEG data contains sensitive information. Existing research has used hash functions and cryptographic schemes to protect EEG data, but they do not provide functions for revoking compromised templates as in cancellable template design. This paper proposes the first cancellable EEG template design for privacy-preserving EEG-based verification systems, which can protect raw EEG signals containing sensitive privacy information (e.g., identity, health and cognitive status). A novel cancellable EEG template is developed based on EEG features extracted by a deep learning model and a non-invertible transform. The proposed transformation provides cancellable templates, while taking advantage of EEG elicitation protocol fusion to enhance biometric performance. The proposed verification system offers superior performance than the state-of-the-art, while protecting raw EEG data. Furthermore, we analyze the system's capacity for resisting multiple attacks, and discuss some overlooked but critical issues and possible pitfalls involving hill-climbing attacks, second attacks, and classification-based verification systems.**

***Index Terms*— EEG biometrics, brain biometrics, verification system, privacy-preserving, cancellable biometrics, non-invertible transformation, template protection.**

## I. Introduction

CONVENTIONAL biometric techniques such as fingerprint and face recognition share vulnerabilities in terms of confidentiality and robustness against circumvention [1] since these biometric traits are observable and can be illegally obtained or forged without the user's awareness, e.g., via high resolution photography [2], [3]. The need for stronger security has given birth to brain biometrics based on electroencephalography (EEG) signals. Meanwhile, the rapid development of brain-computer interface, neuroscience, and sensor technology has created an environment where EEG is readily available for biometric applications. Potential advantages of EEG biometrics include its robustness against circumvention, support for liveness detection, continuous verification, and cognitive information indicators [4], [5]. A typical EEG-based biometric recognition system consists of four major modules: signal acquisition, pre-processing, feature extraction, and decision-making, as illustrated in Fig. 1. During data acquisition, EEG signals, captured by sensors from the user's scalp while he or she engages with the elicitation protocol, are transmitted to the processing unit. Since raw EEG data are usually contaminated with noise and artifacts, it is necessary to preprocess the raw data to enhance signal quality. Then discriminant features are extracted from the preprocessed EEG and fed into a decision-making module.

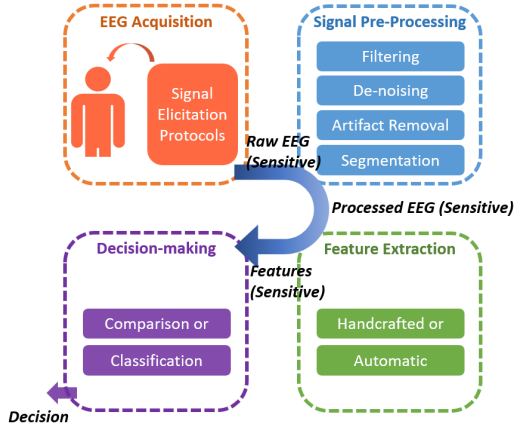


Fig. 1. A typical structure of EEG-based biometric systems.

So far, most studies on EEG biometrics have focused on the improvement of the signal acquisition, feature extraction, and decision-making modules. The acquisition of EEG biometrics requires to specify the corresponding signal elicitation protocols, among which the resting protocol is favorable due to its convenience and minimum requirements for data collection. Ongoing EEG under the resting state protocol does not involve external stimulation to or active response from the user during data acquisition, thus minimizing the impact of cognitive state changes on signal stability [6], [7]. It also supports operation in a continuous and unobtrusive manner. Alternative protocols

Manuscript received 24 February 2022; revised 29 June 2022 and 8 August 2022; accepted 12 August 2022. Date of publication 5 September 2022; date of current version 29 September 2022. This work was supported by the Australian Research Council Discovery Grant DP200103207. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Issa Traore. *(Corresponding author: Jiankun Hu.)*

Min Wang and Jiankun Hu are with the School of Engineering and Information Technology, University of New South Wales, Canberra, ACT 2612, Australia (e-mail: maggie.wang1@adfa.edu.au; j.hu@adfa.edu.au).

Song Wang is with the School of Computing, Engineering and Mathematical Sciences, La Trobe University, Melbourne, VIC 3086, Australia (e-mail: song.wang@latrobe.edu.au).

Digital Object Identifier 10.1109/TIFS.2022.3204222

include the volitional tasks, e.g., the pass-thoughts and various event-related potential (ERP) protocols [8]. The decision-making is achieved by template comparison or classification based on supervised learning models. Template comparison is adopted in many studies as an effective and solid decision-maker in both identification and verification scenarios [7], [8], [9], [10]. Other studies view identification or verification as a classification problem and explore different machine learning models, such as discriminant analysis [11], [12], support vector machines, and neural networks [5], [13]. However, user verification is not merely a classification problem, but also entails security considerations. Classification accuracy does not necessarily reflect real biometric performance. Unfortunately, many studies fail to differentiate the two concepts. We will discuss this issue in Section VI.

A more serious concern is that EEG biometric systems without privacy-preserving mechanisms would pose a huge threat to user privacy. EEG signals contain sensitive information about the user's cognitive and emotional states and health conditions [14]. A recent study examining EEG templates (features) used in biometric applications confirmed that personal characteristics regarding age and gender, as well as information related to medication intake and neurological disorders, can be inferred from the templates [15]. These findings highlight the need to apply privacy-preserving mechanisms to protect user templates when deploying EEG biometric systems. However, the issue of protecting EEG biometric systems from privacy and security breaches has not been fully resolved. The contributions of this study are summarized as follows:

- A cancellable template design is proposed to attain a privacy-preserving EEG-based verification system.
- An innovative transformation is designed to generate cancellable templates from EEG features encoded by a deep neural network via a non-invertible transform. The proposed transformation is tailored for EEG biometrics allowing for elicitation protocol fusion to enhance verification performance, while providing template protection.
- A new concept of second attacks is introduced to examine the possibility of breaking into a system using pre-obtained solutions after the system has revoked the compromised template.
- Pre-image and hill-climbing attacks are widely used criteria to assess cancellable biometrics. We reveal that these criteria do not fully apply to security assessment of cancellable biometrics. Hence, we re-define the concept of pre-image attacks suitable for cancellable biometrics.
- Extensive experiments are carried out to evaluate the effects of pre-image and hill-climbing attacks. The results demonstrate that cancellable template design based on many-to-one mapping is inherently resistant to these attacks, which is contrary to the common understanding in the field.
- In-depth analysis is conducted on pitfalls involved in the evaluation procedure of supervised learning-based verification systems.

The rest of this paper is organized as follows. Section II reviews the state of the art on EEG biometrics and protection mechanisms. Section III presents the proposed methodology. Section IV describes the experimental design, followed by results in Section V and security analysis and discussion in Section VI. The conclusion and future directions are summarized in Section VII.

## II. Related Work

### A. EEG Biometrics

EEG under the resting state has been investigated for biometric applications for over a decade and recent studies on the permanence issue suggested that the resting state protocol presents an effective and robust condition for biometric recognition [6], [9]. In the resting state elicitation protocol, the user remains relaxed with eyes closed (EC) or eyes open (EO) without performing any particular task. The rationale behind it, besides its implementation simplicity, is the neurophysiological evidence which indicates that ongoing EEG under the resting state carries unique identity information (e.g., those related to heritability and personality factors) [16]. In addition, EEG signals present large intra-user variations that could hinder the biometric performance. In order to improve system performance and robustness, the fusion of multiple elicitation protocols is adopted. In many works, this is achieved by decision-level fusion through voting schemes [8]. Another way for protocol fusion is to mix the EEG data collected under different elicitation protocols to create a data set that contains the generalized unique pattern of each user [13]. This strategy has been adopted in many EEG biometrics studies to account for the intra-class variability, especially methods based on supervised learning models [12].

Regarding feature extraction, different methods are proposed considering the various characteristics of EEG signals. Based on whether the relationship information between signals of different channels is captured or not, we can categorize EEG features into univariate features and bivariate features. The univariate features are extracted from single channels of signals considering signal characteristics in the time and frequency domains. Popular ones include the coefficients of autoregressive (AR) models [9], [17], fuzzy entropy [18], and power spectral density (PSD) features [7], [9], which reflect time-dependency, dynamic complexity, and spectral characteristics of EEG, respectively. On the other hand, the bivariate features are based on brain connectivity which captures the interactive or structural information between EEG channels. Different statistical and effective metrics have been used for establishing connectivity between EEG channels, including the Pearson's correlation [5], [19], Granger causality [20], spectral coherence [7], and phase synchronization indices [10], [19]. Moreover, graph features extracted from the brain connectivity networks are also proposed for EEG biometrics [10], [21]. Recent findings suggest that, compared with univariate features, bivariate features are more robust against the intra-user variations across sessions, thus improving biometric performance [5], [10]. The result also shows that the phase synchronization, especially the $\rho$ index, is a sound metric to estimate EEG connectivity for biometric applications. Feature extraction based on deep learning model is also proposed [22], [23], [24].

For classification in EEG biometrics, existing methods can be categorized into comparison-based classification and supervised learning-based classification. In verification, the comparison-based methods predict the class label (genuine user or impostor) of a probe template by calculating its similarity to one or multiple reference templates of the claimed user. The similarity was defined by the Euclidean distance [21], Mahalanobis distance [7], Manhattan distance [9], cosine similarity [9], [17], and cross-correlation [8]. Template comparison is straightforward and computationally fast, yielding interpretable results. The performance depends on the discriminative capacity of the template. Recent studies also explored different machine learning algorithms for classification in EEG biometrics. Popular classifiers include the linear discriminant analysis (LDA) [11], [12], [17], support vector machines (SVMs) [17], and deep neural networks such as multilayer perceptron (MLP) and convolutional neural networks (CNNs) [5], [13]. In these methods, training is an essential step that fits the model to a training dataset. The performance of the model not only depends on the capability of the model itself, but also relies on the training procedure and a good training dataset.

### B. Privacy-Preserving Mechanisms

Non-invertible transformation design for biometric systems renders a vital privacy-preserving mechanism for biometric template protection. This type of method applies a one-way transformation to biometric data such that an adversary cannot retrieve the original biometric data, even if the stored template is compromised. The comparison or classification of the enrolled template and the probe is carried out in the transformed domain to protect the original biometric data from leakage. He *et al.* [25] studied the potential of hashing EEG features for verification. Multi-variate autoregressive coefficients were extracted as features from multi-channel EEG signals and then hashed by the fast Johnson-Lindenstrauss algorithm to obtain compact hash vectors. A naive Bayes probabilistic model was used for decision-making based on the EEG hash vectors. Applying cryptographic hashing to biometrics induces variation, as any slight change to the input would completely alter the hash value produced. Bajwa *et al.* [26] proposed a key generation method with EEG biometrics. The PSD features were extracted from EEG signals using the discrete Fourier transform and discrete wavelet transform, followed by a Neurokey generation procedure which involves feature selection, binarization and hashing. The term 'cancellable' is used in this study to mean that a user's Neurokey can be changed by using the EEG collected in a different cognitive task, if the user's biometric information is compromised. However, such 'cancellable biometrics' cannot protect raw EEG data containing sensitive information. Furthermore, the choice of tasks is limited and different tasks would have vastly different performance [27]. Damaševičius *et al.* [28] developed a cryptographic verification scheme for EEG biometrics using fuzzy commitment and the error-correcting Bose-Chaudhuri-Hocquenghem codes. Although this method protects data privacy, it is not equipped with cancellability to revoke compromised templates. Cognitive biometric cryptosystems based on EEG are also proposed [29]. Cancellable biometric templates based on non-invertible transforms offer a solution to EEG data protection as well as template revocability [30], [31], [32], [33]. To the best of our knowledge, there has been no cancellable EEG template design reported. In EEG biometric systems, most of the work is based on classification models, where it is infeasible to integrate cancellability. Once the model is compromised, the input can be estimated by genetic algorithms so that the system is cracked.

## III. Methodology

In this section, we design a cancellable template to protect EEG biometric data. The proposed privacy-preserving EEG verification system consists of four main components: signal acquisition, feature extraction, feature transformation, and comparison, as illustrated in Fig. 2. In the enrollment stage, EEG signals are collected from each user under the signal elicitation protocol and fed into the feature extraction module, which encodes signals into feature vectors. Then the transformation module takes the features as the input and creates a cancellable template with a user-associated key. This template is a binary representation, and will be stored in the database. In the verification stage, a probe template is generated following the same procedure and the comparison algorithm then outputs a decision to accept or reject the user.

### A. Signal Elicitation Protocol

The resting state EEG elicitation protocol is adopted for signal acquisition. To be specific, two conditions are included, namely the EO and EC states. The user is asked to stay relaxed with eyes closed or eyes open, while the spontaneous EEG signals are recorded. EEG signals present time-varying and non-stationary characteristics, and are sensitive to the cognitive states of the subject, which may affect the biometric performance. Therefore, in order to improve stability, researchers often consider elicitation protocol fusion to get a richer dataset that contains signals in diverse states. We adopt the basic idea of elicitation protocol fusion [13]. However, instead of decision-level fusion with majority voting or directly mixing data collected under different protocols to form training and testing sets as in the existing research, as shown is Fig. 3 (a) and (b), we embed data fusion naturally in the transformation process, as illustrated in Fig. 3 (c). The benefits of our design are twofold: 1) the entropy of extracted features increases, thus the reliability of the biometric system is enhanced, due to the elicitation protocol fusion; and 2) secure cancellable templates are generated at the same time without extra computational costs. Details of the transformation is presented in Section III-C.

### B. Feature Extraction

The state-of-the-art feature extraction method [23], [24] is adopted in our study, where a siamese CNN model is designed to derive discriminative features from the raw EEG time series. The siamese network, as illustrated in Fig. 4, contains two identical CNN subnetworks that share the same architecture,

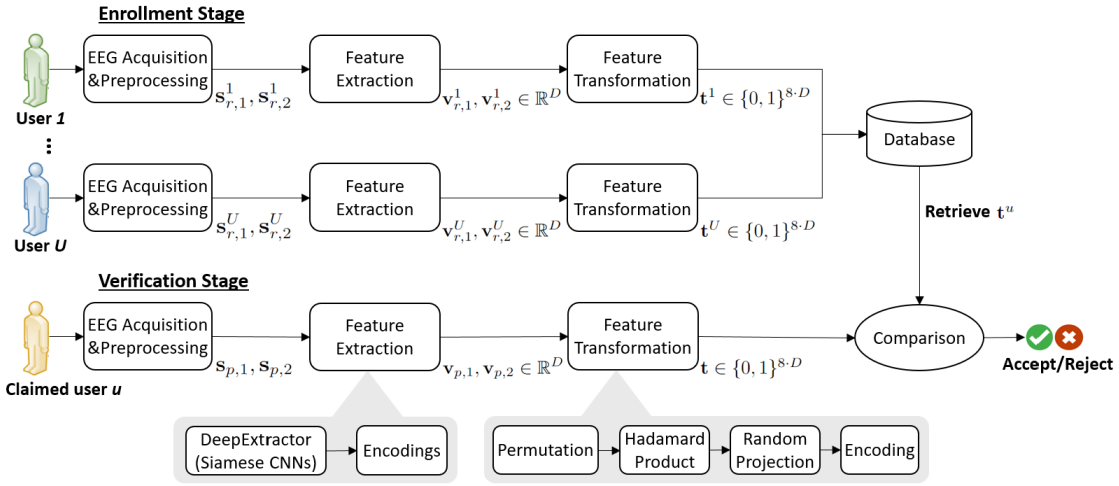


Fig. 2. The proposed privacy-preserving EEG-based verification system.

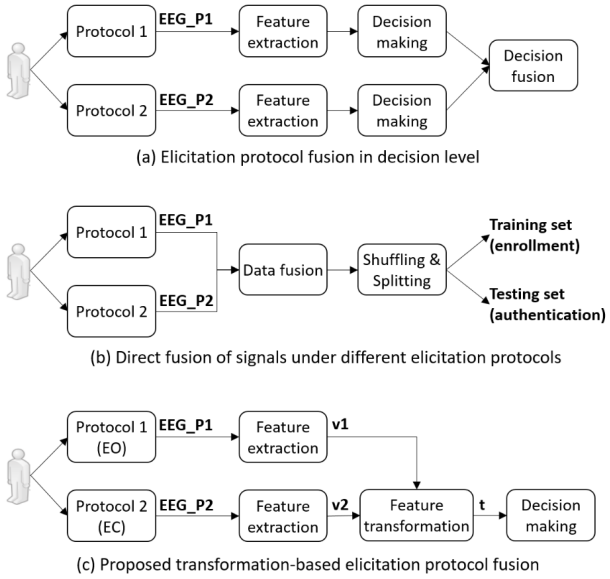


Fig. 3. Signal elicitation protocol fusion.

TABLE I
CONFIGURATION OF THE CNN SUBNETWORKS

| Layer | Filters /Units | Kernel | Activation |
|---|---|---|---|
| Conv | 16 | (1, 5) | ReLu |
| MaxPooling | - | (1, 3) | - |
| Conv | 32 | (1, 5) | ReLu |
| MaxPooling | - | (1, 3) | - |
| Conv | 64 | (1, 3) | ReLu |
| MaxPooling | - | (1, 3) | - |
| Conv | 128 | (1, 3) | ReLu |
| MaxPooling | - | (1, 3) | - |
| Dropout (0.5) | - | - | - |
| Flatten | - | - | - |
| Dense | D | - | ReLu |

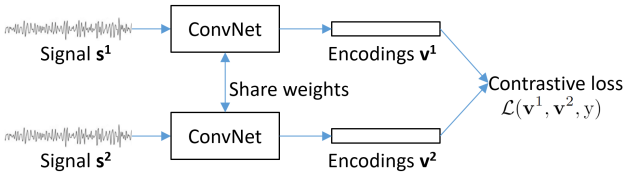


Fig. 4. Siamese network for feature extraction.

parameters and weights. Any parameter updates are mirrored across both subnetworks during the learning process. The input of the siamese network is a pair of signals ($\mathbf{s}_1$, $\mathbf{s}_2$) and the final layers in the subnetworks output encodings ($\mathbf{v}_1$, $\mathbf{v}_2$), where the Euclidean distance between the two encodings, $D^2_{\mathbf{v}^1,\mathbf{v}^2}$, is computed to adjust the weights of the subnetworks. The contrastive loss function is adopted for backpropagation since it is suitable for evaluating how well the siamese network differentiate the input pairs by minimizing distance between encodings derived from signals of the same class and maximizing distance between encodings derived from signals of different classes. The contrastive loss is computed as:

$$\mathcal{L}(\mathbf{v}^1, \mathbf{v}^2, \mathrm{y}) = (1-y)\frac{1}{2}D^2_{\mathbf{v}^1,\mathbf{v}^2} + \mathrm{y}\frac{1}{2}\{max(0, m - D^2_{\mathbf{v}^1,\mathbf{v}^2}\}^2, \tag{1}$$

where y is the label associated with the input pair, with 1 meaning a matching pair (two signals are from the same subject) and 0 a non-matching pair (two signals are from different subjects), and $m$ is the margin that defines the baseline for distance for which pairs should be classified as dissimilar ($m$=1). The configuration of the CNN subnetworks is presented in Table I. After training, the model will be used as a feature extractor to derive a feature vector of length $D$ (here we set $D$ = 1000) from an input signal.

The Adam optimizer with a learning rate of 5e-4 is used for training the siamese network. The batch size is 256, and an early stopping regularization monitoring the validation accuracy with patience of 50 epochs is adopted to avoid overfitting. To train this feature extractor, we split the available subjects into a training set consisting of two-thirds of the subjects and a testing set comprising the remaining subjects, upon which the verification performance is evaluated.

### C. Feature Transformation

Let $\mathbf{v}_1$ and $\mathbf{v}_2$ denote the feature vectors extracted from EEG signals collected under the two elicitation protocols. The

proposed transformation takes the input of $\mathbf{v}_1$ and $\mathbf{v}_2$ and generates a secure cancellable template $\mathbf{t}$, as illustrated in Fig. 3 (c). Both feature vectors have a length of $D$.

At the enrollment stage, each user is assigned a user key $k$ (the key is stored along the template in the database), which is used as the random seed to generate a random permutation of the integers 1 to $D$, as follows:

$$\mathbf{p} = randperm(k, D), \tag{2}$$

where the randperm(seed,integer) is a random permutation function defined on a Pseudorandom number generator (PRNG) which can adopt any generic PRNG algorithm, e.g., the Mersenne Twister algorithm. Then a permuted (re-arranged) version of the feature vector $\mathbf{v}_1$ is obtained and the Hadamard product of it and $\mathbf{v}_2$ is calculated, as follows:

$$\mathbf{v}_1' = \mathbf{v}_1(\mathbf{p}) \tag{3}$$

$$\mathbf{c} = \mathbf{v}_1' \circ \mathbf{v}_2. \tag{4}$$

A vector $\mathbf{r}$ is then generated by projecting vector $\mathbf{c}$ with a matrix $\mathbf{M}$, as follows:

$$\mathbf{r} = \mathbf{c} \cdot \mathbf{M}, \tag{5}$$

where $\mathbf{M}$ is a user-specific random projection matrix with more rows than columns to form an underdetermined system of equations, thus making the transformation non-invertible. Finally, the real-valued vector $\mathbf{r}$ is encoded into a binary template $\mathbf{t}$ through the 8-bit Gray code. The coding process first converts a real vector into a decimal vector which is then converted into the Gray code according to the code book, as in Algorithm 2. The binary template $\mathbf{t}$ is stored in the system for comparison purposes.

EEG is a continuous signal in nature and a moving window of short length is usually applied to segment the data sources into frames for preprocessing and feature extraction. It is natural to use the multiple frames captured during data collection, instead of a single frame, to generate a more stable template. Let $F_e$ and $F_t$ denote the number of frames collected during enrollment and verification, respectively. Each frame corresponds to a vector $\mathbf{r_f}$, hence, for $F$ vectors, $\mathbf{r}_f$ (where $f = 1, \cdots, F$) are obtained. The final template generated by the transformation module is the Gray encoding of the average of these vectors. The complete transformation procedure is summarized in Algorithm 1. The number of frames is adjustable in accordance with application scenarios and requirements.

**Algorithm 1** Transform

**Input** : feature vectors from $F$ frames $\mathbf{v}_{f,1}, \mathbf{v}_{f,2} \in \mathbb{R}^D$, $f = 1, \cdots, F$
user identity $u$
**Output**: template $\mathbf{t}$

**if** *enrollment* **then**
  initialize a key $k_u$
**else**
  retrieve the key $k_u$
**end**
compute $\mathbf{p} \leftarrow randperm(k_u, D)$
compute $\mathbf{M} \leftarrow rand(k_u, [D, \delta D]), \delta \in (0, 1)$
**for** $f = 1$ **to** $F$ **do**
  permutation $\mathbf{v}'_{f,1} \leftarrow \mathbf{v}_{f,1}(\mathbf{p})$
  Hadamard product $\mathbf{c}_f \leftarrow \mathbf{v}'_{f,1} \circ \mathbf{v}_{f,2}$
  projection $\mathbf{r}_f \leftarrow \mathbf{c}_f \cdot \mathbf{M}$
**end**
compute $\mathbf{r} \leftarrow (\sum_{f=1}^{F} \mathbf{r}_f)/F$
encoding $\mathbf{t} \leftarrow GrayCode(\mathbf{r}, 8\text{-bit})$

**Algorithm 2** GrayCode

**Input** : real vector $\mathbf{r}$; bit $M$
**Output**: binary code $\mathbf{t}$

$\mathbf{r} \leftarrow \mathbf{r}/max(\mathbf{r}))$
$\mathbf{r} \leftarrow \mathbf{r} - mean(\mathbf{r}))$
$\mathbf{r} \leftarrow normcdf(\mathbf{r}, 0, std(\mathbf{r})))$
$\mathbf{r} \leftarrow \mathbf{r}$
$\mathbf{r} \leftarrow \mathbf{r} * (2^M)$
$\mathbf{d} \leftarrow round(\mathbf{r})$
$\mathbf{d}(\mathbf{d} == 2^M) \leftarrow 2^M - 1$
$\mathbf{t} \leftarrow dec2gc(\mathbf{d}, M)$

### *D. Transform-Based Comparison*

To verify a user, one or more frames of EEG signals are captured from the user and a probe template is generated following the same procedure as in the registration stage. The Hamming distance is used for comparing the probe template and the reference template (both are binary representations), as follows:

$$d_H(\mathbf{t}_q, \mathbf{t}_r) = sum(\mathbf{t}_q \oplus \mathbf{t}_r), \tag{6}$$

where the symbol $\oplus$ denotes element-wise XOR. Finally, the distance is normalized (percentage of bits that differ) and compared with a pre-defined threshold $\theta$ to make a decision, as follows:

$$\hat{o} = \begin{cases} accept, & \text{if } d_H(\mathbf{t}_q, \mathbf{t}_r) \le \theta \\ reject, & \text{otherwise} \end{cases}. \tag{7}$$

In the analysis, the threshold $\theta$ is automatically adjusted to obtain the equal error rate (EER), which is defined as the error rate when the false match rate (FMR) equals the false non-match rate (FNMR). The FMR reflects the percentage of probe templates in which impostors are incorrectly accepted, and the FNMR reflects the percentage of probe templates in which genuine users are incorrectly rejected [34].

*Remarks:* The proposed transformation provides a concise and elegant solution to the generation of secure and cancellable templates. (i) If a template is compromised, the associated user key can be replaced and a new template can therefore be generated with this new key. (ii) Every time the user key is updated, the random permutation in (3) and the Hadamard product in (4) provide a different set of variables for the random projection in (5). Since the projection matrix is rank-deficient for every set of variables, it is insufficient to inverse the computation, making the system resistant to the ARM. (iii) The transformation takes advantage of EEG signal

TABLE II
DATABASES

| Databases | #Subj. | #Ch | Protocols | Samp. rate | #Sess | Devices |
|---|---|---|---|---|---|---|
| SEEDiv | 15 | 62 | Movie | 200 Hz | 3 | ESI NeuroScan† |
| BED | 21 | 14 | EO EC | 256 Hz | 3 | Emotiv EPOC+‡ |
| MMIDB | 109 | 64 | EO EC | 160 Hz | 1 | Unclear |

†medical-grade ‡consumer-grade

elicitation protocol fusion such that the entropy and reliability of the feature vectors are enhanced. (iv) The encoding procedure is a quantization process alleviating the impact of EEG uncertainties associated with the complexity and variability of brain dynamics.

## IV. Experimental Design

### A. Database and Pre-Processing

Evaluation of the proposed method utilizes three publicly available databases, which are the EEG Motor Movement/Imagery Database (MMIDB) [35], BED database [36], and SEEDiv database [37]. The MMIDB database provides EEG signals collected from 109 subjects in two resting states, EC and EO, and motor imagery tasks including physically opening/closing fists/feet and imagining opening/closing fists/feet without actual body movement. Data acquisition was performed using a BCI2000 system [38] equipped with 64 electrodes with a sampling rate of 160 Hz. The recorded signal is referenced to the earlobes. The BED database contains EEG recordings from 21 individuals under multiple signal elicitation protocols in three sessions. The signals were captured using Emotiv Epoc+, an inexpensive consumer-grade device. The EEG recordings under the resting with eye-open (EO) and eye-closed (EC) protocols are used in this study. The SEEDiv database contains EEG recordings of 15 subjects while watching movie clips in three sessions. This database was originally collected for EEG-based emotion recognition, where the movie clips were used as visual stimuli to induce happiness, sadness, fear and neutral emotions from the subjects. We selected recordings under the neutral emotion for this study. Table II summarizes the details of the three databases.

For signal preprocessing, we apply the Harvard automated processing pipeline for EEG (HAPPE) [39] to remove noise and artifacts originating from muscle and eye movement. It consists of four standard steps, including filtering ([8 30] Hz), bed channel detection and interpolation, artifact component rejection, and common average referencing. The HAPPE pipeline is adopted due to its effectiveness in preprocessing data that are heavily contaminated with noise and artifacts. The alpha and beta bands ([8 30] Hz) are selected since EEG content in these two bands contains most inter-person discriminative characteristics according to existing findings [10], [23]. A downsampling to 100 Hz, 128 Hz and 80 Hz is also applied for SEEDiv, BED, and MMIDB respectively, to improve computational efficiency in the subsequent feature extraction step, considering the Nyquist Shannon sampling theorem. The preprocessed signal is then segmented into two-second frames, so that each frame contains 62, 14, and 64 two-second time series, respectively, for the SEEDiv, BED, and MMIDB databases. Finally, we format the signal frames into an unidimensional representation that concatenates signal time series of each channel.

### B. Evaluation Procedures

*1) Signal Acquisition Protocols:* Two resting state protocols, EO and EC, are selected from databases MMIDB and BED for evaluation. For transformation, when signals under two protocols are available, we enable the transformation-embedded protocol fusion scheme, hence, the input of the transformation module, $\mathbf{v}_1$ and $\mathbf{v}_2$, are feature vectors extracted under the two protocols, respectively. When only one protocol is applied, the feature vector will be divided into two parts of equal length (first half and second half), $\mathbf{v}_1$ and $\mathbf{v}_2$, to fed into the transformation module. For SEEDiv database, the Neutral protocol is selected since the database does not provide data under resting states and Neutral is the most relevant one that available in it.

*2) Handcrafted Features:* We select four representative types of handcrafted features for comparison, which are the reflection coefficients of AR models, band power, fuzzy entropy, and graph features based on the EEG functional connectivity networks. These four types of features capture the time-dependency, power spectral characteristics, dynamic complexity, the functional connectivity characteristics of EEG signals, respectively. They are classic and important EEG features in time, frequency, and space domains, and have been widely used for EEG biometric applications. In the following analysis, we refer to them as AR, PSD, FuzzEn, and Graph, respectively.

- AR: An AR model describes the time-varying processes in EEG by specifying that the value of the timeseries at a certain time depends linearly on its own previous values and on a stochastic term (white noise), i.e., $s(t) = \sum_{i=1}^{p} \theta_i s(t-i) + \varepsilon(t)$, where $\theta$ is the coefficients of the AR model. In this study, we use an AR model of order 5 to fit the signal timeseries, and derive the reflection coefficients as features using the Burg method [40]. The final feature vector has a length of $5 \times N$, where $N$ is the number of channels.
- PSD: We estimate the power spectral density of EEG signals using a non-parametric approach based on the fast Fourier transform. This approach is selected because it directly corresponds to the physical interpretation in terms of EEG rhythms [7]. Based on the PSD, the average band power over the delta, theta, alpha, beta, and gamma bands are extracted as features. The length of the final feature vector is $5 \times N$, where $N$ is the number of channels.
- FuzzEn: Entropy quantifies the amount of uncertainty in the EEG amplitudes. Among the existing entropy estimation methods such as approximate entropy and sample entropy, we select the FuzzEn [18], which was shown to be a more reliable measure than others for biological data, since the uncertainty at the boundaries between classes can provide a shade of ambiguity [41]. The final feature vector has a length of $N$, the number of channels.

- Graph: The $\rho$-index, a phase synchronization measure based on Shannon entropy, is used for computing the functional connectivity from the beta band (13-30) multi-channel EEG signals. This index and frequency band are selected based on previous findings [10]. The calculation of $\rho$ is based on the relative phase of two signals, $\Delta\phi_r(t) = |\phi_{x_i}(t) - \phi_{x_j}(t)| \mod 2\pi$, where $\phi_{x_i}(t)$ and $\phi_{x_j}(t)$ are the instantaneous phases of signals $x_i(t)$ and $x_j(t)$, respectively, calculated by Hilbert transform. For EEG signals of $N$ channels, we compute the $\rho$ connectivity on every two channels to construct an $N \times N$ network, on which nodal and global features are extracted. The features include pagerank centrality for each node, transitivity, modularity, network characteristic path length, global efficiency, network radius and diameter. The final feature vector has a dimension of $N + 6$, where $N$ is the number of channels.

*3) Verification Performance:* We test the proposed transformation on different feature extraction methods, and compare the transformed-domain performance with its corresponding performance in the non-transformed domain. Manhattan distance is used for comparison in the non-transformed domain where the features are real-valued. The non-invertible transformation often needs to reset the order or position of the feature set, which is likely to weaken the discriminant power of the feature set and introduce extra intra-user variations, thus affecting biometric performance [42]. A good cancellable template design should enhance the security of the template without compromising the biometric performance. The non-transformed domain performance is used to show whether and to what extent the proposed transformation has an impact on biometric performance.

### C. Cross-Session Evaluation

Two experimental setups are considered, the within-session evaluation uses data collected in one session for enrollment and verification, and the cross-session evaluation tests verification performance using data collected in a different session than the one used for enrollment. The cross-session stability is important for practical EEG biometric systems. In the cross-session setup, the siamese model is trained using input pairs generated within and across sessions in the training subject set, and the verification performance is evaluated by comparing data of each user in the third session against that in the first session. Then through the proposed transformation method, the features extracted from each user during the enrollment stage (session 1) are transformed into a reference template which is then stored in the system. During the verification stage (session 3), the same feature extractor and transform are used to derive probe templates which are compared with the reference templates to compute the EER performance.

## V. Results

Let $F_e$ and $F_t$ denote the number of consecutive frames involved in a template during enrollment and verification, respectively. The verification performance is measured by the EER.

TABLE III

Within-Session Verification Performance EER (%). All Methods Are With the Same Frame Configuration ($F_e = 10$ and $F_t = 5$)

*MMIDB database*

| Features | Non-transformed | | Transformed domain | | |
|---|---|---|---|---|---|
| | EO | EC | EO | EC | Fusion |
| AR | 9.51 | 9.83 | 9.58 | 11.73 | 3.47 |
| PSD | 22.31 | 23.64 | 19.12 | 18.21 | 8.43 |
| FuzzyEn | 14.81 | 17.5 | 13.62 | 15.83 | 7.68 |
| AR+PSD+FuzzyEn | 8.23 | 8.67 | 3.53 | 4.87 | 2.41 |
| Graph | 1.89 | 6.37 | 1.32 | 2.44 | 1 |
| HandcraftedAll | 3.04 | 5.28 | 2.12 | 4.42 | 0.46 |
| DeepExtractor | 2.65 | 7.68 | 2.54 | 7.24 | 4.27 |

*BED database*

| Features | Non-transformed | | Transformed domain | | |
|---|---|---|---|---|---|
| | EO | EC | EO | EC | Fusion |
| AR | 27.53 | 23.75 | 26.9 | 18.86 | 22.9 |
| PSD | 10.59 | 13.6 | 11.6 | 12.29 | 9.47 |
| FuzzyEn | 17.81 | 16.14 | 21.15 | 17.76 | 20.97 |
| AR+PSD+FuzzyEn | 11.25 | 10.39 | 12.96 | 8.97 | 10.53 |
| Graph | 2.81 | 6.18 | 5.5 | 7.05 | 6.83 |
| HandcraftedAll | 5.34 | 5.89 | 5.1 | 6.33 | 5.67 |
| DeepExtractor | 0.86 | 0.42 | 0.86 | 0.94 | 0 |

*SEEDiv database*

| Features | Non-transformed | Transformed domain |
|---|---|---|
| | Neutral | Neutral |
| AR | 9.89 | 5.93 |
| PSD | 5.78 | 4.56 |
| FuzzyEn | 11.46 | 11.24 |
| AR+PSD+FuzzyEn | 3.38 | 3 |
| Graph | 0.08 | 0.34 |
| HandcraftedAll | 0.93 | 1.23 |
| DeepExtractor | 0 | 0 |

### A. Performance in the Lost Key Scenario

The lost key scenario is considered the worst case as the user loses his/her parameter key. This means that the attacker can take this advantage to penetrate the verification system. In order to simulate this scenario, we use the same parameter key to generate the permutation vector $\mathbf{p}$ in (3) and projection matrix $\mathbf{M}$ in (5) for all users in the transformation module. Table III presents within-session verification performance EER (%). All methods are with the same frame configuration ($Fe = 10$ and $Ft = 5$). Table IV summarizes the cross-session verification performance EER (%) under the same frame configuration ($F_e = 10$ and $F_t = 5$). The corresponding DET plots are summarized in Fig. 5 and Fig. 6. From the results, we can observe that the proposed cancellable template design (DeepExtractor+transformation) demonstrates a superior verification performance while protecting the raw EEG biometrics.

Comparing the results of elicitation fusion (embedded in the transformation) with those of single elicitation protocols, we can see an improvement in the verification performance for most of the cases, which shows the effectiveness of embedding the protocol fusion in the transformation for enhancing the verification performance. The results also show that, although

TABLE IV

CROSS-SESSION VERIFICATION PERFORMANCE EER (%). ALL METHODS ARE WITH THE SAME FRAME CONFIGURATION ($F_e = 10$ AND $F_t = 5$)

| Features | *BED database* | | | | | *SEEDiv database* | |
|---|---|---|---|---|---|---|---|
| | Non-transformed | | Transformed domain | | | Non-transformed | Transformed |
| | EO | EC | EO | EC | Fusion | Neutral | Neutral |
| AR | 38.26 | 37.4 | 35.25 | 39.77 | 38 | 29.7 | 35.52 |
| PSD | 39.61 | 27.07 | 43.55 | 36.21 | 36.1 | 38.09 | 34.38 |
| FuzzyEn | 37.91 | 35.35 | 44.87 | 41.6 | 44.65 | 38.59 | 41.37 |
| AR+PSD+FuzzyEn | 32.56 | 28.83 | 37.86 | 27.31 | 32.08 | 28.58 | 29.34 |
| Graph | 39.39 | 37 | 43.87 | 41.36 | 41.33 | 38.56 | 39.99 |
| HandcraftedAll | 32.94 | 31.92 | 38.55 | 34.28 | 36.52 | 28.13 | 30.24 |
| DeepExtractor | 8.59 | 11.91 | 6.78 | 11.57 | 3.75 | 1.29 | 0.14 |
| DeepExtractor-UserSpecific (mean) | 1.43 | 0.17 | 1.59 | 0.54 | 0.15 | 0.91 | 0.91 |
| *FNMR (%) when FMR=0* | | | | | | | |
| DeepExtractor | 98.61 | 99.32 | 98.37 | 98.62 | 97.5 | 31.1 | 0.14 |
| DeepExtractor-UserSpecific (mean) | 4.53 | 4.52 | 3.81 | 4.19 | 0.5 | 1.67 | 1.42 |

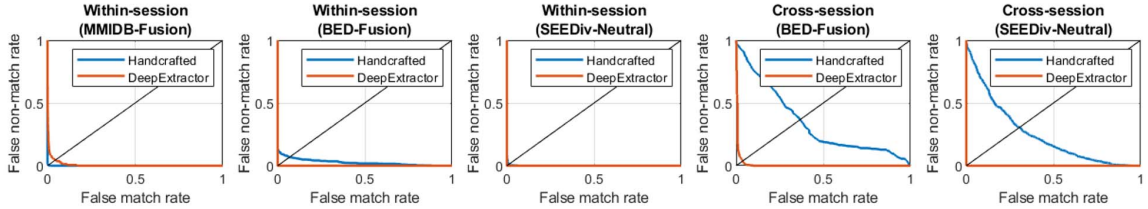


Fig. 5. DET curves of Handcrafted and DeepExtractor in the transformed domain for within-session and cross-session evaluation.

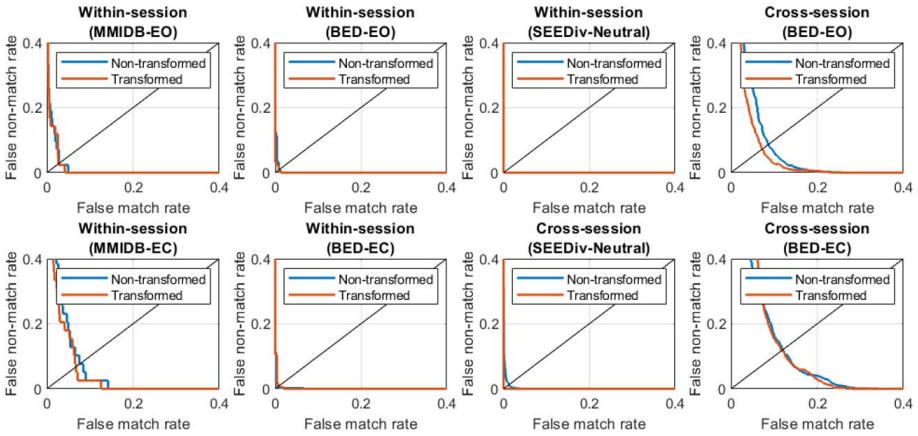


Fig. 6. DET curves of DeepExtractor in both non-transformed and transformed domains under EO and EC protocols for within-session and cross-session evaluation.

the proposed method uses resting state protocols and DeepExtractor features, the transformation itself is not confined to specific signal elicitation protocols or features. In high-security scenarios, it is often required to have a very low FMR. Therefore, we also report the FNMR of the proposed method when FMR=0 in Table IV. Comparing results in non-transformed and transformed domains, we can notice that the proposed transformation reduces FNMR under the same conditions. However, it is observed that the FNMR is relatively high on the BED database, although it is reasonable on the SEEDiv database. This issue can be easily addressed by some practical approaches, for example, using user-specific models and thresholds. A significant improvement is observed with user-specific models and comparison thresholds. When the system is operating at an extremely secure level (FMR=0), the DeepExtractor-UserSpecific provides FNMRs of 3.81%, 4.19%, and 0.5% for EO, EC, and state fusion on BED database, and 1.42% on SEEDiv database. The results indicate that the proposed method works well with high usability (low FNMR) when operating at high security level (FMR=0) in the classical user authentication scenario. The corresponding EER is also improved.

### *B. Decidability Analysis*

Biometric verification is essentially a decision task to discriminate the user from impostors. In this analysis, we adopt the decidability index $d'$ [43] to measure the discriminant capacity of the designed cancellable template. The $d'$ is defined as:

$$d' = (m_{intra} - m_{inter})/\sqrt{(s_{intra}^2 + s_{inter}^2)/2} \tag{8}$$

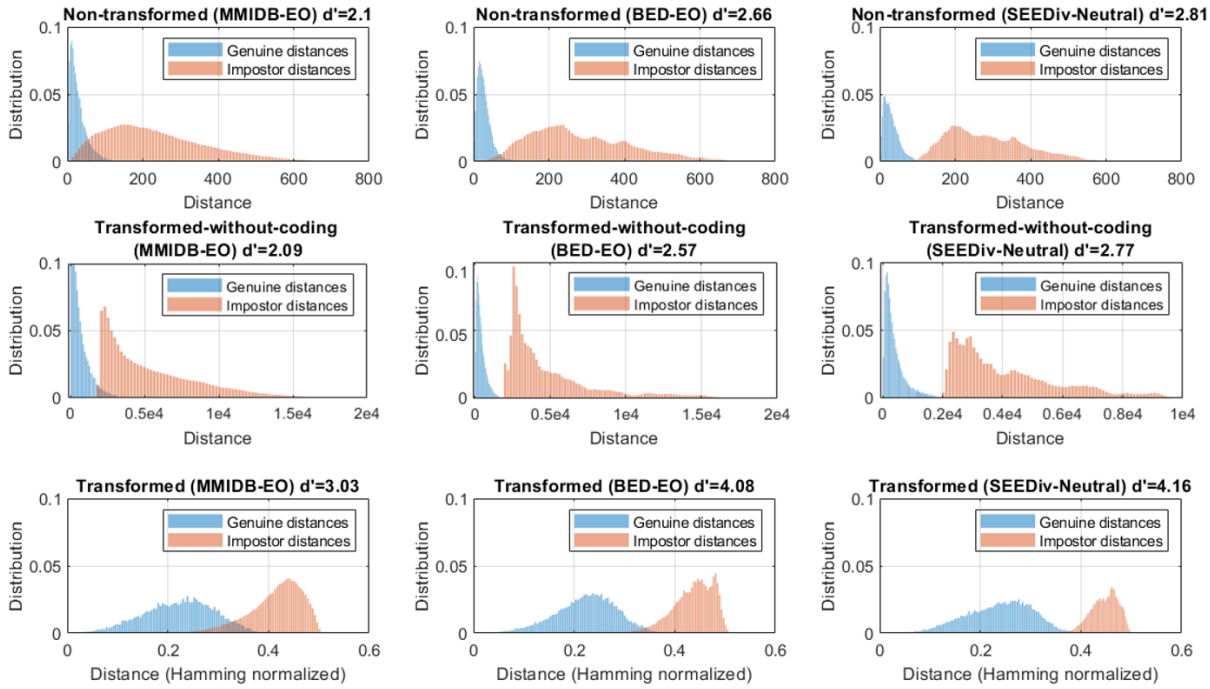


Fig. 7. Decidability analysis of DeepExtractor with and without transform. Distributions of the genuine distances and impostor distances, and the corresponding decidability index $d'$.

where $(m_{intra}, s_{intra})$ and $(m_{inter}, s_{inter})$ denote the mean and standard deviation of the comparison distances between user samples and the comparison distances between user samples and impostor samples, respectively. For each user, we generate a genuine distance distribution by comparing every possible pair of the user samples, and an impostor distance distribution by comparing each user sample with each sample of other subjects. Fig. 7 presents the distance distributions of the proposed method, DeepExtractor+transform (with and without coding), and its non-transformed version, DeepExtractor, under EO protocol on databases MMIDB and BED, and under Neutral protocol on database SEEDiv. The results of transformation without Gray code encoding are also presented to provide a better visual comparison of the corresponding real-valued distributions in the non-transformed domain. The observation is that the proposed transformation enhances system decidability, reducing the overlap between the genuine and impostor distance distributions. Without the coding component, it still provides the same level of decidability as the original system without transformation.

### C. Revocability and Diversity

The revocability and diversity criteria specify that templates generated from the same biometric features by different parameter keys should have no correlation. To evaluate the capacity of the proposed cancellable template design in terms of revocability and diversity, we follow the common practice in relevant studies [44] and calculate the pseudo-impostor distances. For each user, 50 additional transformed templates (i.e., the pseudo-impostor) are generated from the first feature template using different parameter keys. A pseudo-impostor distance distribution can then be obtained by comparing the original user templates with the pseudo-impostor templates of the same user. Fig.8 shows the pseudo-impostor distance distribution of the proposed method, along with the genuine and impostor distance distributions. The results show that the pseudo-impostor distance distribution has almost no overlap with the genuine distance distribution, and at the same time, having significant overlap with the impostor distance distribution. In other words, the system satisfies the revocability and diversity requirements.

### D. Unlinkability

For a cancellable biometric template design, the unlinkability property requires that the transformed templates originated from the same EEG data of the same subject are as different as those from different subjects [45]. To evaluate the unlinkability of the proposed method, we adopted two measures, i.e., the score/distance-wise linkability $D_{\leftrightarrow}(d)$ (a local measure) and system overall linkability $D_{\leftrightarrow}^{sys}$ (a global measure) [45], which are popular tools for unlinkability assessment in cancellable biometrics research. The calculation of $D_{\leftrightarrow}(d)$ and $D_{\leftrightarrow}^{sys}$ is based on the mated and non-mated sample score/distance distributions. The mated sample score/distance is obtained by comparing two templates generated from the same EEG data using different parameter keys. The non-mated sample score/distance is obtained by comparing two templates generated from the EEG of different subjects using different parameter keys. We followed the same procedure in a recent study [44] and generated six transformed databases using six different keys. The value range of $D_{\leftrightarrow}(d)$ and $D_{\leftrightarrow}^{sys}$ is [0, 1] with 0 indicating fully unlinkable and 1 indicating fully linkable. Fig. 9 presents the analysis results, where we tested the proposed method 'DeepExtractor+transformation' on the three databases. The proposed method provides high unlinkability, with very low global linkability indices around

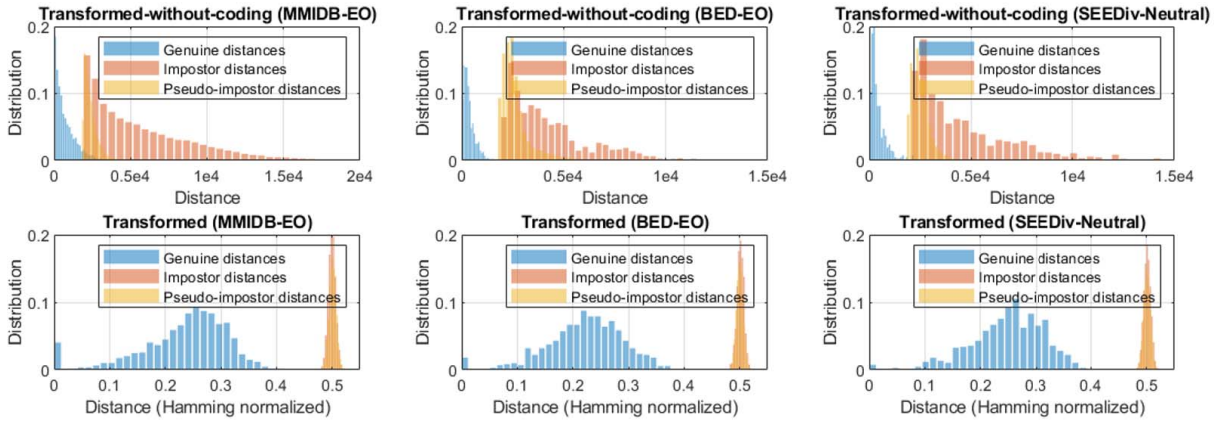


Fig. 8. Revocability and diversity analysis of DeepExtractor+transform. Distributions of the genuine, impostor, and pseudo-impostor distances.

0.01-0.05. With the complete transformation, the mated and non-mated distance distributions are highly overlapped. It is also observed that the Gray code encoding helps reduce the difference between the mated and non-mated distributions. This is because the scaling and normalization steps in Gray code encoding process enhances data consistency, and meanwhile, the coding procedure is a quantization process alleviating the impact of EEG uncertainties associated with the brain dynamics.

## VI. Security Analysis and Discussion

### A. Attacks via Record Multiplicity (ARM)

A cancellable biometric template design allows distinct transformed templates $\{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n\}$ to be generated from the same raw biometric template $\mathbf{x}$ by applying different transformation parameters $\{k_1, k_2, \cdots, k_n\}$. ARM refers to the attack that aims to retrieve the raw biometric template $\mathbf{x}$ by correlating multiple transformed templates $\{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n\}$, assuming that these transformed templates as well as information about the transformation method $F$ and corresponding parameters $\{k_1, k_2, \cdots, k_n\}$ are available [46].

The resistance of the proposed method to the ARM attack is guaranteed by the non-invertible transformation with three key points, i.e., random projection, random permutation and Hadamard product. First of all, the random projection procedure $\mathbf{y} = \mathbf{x} \cdot \mathbf{M}$ in (5) provides one-time-pad security so that each individual $\mathbf{y}_i$ cannot be reversed to obtain $\mathbf{x}$, as proved and demonstrated in previous studies [47]. However, the random projection itself is exposed to ARM because a unique solution can be determined by solving a well-defined system of linear equations $\{\mathbf{y}_i = F(\mathbf{x}, k_i)\}, i = 1, 2, \cdots, n$. To address this issue, the random permutation and Hadamard product operation is performed before the random projection. Note that the input of the random projection is actually the Hadamard product of $\mathbf{v}_2$ and permutation of $\mathbf{v}_1$ as in $\mathbf{c} = perm(\mathbf{v}_1, \mathbf{p}) \circ \mathbf{v}_2$, where $\mathbf{v}_1$ and $\mathbf{v}_2$ are two feature vectors; see (3) and (4). For different values of $k$ in (3), the random permutation and Hadamard product would produce different sets of variables for the random projection-based linear equations, thus a well-defined system of linear equations cannot be established. Since the projection matrix $\mathbf{M}$ is rank-deficient for every set of variables, it is insufficient to inverse the computation in (5).

Below is a representative example to show how the proposed method protects the system from the ARM attack. For demonstration purposes, we use low dimensional real vectors $\mathbf{v}_1 = [v_{11}, v_{12}, v_{13}, v_{14}]$ and $\mathbf{v}_2 = [v_{21}, v_{22}, v_{23}, v_{24}]$ to represent the real-valued feature vectors under protection. Suppose that the feature vectors are $\mathbf{v}_1 = [0.19, 0.54, 0.37, 0.84]$ and $\mathbf{v}_2 = [0.59, 0.18, 0.04, 0.92]$. Given two transformation parameters $k_1 = 1$ and $k_2 = 10$, we can produce two sets of permutations $\mathbf{p}_1 = [3, 4, 1, 2]$ and $\mathbf{p}_2 = [2, 3, 4, 1]$ and projection matrices $\mathbf{M}_1 = [0.15, 0.40; 0.09, 0.54; 0.19, 0.42; 0.35, 0.69]$ and $\mathbf{M}_2 = [0.50, 0.17; 0.22, 0.09; 0.20, 0.69; 0.76, 0.95]$. Applying the proposed transformation, we can get two transformed templates $\mathbf{t}_1 = [0010001110100010]$ and $\mathbf{t}_2 = [1010001000100011]$, which are the codes of real vectors $\mathbf{r}_1 = [0.22, 0.51]$ and $\mathbf{r}_2 = [0.31, 0.25]$, respectively. Now, suppose that an adversary gets $\mathbf{t}_1$, $\mathbf{t}_2$, $k_1$, $k_2$, knows the transformation function, and wants to retrieve $\mathbf{v}_1$ and $\mathbf{v}_2$. The first step taken by the adversary would be to decode the binary templates into the corresponding real values, which can be possible in the worse case, assuming that the adversary is able to collect massive amounts of encoded data and get the distribution of the values through statistical tools. Suppose that the estimated real vectors are $\hat{\mathbf{r}}_1 = [0.2, 0.5]$ and $\hat{\mathbf{r}}_2 = [0.3, 0.2]$, then the key step is to solve the following equations:

$$\begin{cases} 0.2 = 0.15v_{13}v_{21} + 0.09v_{14}v_{22} + 0.19v_{11}v_{23} + 0.35v_{12}v_{24} \\ 0.5 = 0.4v_{13}v_{21} + 0.54v_{14}v_{22} + 0.42v_{11}v_{23} + 0.69v_{12}v_{24} \\ 0.3 = 0.5v_{12}v_{21} + 0.22v_{13}v_{22} + 0.2v_{14}v_{23} + 0.76v_{11}v_{24} \\ 0.2 = 0.17v_{12}v_{21} + 0.09v_{13}v_{22} + 0.69v_{14}v_{23} + 0.95v_{11}v_{24} \end{cases}$$

However, the above is an ill-posed problem in that there is no unique solution and the solution is highly sensitive to changes in the estimated $\hat{\mathbf{r}}$. Using Matlab pseudo-inverse function, we get $\hat{\mathbf{v}}_1 = [0.39, 0, 0.85, 0.16]$ and $\hat{\mathbf{v}}_2 = [1, 0, 1, 0]$. However, the cosine similarity of the estimated value and ground truth value is $\frac{\mathbf{v}\hat{\mathbf{v}}^\intercal}{\|\mathbf{v}\|_2\|\hat{\mathbf{v}}\|_2} = 0.44$, which indicates that the obtained value is far from the true biometric data. Our analysis shows that if a transformed template stored in the database is compromised, it reveals no clue about the original biometric data. Even in the worst-case scenario where multiple sets of

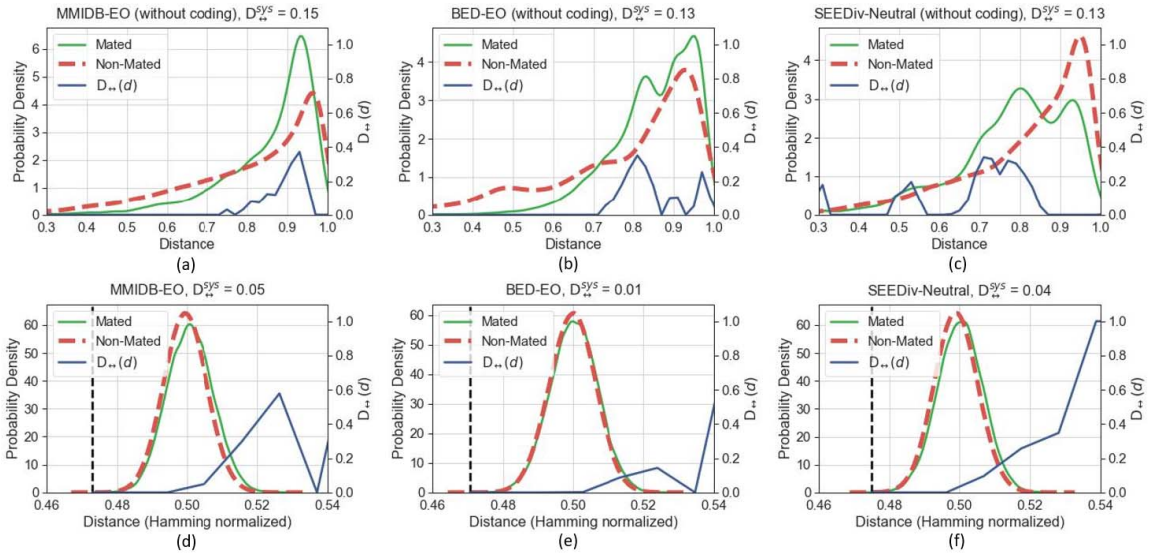


Fig. 9. Unlinkability analysis of DeepExtractor+Transformation. The mated and non-mated distance distributions and the linkability measures.

templates and the corresponding parameter keys are exposed, it would be highly unlikely to retrieve the true biometric data from infinite solutions.

### B. Pre-Image Attacks

The original definition of a pre-image attack on a cryptographic hash function refers to an attacker trying to determine an input that has a specific hash value. A cryptographic hash function $f(\cdot)$ should resist attacks on its pre-image, that is, given $y$, it is difficult to find $x$ such that $y = f(x)$. Such a definition does not fully apply in the context of transformation-based (i.e. non-cryptographic) cancellable design schemes. Strictly speaking, for a transformation-based cancellable design, it is possible to find an input $x$ given $y$ such that $y = f(x)$. However, it will be of little value if the solution $x$ is not the original biometric feature under protection and the compromised template is revoked. Considering the properties of transformation-based cancellable schemes, we therefore redefine the pre-image attack as follows:

*Given a transformed template* $\mathbf{y}$*, it is difficult to find a solution* $\mathbf{x}$ *such that* $\mathbf{y} = f(\mathbf{x}, K) = f(\mathbf{x}_0, K)$ *and* $\mathbf{x} = \mathbf{x}_0$*, where* $f(\cdot)$ *is the transformation function with parameter key* $K$*, and* $\mathbf{x}_0$ *is the original biometric feature.*

The proposed transformation is a many-to-one mapping function, and we have demonstrated in the ARM attack analysis that it would be difficult to find the real input in a systematical way. In the following hill-climbing attack analysis, Case I can also be considered as a pre-image attack. We will show that the solution found by the hill-climbing attack is far away from the real input, and therefore, the solution becomes insignificant once the compromised template is revoked.

### C. Hill-Climbing Attacks

This refers to an adversary exploiting the comparison scores/distances to generate synthetic biometric data that would allow a false acceptance [48]. In the context of cancellable biometrics, the hill-climbing attack can be launched in two ways, as illustrated in Fig. 10. Case I – the adversary submits and tries to obtain feature vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ as in the conventional non-cancellable context [48]. Case II – the adversary submits and tries to obtain template $\mathbf{t}$ stored in the system. Hill-climbing attacks are a threat to conventional non-cancellable biometric systems as the adversary is able to get a synthetic feature vector that is very close to the true feature vector and compromise the system with it. However, this is not necessarily true for cancellable biometric systems. In the following, we will demonstrate that cancellable biometric systems, especially those based on many-to-one mapping, are naturally resistant to hill-climbing attacks.

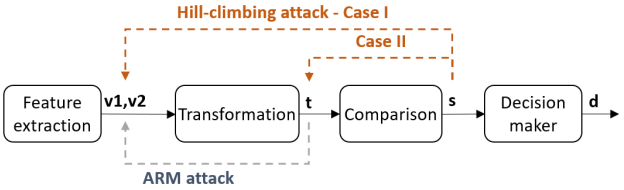


Fig. 10. Hill-climbing attacks on the system.

The Nelder-Mead algorithm was used to implement the hill-climbing attack. It is a downhill simplex method that is among the most well-known algorithms for derivative-free optimization [48]. The evaluation of the objective function $\mathcal{F}(\cdot)$ represents the difference between the input probe and the reference template. The process ends either when the minimum value of the objective function is equal to or less than the system threshold (here we set the threshold to the EER operating point) or when the maximum number of attempts is reached (here set to 20,000). The system's vulnerability to hill-climbing attacks is measured by the success rate (SR), defined as the percentage of users whose accounts are compromised within 20,000 attempts. The efficiency of the attack is measured by $N_{att}$, the average number of attempts required to successfully crack an account. We run the hill-climbing attack on two

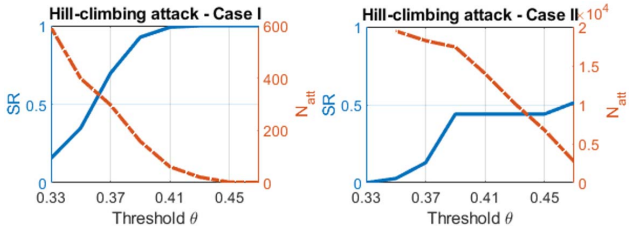


Fig. 11. The SR and $N_{att}$ of the hill-climbing attack on the system (Graph+transform).

methods, DeepExtractor+transform and Graph+transform. The SR of hill-climbing attack on DeepExtractor+transformation is 0, and result of the hill-climbing attack on Graph+transform is presented in Fig. 11. The hill-climbing attack failed to break the DeepExtractor+transformation because the dimension of DeepExtractor feature is large ($D$=1000) and the algorithm was not able to converge and to find a solution within the maximum number of attempts (20,000). We can see that it is possible to find a solution to temporarily break in user accounts with hill-climbing attacks [49], [50]. At the EER operating point, the SR of hill-climbing attacks is around 0.899 and 0.358 in Cases I and II, respectively. It is also worth noting that when adjusting the system operating threshold towards a lower FMR, the SR and efficiency of launching hill-climbing attacks decrease significantly.

Now, let us assume the adversary has successfully found a solution to pass the system through the hill-climbing attack. We will demonstrate that this solution will fail once the system changes the cancellable template. Let $\mathbf{t}_0$ denote the transformed template stored in the system before attack; and $\hat{\mathbf{v}}_1$, $\hat{\mathbf{v}}_2$, and $\hat{\mathbf{t}}_0$ denote the feature vectors and template obtained through the hill-climbing attack.

In Case I, the adversary obtained an estimated solution $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ that generates a template close enough to $\mathbf{t}_0$ to pass the system. To defend, the system will replace the compromised template $\mathbf{t}_0$ with a new one $\mathbf{t}_1$ using a new set of transformation parameters. Let $\hat{\mathbf{t}}_1$ denote the probe template generated from the estimated $\hat{\mathbf{v}}_1$ and $\hat{\mathbf{v}}_2$ with the same new transformation parameters. We now demonstrate that $\hat{\mathbf{t}}_1$ is not a valid solution for $\mathbf{t}_1$. In our experiment, there were 98/109 users (SR=0.899 at operating point of EER) whose original templates were successfully attacked by hill-climbing attacks. For each of those 98 users, we replaced the compromised template with a new one using a new key and tested whether the probe generated from the obtained solution using the new key is able to match the new template. To have a reliable analysis, we randomly generated 200 keys for each user, which yields a total number of 19600 ($98 \times 200$) tests. The results in Table V show that it is highly unlikely (0/19600) that the $\hat{\mathbf{t}}_1$ (generated from the adversary's obtained solution) can match the corresponding true template $\mathbf{t}_1$ (generated from the true feature vectors) to break in the account. A further examination shows that the hill-climbing solution $\{\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2\}$ is far away from the true feature vectors $\{\mathbf{v}_1, \mathbf{v}_2\}$, with a very low similarity score of $0.193 \pm 0.004$. This is because the proposed transformation is a many-to-one mapping and irreversible function, thus it is hard to hit the true solution through hill-climbing. Our results prove that even if a temporary solution is found, it is unlikely that this fake solution can pass the system after the compromised template is revoked.

TABLE V

RESULTS FOR ATTACKING THE SYSTEM USING HILL-CLIMBING SOLUTIONS AFTER THE COMPROMISED TEMPLATES ARE REVOKED (GRAPH+TRANSFORMATION)

| | Solution similarity | comparison scores (distance) | #Success/#Tests (SAR) |
|---|---|---|---|
| Case I | $0.193 \pm 0.004$ | $0.489 \pm 0.025$ | 0 / 19600 |
| Case II | $0.219 \pm 0.036$ | $0.781 \pm 0.036$ | 0 / 7800 |

In Case II, the adversary obtained $\hat{\mathbf{t}}_0$, which is an approximation of $\mathbf{t}_0$. To defend, the system can simply replace the compromised $\mathbf{t}_0$ with a new template $\mathbf{t}_1$ using a new set of transformation parameters. In our experiment, there were 39/109 users (SR=0.358 at the operating point of EER) whose original templates were successfully attacked by hill-climbing attacks. For each of those 39 users, we replaced the compromised template with a new one using a new key and tested whether the obtained solution can match the new template to break in the account again. To have a reliable analysis, we randomly generated 200 keys for each user, which yields a total number of 7800 ($39 \times 200$) tests. The results in Table V show that the hill-climbing solution $\hat{\mathbf{t}}_0$ is not similar to any of these new templates (with a low similarity score of $0.219 \pm 0.036$), and none of them can successfully match the new templates. Since $\hat{\mathbf{t}}_0$ neither reveals clues about the raw biometric data nor correlates with the new template, obtaining $\hat{\mathbf{t}}_0$ through hill-climbing attacks would be meaningless.

Hill climbing attacks have also been investigated in other cryptography-based privacy-preserving biometrics comparison schemes [3]. Our cancelable template design follows a different path: Instead of increasing the difficulty of finding the solutions leading to the match of the template, our many-to-one mapping scheme allows attackers to find many solutions where the genuine feature is hidden within. Little additional information is provided to the attackers to filter out the true solution.

### *D. Second Attack Rate - an Extension to the Classical Lost Key Scenario Analysis*

Spawned from hill-climbing attacks is a new concept, which we call *Second Attacks*. It is defined as an attempt to break into a system using pre-obtained solutions after the system has revoked the compromised templates. Accordingly, the *Second Attack Rate (SAR)* is the success rate of these second attack attempts. Here we examine three ways to obtain a valid solution to break in a user account.

- Mathematical solutions. Assume that the attacker acquires the user template $t$ and knows the transformation function $F$, then it is possible to find a solution $\hat{x}$ such that $t = F(x)$.
- Public data solutions. Assume that the attacker has a public database $X$, then for a system with a non-zero FMR, it is likely to find $\hat{x}$ that can pass the system by testing each data.
- Computational solutions via hill-climbing attacks. Assume that an attacker can submit $x$ to the system and

TABLE VI
SAR RESULTS UNDER THREE TYPES OF SOLUTIONS

| Graph+transformation | | | | |
|---|---|---|---|---|
| Solution | Solution similarity | comparison score (dist.) | #Tests | SAR |
| Mathematical | 0.167 ± 0.129 | 0.497 ± 0.025 | 109×200 | 0 |
| Public data | 0.033 ± 0.005 | 0.499 ± 0.025 | 109×200 | 0 |
| Computational | 0.193 ± 0.004 | 0.489 ± 0.025 | 98×200 | 0 |
| DeepExtractor+transformation | | | | |
| Solution | Solution similarity | comparison score (dist.) | #Tests | SAR |
| Mathematical | 0.154 ± 0.052 | 0.499 ± 0.034 | 109×200 | 0 |
| Public data | 0.047 ± 0.008 | 0.496 ± 0.031 | 109×200 | 0 |
| Computational | – | – | 0 | 0 |

get the corresponding comparison score/distance, then it is possible to find $\hat{x}$ that can pass the system.

Table VI summarizes the performance of the proposed cancellable biometrics design in terms of SAR using the aforementioned three types of solutions. An interesting finding is that the computational solution obtained through the hill-climbing attack is not better than the public data solution or the mathematical solution. This finding validates our hypothesis from an experimental point of view that a good cancellable biometrics design based on many-to-one mapping is inherently resistant to hill-climbing attacks.

### *E. Entropy and Brute Force Attacks*

There are two types of formal proof techniques in data security. One refers to the provable security in cryptography which depends on the reduction into a problem's complexity. The problem in our paper does not belong to this category. Another formal proof is related to the information-theoretic leakage, which is virtually about the conditional probability related to the information leakage. As our feature variables are real numbers in an under-defined system of nonlinear equations, it has theoretically infinite number of solutions, leading to the zero conditional probability of finding the genuine feature. Therefore, we provided a close to the worst-case information leakage estimation based on the search space of the encoded input feature (binary), which is the brute force attack-based entropy.

A brute force attack attempts to guess the elements of the original EEG feature vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ through exhaustive search. We analyze the search space to show the likelihood of finding the secret successfully from the search space. In the proposed system, both $\mathbf{v}_1$ and $\mathbf{v}_2$ have a length of $D \times n$ bits, where we have $D = 1000$ and $n = 8$ in our experimental setup. Therefore, the number of trials needed to attack the system would be $2^{8000}$, which is computationally expensive. In actual deployment, the settings of $D$ and $n$ can be adjusted according to the requirements. For example, a larger $D$ can further enhance the security level, however, it is worth noting that a larger $D$ also implies a higher computational cost or less efficiency in data collection. There is a trade-off between performance, security and system efficiency.

### *F. Pitfalls in the Evaluation Procedure of Supervised Learning-Based Verification Systems*

Many papers on EEG biometrics treat verification as a pure classification problem without considering the differences between classification and biometric verification. In the following analysis, we take the LDA and SVM, two popular classifiers widely used in EEG biometrics, as an example to show that user verification is not merely a classification problem, and the standard evaluation procedure for classification systems based on supervised learning (SL) models are not fully applicable to the evaluation of user verification systems. Assuming that the database has $N$ subjects, the standard evaluation procedure for SL-based systems is illustrated in Algorithm 3.

**Algorithm 3** Standard Evaluation for Classification

```
predictions = []
for n = 1 to N do
    re-label data of subject n as 1 (user)
    re-label data of other subjects as 0 (intruder)
    train-test-split (80%, 20%) (or 5-fold cross-validation)
    SL model ← SL model training on train set
    predictions ← SL model testing on test set
end
accuracy ← confusion matrix (predictions, ground truth)
```

In the context of user verification, the evaluation process described above has two major issues. 1) Training a binary SL model requires samples from both classes, and the way it splits train/test sets has provided the model with all intruders' data during the training phase. This is incorrect because no data from any test intruder should be seen by the verification system until the system is tested. The correct procedure should separate the intruder set from the user set, as applied in [11]. Adopting the basic idea of separating user and intruder sets, we suggest a more reliable evaluation procedure for EEG-based verification system (i.e., Algorithm 4). 2) Verification systems based on SL classification models (e.g., LDA and SVM) do not have a system operating threshold. An individual classification model is trained for each user, and the overall system performance is embodied as a pair of FMR and FNMR. The ROC curve and EER are reported in some studies, but such ROC or EER is a measure to examine the output of the classifier, not the ROC or EER of the verification system.

Table VII reports the results of PSD features with LDA and SVM classifiers under the two evaluation procedures in different settings. The first observation is that the standard classification evaluation procedure gives false high performance, especially the FMR, because it erroneously feeds intruder data to the model during the training phase. In addition, the performance of classification-based verification systems relies on a good training set. With less training data (a smaller number of users in the system dropped from 80 to 30), the performance would degrade. Another concern is that the 80%-20% data split, a very common setup in classification tasks, may be too lax for testing an verification system. For example, given a database where each subject has 60 samples, the 80-20 split means 48 samples are used for registration (which would take a while for data collection) and only 12 samples are used for positive testing. Our results show that reducing the split ratio (from 80% to 33%) also degrades

**Algorithm 4** Evaluation Procedure for Supervised Learning-Based Verification

```
Split database into user set U and intruder set I
IntruderTests ← intruder set I
predictions = []
for u = 1 to U do
    re-label data of subject u as 1 (user)
    re-label data of other subjects as 0 (non-user)
    train-test-split (80%, 20%) (or 5-fold cross-validation)
    SL model ← SL model training on train set
    UserTests ← testing data of label 1
    predictions ← SL model testing on UserTests
    predictions ← SL model testing on IntruderTests
end
accuracy ← confusion matrix (predictions, ground truth)
```

TABLE VII
RESULTS (%) UNDER DIFFERENT EVALUATION PROCEDURES FOR CLASSIFICATION-BASED SYSTEMS

| Method | Evaluation | Accuracy | FMR | FNMR | EER (classi.) |
|---|---|---|---|---|---|
| PSD+SVM | Classi. (80%) | 99.56 | 0.37 | 7.80 | 1.68 |
| PSD+SVM | Authen. (80%, 80 users) | 96.73 | 3.23 | 8.75 | 4.06 |
| PSD+SVM | Authen. (33%, 80 users) | 96.34 | 3.44 | 12.94 | 4.7 |
| PSD+SVM | Authen. (33%, 30 users) | 95.37 | 4.58 | 10.33 | 5.89 |
| PSD+LDA | Classi. (80%) | 97.16 | 2.76 | 11.16 | 7.80 |
| PSD+LDA | Authen. (80%, 80 users) | 91.46 | 8.52 | 11.46 | 9.58 |
| PSD+LDA | Authen. (33%, 80 users) | 89.74 | 10.09 | 17.38 | 12.91 |
| PSD+LDA | Authen. (33%, 30 users) | 77.51 | 22.52 | 18.50 | 21.65 |

The number in parentheses indicates the split ratio for training set.

performance. We hope to use this demonstration to rectify the misconception of many existing studies evaluating classifier-based verification systems.

## VII. CONCLUSION

In this study, we proposed a cancellable biometrics scheme for privacy-preserving EEG-based verification systems. To be specific, an innovative non-invertible transformation was designed to generate cancellable templates from EEG features extracted by a deep learning model while taking advantage of signal elicitation protocol fusion to enhance biometric performance. The results demonstrated that the proposed method provides a superior verification performance than the state-of-the-art, prevents the leakage of the sensitive information contained in the EEG data, and is secure against the ARM attack, pre-image attack, hill-climbing attack, and brute force attack. In particular, we examined two ways to perform hill-climbing attacks, and demonstrated that the solution found through hill-climbing attacks would fail once the system revokes the compromised template. In other words, cancellable biometric systems, especially those based on many-to-one mapping, are naturally resilient against hill-climbing attacks. We also introduced the concept of second attacks for cancellable biometric systems. Finally, we discussed the evaluation procedure of supervised learning-based verification systems and the pitfalls involved. Our future work will further this line of research to explore the possibility of integrating cryptographic schemes into verification systems [49].

## REFERENCES


[1] A. Hadid, N. Evans, S. Marcel, and J. Fierrez, "Biometrics systems under spoofing attack: An evaluation methodology and lessons learned," *IEEE Signal Process. Mag.*, vol. 32, no. 5, pp. 20–30, Sep. 2015.

[2] E. Marasco and A. Ross, "A survey on antispoofing schemes for fingerprint recognition systems," *ACM Comput. Surveys*, vol. 47, no. 2, pp. 1–36, Jan. 2015.

[3] N. D. Sarier, "Multimodal biometric authentication for mobile edge computing," *Inf. Sci.*, vol. 573, pp. 82–99, Sep. 2021.

[4] Q. Gui, M. V. Ruiz-Blondet, S. Laszlo, and Z. Jin, "A survey on brain biometrics," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 1–38, 2019.

[5] M. Wang, H. El-Fiqi, J. Hu, and H. A. Abbass, "Convolutional neural networks using dynamic functional connectivity for EEG-based person identification in diverse human states," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 12, pp. 3259–3272, Dec. 2019.

[6] D. La Rocca, P. Campisi, and G. Scarano, "EEG biometrics for individual recognition in resting state with closed eyes," in *Proc. Int. Conf. Biometrics Special Interest Group (BIOSIG)*, Sep. 2012, pp. 1–12.

[7] D. La Rocca *et al.*, "Human brain distinctiveness based on EEG spectral coherence connectivity," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 9, pp. 2406–2412, Sep. 2014.

[8] M. V. Ruiz-blondet, Z. Jin, and S. Laszlo, "CEREBRE: A novel method for very high accuracy event-related potential biometric identification," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 7, pp. 1618–1629, Jul. 2016.

[9] E. Maiorana, D. La Rocca, and P. Campisi, "On the permanence of EEG signals for biometric recognition," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 1, pp. 163–175, Jan. 2016.

[10] M. Wang, J. Hu, and H. A. Abbass, "BrainPrint: EEG biometric identification based on analyzing brain connectivity graphs," *Pattern Recognit.*, vol. 105, Sep. 2020, Art. no. 107381.

[11] A. Riera, A. Soria-Frisch, M. Caparrini, C. Grau, and G. Ruffini, "Unobtrusive biometric system based on electroencephalogram analysis," *EURASIP J. Adv. Signal Process.*, vol. 2008, no. 1, pp. 1–8, Dec. 2007.

[12] S. Yang, F. Deravi, and S. Hoque, "Task sensitivity in EEG biometric recognition," *Pattern Anal. Appl.*, vol. 21, no. 1, pp. 105–117, 2016.

[13] E. Debie, N. Moustafa, and A. Vasilakos, "Session invariant EEG signatures using elicitation protocol fusion and convolutional neural network," *IEEE Trans. Dependable Secure Comput.*, vol. 19, no. 4, pp. 2488–2500, Jul. 2022.

[14] O. Landau, R. Puzis, and N. Nissim, "Mind your mind: EEG-based brain-computer interfaces and their security in cyber space," *ACM Comput. Surveys*, vol. 53, no. 1, pp. 1–38, Jan. 2021.

[15] Y. Höller and A. Uhl, "Do EEG-biometric templates threaten user privacy?" in *Proc. 6th ACM Workshop Inf. Hiding Multimedia Secur.*, Jun. 2018, pp. 31–42.

[16] P. Campisi and D. La Rocca, "Brain waves for automatic biometric-based user recognition," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 5, pp. 782–800, May 2014.

[17] T. Nakamura, V. Goverdovsky, and D. P. Mandic, "In-ear EEG biometrics for feasible and readily collectable real-world person authentication," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 3, pp. 648–661, Mar. 2018.

[18] Z. Mu, J. Hu, and J. Min, "EEG-based person authentication using a fuzzy entropy-related approach with two electrodes," *Entropy*, vol. 18, no. 12, p. 432, 2016.

[19] M. Fraschini, S. M. Pani, L. Didaci, and G. L. Marcialis, "Robustness of functional connectivity metrics for EEG-based personal identification over task-induced intra-class and inter-class variations," *Pattern Recognit. Lett.*, vol. 125, pp. 49–54, Jul. 2019.

[20] B.-K. Min, H.-I. Suk, M.-H. Ahn, M.-H. Lee, and S.-W. Lee, "Individual identification using cognitive electroencephalographic neurodynamics," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 9, pp. 2159–2167, Sep. 2017.

[21] M. Fraschini, A. Hillebrand, M. Demuru, L. Didaci, and G. L. Marcialis, "An EEG-based biometric system using eigenvector centrality in resting state brain networks," *IEEE Signal Process. Lett.*, vol. 22, no. 6, pp. 666–670, Jun. 2015.

[22] E. Maiorana, "Deep learning for EEG-based biometric recognition," *Neurocomputing*, vol. 410, pp. 374–386, Oct. 2020.

[23] E. Maiorana, "Learning deep features for task-independent EEG-based biometric verification," *Pattern Recognit. Lett.*, vol. 143, pp. 122–129, Mar. 2021.

[24] E. Maiorana, "EEG-based biometric verification using Siamese CNNs," in *Proc. Int. Conf. Image Anal. Process.* Cham, Switzerland: Springer, 2019, pp. 3–11.
[25] C. He, X. Lv, and Z. J. Wang, "Hashing the mAR coefficients from EEG data for person authentication," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 1445–1448.
[26] G. Bajwa and R. Dantu, "Neurokey: Towards a new paradigm of cancelable biometrics-based key generation using electroencephalograms," *Comput. Secur.*, vol. 62, pp. 95–113, Sep. 2016.
[27] S. Yang and F. Deravi, "On the usability of electroencephalographic signals for biometric recognition: A survey," *IEEE Trans. Human-Mach. Syst.*, vol. 47, no. 6, pp. 958–969, Dec. 2017.
[28] R. Damaševicius, R. Maskeliunas, E. Kazanavicius, and M. Wozniak, "Combining cryptography with EEG biometrics," *Comput. Intell. Neurosci.*, vol. 2018, pp. 1–11, May 2018.
[29] E. Maiorana, D. L. Rocca, and P. Campisi, "Cognitive biometric cryptosystems a case study on EEG," in *Proc. Int. Conf. Syst., Signals Image Process. (IWSSIP)*, Sep. 2015, pp. 125–128.
[30] K. Simoens, J. Bringer, H. Chabanne, and S. Seys, "A framework for analyzing template security and privacy in biometric authentication systems," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 833–841, Apr. 2012.
[31] A. Manisha and N. Kumar, "Cancelable biometrics: A comprehensive survey," *Artif. Intell. Rev.*, vol. 53, pp. 3403–3446, Oct. 2019.
[32] H. Zhang, W. Bian, B. Jie, D. Xu, and J. Zhao, "A complete user authentication and key agreement scheme using cancelable biometrics and PUF in multi-server environment," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 5413–5428, 2021.
[33] N. D. Sarier, "Practical multi-factor biometric remote authentication," in *Proc. 4th IEEE Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Sep. 2010, pp. 1–6.
[34] A. Mansfield, *Information Technology—Biometric Performance Testing and Reporting—Part 1: Principles and Framework*, Standard ISO/IEC 19795-1, 2006.
[35] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet," *Circulation*, vol. 101, no. 23, Jun. 2000.
[36] P. Arnau-Gonzalez, S. Katsigiannis, M. Arevalillo-Herraez, and N. Ramzan, "BED: A new data set for EEG-based biometrics," *IEEE Internet Things J.*, vol. 8, no. 15, pp. 12219–12230, Aug. 2021.
[37] W.-L. Zheng, W. Liu, Y. Lu, B.-L. Lu, and A. Cichocki, "EmotionMeter: A multimodal framework for recognizing human emotions," *IEEE Trans. Cybern.*, vol. 49, no. 3, pp. 1110–1122, Mar. 2019.
[38] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain-computer interface (BCI) system," *IEEE Trans. Biomed. Eng.*, vol. 51, no. 6, pp. 1034–1043, Jun. 2004.
[39] L. J. Gabard-Durnam, A. S. M. Leal, C. L. Wilkinson, and A. R. Levin, "The Harvard automated processing pipeline for electroencephalography (HAPPE): Standardized processing software for developmental and high-artifact data," *Frontiers Neurosci.*, vol. 12, p. 97, Feb. 2018.
[40] P. Campisi *et al.*, "Brain waves based user recognition using the 'eyes closed resting conditions' protocol," in *Proc. IEEE Int. Workshop Inf. Forensics Secur.*, Nov. 2011, pp. 1–6.
[41] W. Chen, Z. Wang, H. Xie, and W. Yu, "Characterization of surface EMG signal based on fuzzy entropy," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 15, no. 2, pp. 266–272, Jun. 2007.
[42] A. B. J. Teoh, Y. W. Kuan, and S. Lee, "Cancellable biometrics and annotations on BioHash," *Pattern Recognit.*, vol. 41, no. 6, pp. 2034–2044, 2008.
[43] G. O. Williams, "The use of d' as a 'decidability' index," in *Proc. 30th Annu. Int. Carnahan Conf. Secur. Technol.*, 1996, pp. 65–71.
[44] Q. N. Tran and J. Hu, "A multi-filter fingerprint matching framework for cancelable template design," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 2926–2940, 2021.
[45] M. Gomez-Barrero, J. Galbally, C. Rathgeb, and C. Busch, "General framework to evaluate unlinkability in biometric template protection systems," *IEEE Trans. Inf. Forensics Security*, vol. 13, no. 6, pp. 1406–1420, Jun. 2018.
[46] C. Li and J. Hu, "Attacks via record multiplicity on cancelable biometrics templates," *Concurrency Comput., Pract. Exper.*, vol. 26, no. 8, pp. 1593–1605, Jun. 2014.
[47] S. Wang and J. Hu, "Alignment-free cancelable fingerprint template design: A densely infinite-to-one mapping (DITOM) approach," *Pattern Recognit.*, vol. 45, no. 12, pp. 4129–4137, 2012.
[48] E. Maiorana, G. E. Hine, D. L. Rocca, and P. Campisi, "On the vulnerability of an EEG-based biometric system to hill-climbing attacks algorithms comparison and possible countermeasures," in *Proc. IEEE 6th Int. Conf. Biometrics, Theory, Appl. Syst. (BTAS)*, Sep. 2013, pp. 1–6.
[49] D. Sarıer, "Biometric cryptosystems: Authentication, encryption and signature for biometric identities," Ph.D. dissertation, Fac. Math. Natural Sci., Rheinische Friedrich-Wilhelms-Univ. Bonn, Bonn, Germany, 2013.
[50] E. Maiorana, G. E. Hine, and P. Campisi, "Hill-climbing attacks on multibiometrics recognition systems," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 900–915, May 2015.

**Min Wang** (Member, IEEE) received the Ph.D. degree in computer science from the University of New South Wales, Canberra, Australia, in 2020. She is currently a Post-Doctoral Research Fellow with the School of Engineering and Information Technology, University of New South Wales. Her research interests include biometrics, pattern recognition, machine learning, brain–computer interface, and bio-cryptography.

**Song Wang** received the B.Eng. degree in electrical engineering from Xi'an Jiaotong University, China, in 1991, and the Ph.D. degree in control theory from the University of Melbourne, Australia, in 2001. She is currently a Senior Lecturer with the Department of Engineering, La Trobe University, Australia. She has published many high-quality articles in highly ranked journals, such as *Pattern Recognition*, IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS, and IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY. Her main research interests include biometric security.

**Jiankun Hu** (Senior Member, IEEE) is currently a Full Professor in cyber security with the School of Engineering and Information Technology, University of New South Wales, Canberra, Australia. His main research interests are in the field of cyber security, including biometrics security, where he has publications at top venues, including the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. He has served on the Editorial Boards of up to seven international journals, including the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY.