Digital Signal Processing ••• (••••) ••••••

Contents lists available at ScienceDirect

Digital Signal Processing

www.elsevier.com/locate/dsp

Recent advances in biometrics-based user authentication for wearable devices: A contemporary survey

Shuqi Liu [a,b,*], Wei Shao [a,*], Tan Li [a,*], Weitao Xu [a,*], Linqi Song [a,b,*]

[a] Department of Computer Science, City University of Hong Kong, Hong Kong
[b] City University of Hong Kong Shenzhen Research Institute, China

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Wearable devices
User authentication
Signal processing
Machine learning

ABSTRACT

In recent years, wearable technology is interwoven with our everyday lives because of its commoditization and comfort. Security and privacy become a big concern as many user-sensitive data have been stored in such devices, such as personal emails and bank accounts. Traditional user authentication techniques like PIN entry are unfriendly and vulnerable to shoulder surfing attacks. To address these problems, a number of new authentication methods have been proposed. In this survey, we review and categorize recent advances in user authentication for wearable devices. We classify existing studies into physiological biometrics based and behavioral biometrics based methods. For each category, we review how signal processing techniques have been used to extract features in various wearable devices. Leveraging these extracted features, specifically designed classification methods can be used to realize user authentication. Finally, we review evaluation metrics for user authentication in wearable devices. Overall, in this survey, we systematically study assorted state-of-the-art user authentication methods for wearable devices, aiming to provide guidance and directions for future research in this area.

1. Introduction

Wearable devices are becoming increasingly popular in recent years. Commercial wearable products such as smartwatches and smart wristbands are penetrating everyone's daily life. Through these easy-to-wear devices, people can easily access social media and enjoy e-pay services. According to a recent report [1], the wearable market is expected to reach $57,653 million by 2022, which is three times larger than that in 2016 ($19,633 million). Although wearable applications show significant potential in terms of features - such as accuracy, usability, and comfort - new protection, authentication and privacy concerns are also raised. This is because much of the appeal of wearable device services is focused on the sensitive and personal nature of the data they collect, store, manipulate and transmit. Attackers will steal confidential data from the computer to snoop, resulting in dreadful consequences. Therefore, a robust, rapid and friendly user authentication is of great significance.

For those reasons, user authentication in wearables has attracted much attention in the literature [2–4] and deserves a thorough survey. In this survey, we review recent advances in user authentication for wearable devices, especially how to combine recent signal processing and machine learning approaches with wearable applications. There are two categories of authentication systems in the computer security field that can enable the link to a person and his/her identity: user identification and user verification. User identification refers to the process of establishing the identity of a user, while user verification means the system accepts or declines his/her claim when the user claims an identity. In any security system, the identification or verification can be done with the following three factors [3,5]: a knowledge factor ("something the user knows") such as password and Personal Identification Number (PIN) code, a possession factor ("something the user owns") such as card and key, and an inherence factor ("something the user is") such as face and fingerprint. While the former two factors are widely used in our life, there are several disadvantages. For example, the password can be forgotten and the card can be stolen or duplicated.

This survey will focus on inherence-factor-based user authentication systems, namely, something the user is. These inherence factors obey the rule of uniqueness, so each individual has their own signal pattern that can be distinguished from other users. "Something the user is" represents the unique biometric signals produced by a human. The uniqueness could either be determined by human DNA like fingerprint or determined by an individual's

* Corresponding authors.
E-mail addresses: shuqiliu4-c@my.cityu.edu.hk (S. Liu), weishao4-c@my.cityu.edu.hk (W. Shao), tanli6-c@my.cityu.edu.hk (T. Li), weitaoxu@cityu.edu.hk (W. Xu), linqi.song@cityu.edu.hk (L. Song).

https://doi.org/10.1016/j.dsp.2021.103120
1051-2004/

Please cite this article as: S. Liu, W. Shao, T. Li et al., Recent advances in biometrics-based user authentication for wearable devices: A contemporary survey, Digital Signal Processing, https://doi.org/10.1016/j.dsp.2021.103120

long-standing habits that are hard to be imitated. Therefore, the research in this area can be further divided into two subfields: physiological biometrics and behavioral biometrics. Physiological biometrics refers to the internal signal produced by the human body that varies among different people. In comparison, behavioral biometrics refers to a period of continuous human behavior, which relates to an individual's unique behavior pattern determined by long-term habits. Physiological biometrics-based authentication systems could identify different users by their physiological signals. Fingerprint and FaceID are commonly used in smartphones to unlock the system, but these signals are vulnerable since they are easy to be replicated [6–8]. The fingerprint can be collected from the object surface that the user has touched, and the replicated fingerprint model could unlock the system as well [9,10]. The 2D-FaceID authentication system can be attacked by user face images. 3D-FaceID authentication system is hard to be attacked but has higher price [11,12].

There are several survey works focusing on biometric user authentication that introduce which biometrics could be used in user authentication, how these biometrics are utilized and their performance. In detail, Kataria et al. [13] list a series of biometrics for user authentication in their work. Yusuf et al. [14] give a classification for some biometrics and thoroughly describe fingerprint biometric authentication and password biometric authentication. Mahfouz et al. [15] make a comprehensive survey on biometric user authentication on the smartphone. Zhang et al. [16] introduce the advances of biometric user authentication from a perspective of secure and privacy and provide detailed performance comparison between biometrics based on different criteria. Also, some works [17,18] only focus on a kind of biometric for user authentication, like keystroke. These works are more or less insufficient in the comprehensiveness of biometrics, content organization, the depth of analysis, and the comparison of the effects of characteristic methods.

Different from these previous works, our work concentrates on the user authentication for wearable devices. Because wearable devices can collect physiological signals and some signals are unique for each user, these unique physiological signals could be utilized in user authentication. In this case, some unique physiological signals-based user authentication methods are grouped. In detail, the difference between our work and previous related works could be described as the following:

1) Features: In this survey, we put an emphasis on the biometrics-based user authentication methods that are hard to be attacked. Besides the including of regular biometrics, like fingerprints and faces, our work pays more attention to Physiological signals inside the human body, like a heartbeat, breathing, and muscle signals. At the same time, we take the human body's behavior and movement trajectory as a unique physiological signal, including gait, signature, keystroke, and mouse dynamics. However, other works only contain the former part.
2) Structure: In the structure of the content, a five-level logical structure is used to summarize these surveys instead of simply listing the methods for each task. In this logical structure, we reformulate the user authentication as a classification task and cut this task to several sequential modules. These modules usually have a higher level than parts of other works, which could help us understand the relationships among mentioned works in our paper easily.

Instead of just listing or grouping these biometrics, we indicate the usage situation of these metrics and related methods and how they are combined together. Furthermore, we present the best performance of each kind of method on different evaluation metrics and datasets.

Studies on biometrics-based user authentication systems have gone through several stages, from using fiducial features detected from raw signals to machine learning methods using extracted features and deep learning methods without feature engineering. We depict Fig. 1 to illustrate the leading architecture and commonly used techniques of a biometrics-based authentication system. The authentication pipeline mainly contains five parts: data acquisition, data preprocessing, feature extraction, classification, and evaluation.

1) Data acquisition module aims to obtain raw biometric signals, either obtained from wearable devices or directly acquired from available public datasets. Wearable devices are typically placed on the human body to capture signals through various kinds of sensors continuously. Popular wearable devices have been verified to be promising approaches for user authentication, such as brain-computer interfaces to capture brain waves, tight clothes embedded with sensors or smart watches to capture heart or breath waves, and accelerometers to capture gait or keystroke dynamics. Moreover, researchers could also utilize open-access datasets to focus on model design and verification.
2) Data pre-processing module is applied to filter high-frequency noise and segment signals into periodic cycles. Wearable device-based biometric signals are generally time-series waveform signals, so we mainly introduce the commonly used waveform signal pre-processing methods. The first step in data pre-processing is signal filtering, which filters the high-frequency noise caused by unstable measurement circumstances in the signal capturing stage. Then, filtered signals are segmented into cycles with the same time interval. Since Biometric or behavioral signals are roughly periodic, we usually regard a signal circle as a sample. Next, the key points and key intervals are detected in the signal as fiducial features.
3) Feature extraction module aims to extract useful features for classification. The features could be divided into three categories, statistical features, transform domain features, and distance-based features. Statistical features like variance and skewness are calculated through commonly used statistical methods to analyze the data distribution. Shannon energy and spectral slope mainly focus on the signal power information to reflect signal attributes. Other features like Zero Cross Rate (ZCR), Short Time Energy (STE) are considered to be key features to identify a particular voice in the speech signal. Signal processing features focus on the frequency information of time-domain waveform, obtained through various transformation methods in the signal processing domain. Fast Fourier Transform (FFT) and Discrete Cosine Transform (DCT) are the basic transformation techniques to transfer the time-domain signal into a frequency domain signal with individual spectral components. Discrete Wavelet Transform (DWT) decomposes a signal into a set of the orthogonal waveform in both time and frequency domain. Mel-frequency Cepstral Coefficient (MFCC) is considered the most evident and mainstream feature extraction technique constructed using the speech frequency information. Power Spectral Density (PSD) measures the relationship between signal power and frequency. Fiducial features occasionally are insufficient to form a signal template. Therefore distance-based features are obtained by calculating the distance between the pairs of raw training data. The commonly used distance measurement methods include discrete-time warping, Manhattan distance, Euclidean distance.
4) Classification module utilizes template matching method, machine learning algorithm, or deep neural network to classify the given biometrics signals and returns the authentication re-

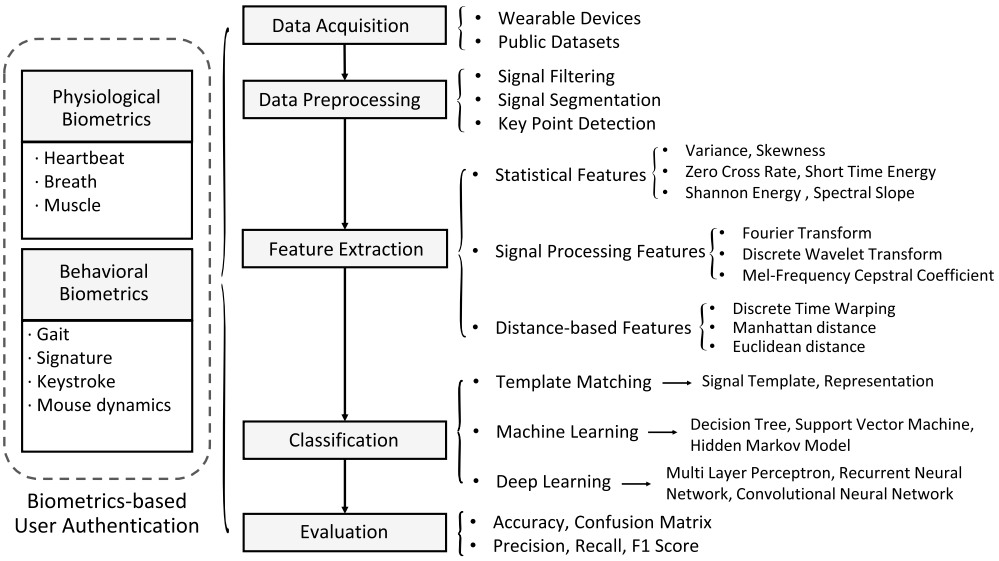

Fig. 1. The architecture of a biometrics-based user authentication method.

sult. The template matching method compares the distance between the signal representation and the template. The user is authenticated when the calculated distance is lower than the threshold. Machine learning methods utilize the extracted features as input and learn a parameterized classification model to obtain the authentication result. Because machine learning methods require heavy task-oriented feature engineering work, further research starts to extracted features by utilizing neural networks that directly transfer a signal sample into a feature vector. Besides, the recurrent neural network is adopted to deal with sequential data.

5) Evaluation module is an essential part of evaluating the proposed classification models through various metrics such as accuracy, confusion matrix, precision, recall, F1 score, etc. Using only one metric is insufficient to judge the model. Thus several performance evaluation results should be analyzed simultaneously.

The rest part of the paper is organized as follows. Section 2 summarizes the recent developments of physiological biometrics and behavioral biometrics, involving the signal description, signal acquisition methods, signal fiducial features, and current authentication methods. Then, Section 3 summarizes the common methods of signal preprocessing and feature extraction. Signal preprocessing involves the methods of signal filtering, signal segment, and key points detection. Feature extraction contains features in the time and frequency domain, involving both statistical features and transform domain features. Next, Section 4 illustrates the common classification methods of template matching methods, machine learning methods, and deep learning methods. Section 5 reviews the evaluation metrics to test model performance. Finally, Section 6 concludes the paper.

2. Wearable devices-based biometrics

In this section, we focus on the data acquisition module in the user authentication system and summarize various kinds of data collection equipment and methods. Biometrics collected from wearable devices can be categorized as physiological biometrics and behavioral biometrics. Physiological biometrics refers to the internal signal produced by the human body that varies among different people. In comparison, behavioral biometrics refers to a period of continuous human behavior, which relates to an individual's unique behavior pattern determined by long-term habits.

2.1. Physiological biometrics

In this subsection, we introduce the data acquisition methods of various human internal biometrics, including heartbeat signal, breath signal, and muscle signal. Unlike traditional biometric signals like fingerprint or face ID, which are easy to be reproduced. Finger-print could be collected from object surfaces, and faces could be obtained from photos. Physiological Biometric signals demonstrate a specific electrical trace, making the authentication system hard to be attacked. Additionally, Physiological Biometrics could be easily collected through wearable devices without any discomfort and burden.

2.1.1. Heartbeat

Electrocardiogram (ECG) signals are the most widely used heartbeat signals in the user authentication system. ECG signals reflect an individual's unique heartbeat pattern based on physiological characteristics, which measure a periodic waveform recording certain heart events in a cardiac cycle. Due to its uniqueness, easy accessibility, and hard reproducibility, ECG signals become a promising biometric ID for robust authentication system [19].

ECG signals have been widely used in heart disease diagnosis [20,21] and heart rate variability assessment [22,23]. ECG signals provide the most accurate heartbeat measurement, detecting the timing and strength of heart electrical activity, usually recorded by the electrodes attached to the body surface. ECG signals depict heartbeat pattern through several fiducial features involving three fundamental waves (P-wave, T-wave, and QRS-complex) and five major intervals (PR-interval, PR-segment, QT-interval, ST-interval, and ST-segment). During each heartbeat, atria and ventricle depolarize cause the emergence of P-wave and QRS-complex. In contrast, T-Wave occurred as the ventricle starts to repolarize. Five core intervals, also known as ECG fiducial features, demonstrates the peaks, boundaries, and intervals among three waves.

Standard ECG acquisition device is generally used in medical institutions. The wearable instrument includes 12 leads placed on the user's skin from chest to wrist and ankle. These electrodes could detect the small electrical changes caused by cardiac muscle depolarization and repolarization throughout each cardiac cycle.

For the convenience of daily use, multiple methods have been proposed to overcome this issue. [24] reduced the number of required electrodes by using a single electrode placed on each side of the heart to measure ECG signals. Sun et al. [25] further proposed a wearable T-shirt embedded with a conductive fabric electrode. The electrode collects ECG signals by pressing against the chest inside of the T-shirt. Then motivated by the healthy monitoring application based on wearable devices, Beach et al. [26] presented an ECG smart wrist embedded with an IoT platform to monitor the heart condition. Nowadays, the Apple watch, the best-selling wearable device designed for a healthy life, is embedded with 1-lead ECG, which is sufficient for user authentication without wearing an extra device [19].

2.1.2. Breath

Breath is another physiological biometric for user authentication since each individual displays a unique breath pattern of breath rate and amplitude according to the individual's age, weight, height, or gender. Breath is a physiological process that is periodically moving airflow into (inhalation) and out of (exhalation) the lungs, controlled by the body parts of the trachea, lung, and diaphragm. Breathing occurs continuously and naturally in every living life without extra corporeal or cognitive efforts. A healthy adult breathes approximately 12–20 times per minute, accompanied by the easy-measurable breath sounds. Therefore, the uniqueness, constituency, accessibility of breath sound makes it a convincing signal for individual identification.

Generally, breath sounds could be divided into two categories according to the detection positions. Vesicular breath sounds are frequently detected from the chest wall, and bronchial breath sounds are measured upon the large airways near each side of the neck. The chest wall is far away from large airways and can be regarded as a low-pass filter to remove high-frequency elements in sound. Therefore, Vesicular breath sounds mainly contain low-frequency components while bronchial breath sound includes both low and high-frequency components [27]. Breath sounds could also be detected under the nose or identified through a period of speech with breath intervals.

BreathPrint designed a breathing acoustics-based user authentication system based on breathing gestures [28]. Breathing audio is collected through three distinct breath gestures of sniff breath, normal breath, and deep breath measured by a microphone placed 1–2 cm under the user's nose. The combination of three different types of breath signals provided sufficient information to distinguish individuals. According to both statistical features and extracted features through algorithms, the signal pattern among different users may similar under one breath gesture but hard to be alike in three-dimensional breath gestures. However, system performance is largely affected by the measurement distance and environmental noise.

Besides, to avoid additional breath sounds collection process, other studies shift to focus on identifying breath sounds in a speech signal period. These studies' main challenge is how to automatically detect the breath sounds in an audio clip, referring to the correct breath demarcation procedure.

We depict Fig. 2 to illustrate the accurate breath demarcation algorithm, which generally involves three steps. A general breath template is first constructed from a small number of breath examples, and the Mel frequency cepstral coefficients is commonly used as features to construct a template. For each consecutive overlapping frame, a parameter matrix is calculated and compared with the template. The frame is determined as breath when the similarity between template and signal frame exceeds the predefined threshold. After the breath is initially detected, a refinement procedure is executed for accurate breath boundaries detection, involving the features of zero-cross rate, short-time energy (STE), B index, etc. [29]. Finally, the accurate breath signal is segmented through edge detection, spectral scope, and time duration [30].

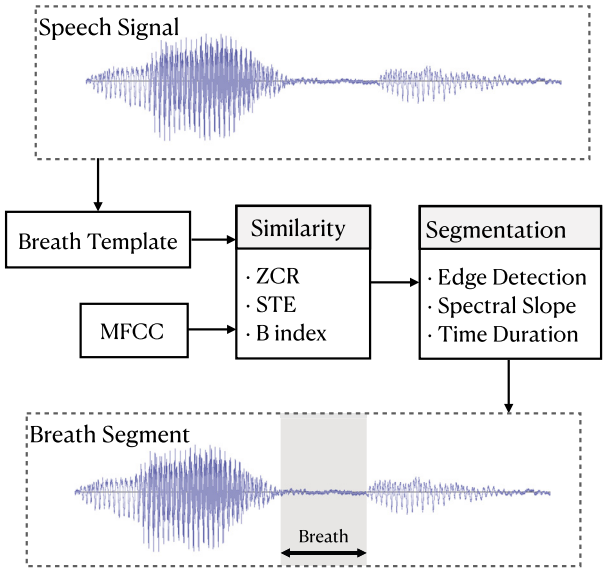

Fig. 2. Accurate breath demarcation process.

However, the datasets used are basically of high quality collected from professional speakers in a controlled recording environment, making the methods too sensitive to handle spontaneous speech directly. Therefore, Dumpala et al. [31] proposed a rule-based algorithm to detect breath sounds in spontaneous speech using both excitation source and vocal tract system-based characteristics. However, it still needs to manually distinguish breathing sounds in noise like breath, laughter, cough, etc., from the Buckeye corpus.

2.1.3. Muscle

As an important physiological biometric, muscle-related features are frequently applied for user authentication. In detail, these features include finger muscle isometric contraction password (FMICP) [32], high-density surface electromyogram (HD-sEMG), and electrical muscle stimulation (EMS) [33]. In FMICP mode, people can enter their password by isometric contraction of different finger muscles in a prescribed order, without actual finger movements. For HD-sEMG, it could be acquired from the dorsum of the user's; the subject performed isometric contraction of different finger muscles to enter the FMICP. With the help of HD-sEMG, the isometric contraction patterns of different finger muscles can be recognized between individuals. These two features are usually combined as a whole neuromuscular password, which can realize double security. EMS is a *challenge-response* as a form of active biometric authentication method. Many biometrics, such as the fingerprint, have a common shortcoming: when they are stolen, people can not reuse the same data. EMS can overcome this drawback by replacing the breached challenge-response pairs with new responses to a series of challenges.

For the collecting of HD-sEMG, a SAGA 64+ system of Twente Medical Systems International BV at a sampling rate of 4000 Hz is utilized to collect the 64-channel monopolar HD-sEMG signals. In this process, the 8 $\times$ eight flexible high-density electrode array with 8-mm interelectrode distance was placed in the center of the dorsal aspect of a subject's right hand. During isometric contraction of different finger muscles, sEMG signals can be recorded on the forearm, palm, and dorsal hands.

For the acquisition of EMS data, the Hasomed Rehastim, a medical compliant device with eight individually controllable channels, is applied in the delivery of EMS impulses. Before collecting EMS

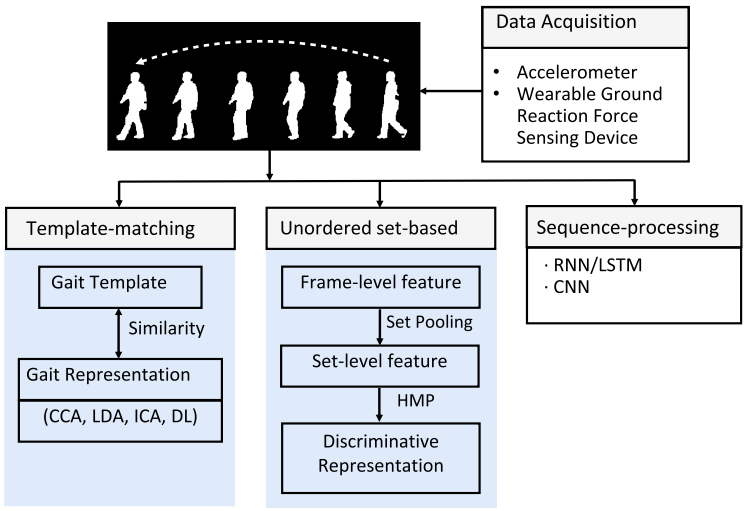

Fig. 3. The architecture of gait recognition methods.

data with this tool, it is necessary to calibrate the electrode placement for each user. After sending EMS challenges to the user, a motion sensor, such as IMU, can capture the limb's movements. In this case, challenge-response pairs could be sampled.

After obtaining HD-sEMG signals, macroscopic features and microscopic features can be extracted based on these signals. Usually, independent component analysis (ICA) can extract microscopic features. Macroscopic feature extraction methods include sample entropy, spectral entropy, frequency median (FMD), waveform length (WL), and root mean square (RMS).

2.2. Behavioral biometrics

In this subsection, we summarize the behavioral biometrics in the user authentication system. Behavioral biometrics reflects user identity by their unique behavior patterns. We mainly focus on biometrics such as gait, gesture, keystroke, and mouse dynamics.

2.2.1. Gait

Gait recognition stands for authenticating individuals by the manner of their walking style. Gait is a behavioral biometric that is convenient to detect through simple instrumentation like a camera or accelerometer. It remains unique among different individuals due to the specific arm swing amplitude, step frequency, and length. Gait recognition especially plays an essential role in criminal investigation to determine the identity of the suspect. Generally, walking pattern captured through cameras is analyzed through a period of gait signal, either regarding gait as an image including all gait silhouettes or the original gait silhouettes video sequence.

We depict Fig. 3 to demonstrate the architecture of gait recognition methods involving data acquisition sensors and three kinds of mainstream gait recognition methods. Gait could be collected from both wearable devices like accelerometers and remote sensors like cameras and radars. Apart from the general approach that follows the pipeline of feature extraction and classification, a high-performance gait recognition system utilizes template-matching methods, unordered set-based methods, and sequential deep learning models. The details of these classification methods will be illustrated in Section 4.

The commonly used wearable devices in gait recognition are described below:

Accelerometer. Accelerometer sensors are usually placed on an individual's body to record three-dimensional accelerations in a gait recognition system. Early studies developed goal-oriented wearable devices placed on the user's waist or user's lower leg to collect gait samples from various directions, including vertical, sideways, and forward-backward accelerations. Later studies start to use the accelerometers embedded in commercial mobile phones and execute real-time gait recognition based on mobile computation resources [34,35]. Further studies commit to optimize the gait identification system by building a speed-adaptive algorithm [36] and adapting to a more challenging wild environment where individuals do not need to walk along a specific road at an average speed [37]. Recently, Xu et al. also designed several user and device authentication systems for wearable devices such as smart glass and smart watch [38–40]. Moreover, because continuously sampling accelerometer quickly drains the battery of wearable devices, researchers also started to use wearable energy harvesting to authenticate users based on their walking patterns [41,42].

Wearable Ground Reaction Force Sensing Device. A wearable Ground Reaction Force (GRF) sensor is a force plate attached to the bottom of the shoe constructed by several small triaxial force sensors. The sensing system constructs a global coordinate system to evaluate gait through the relationship between triaxial position and force. The X axis is in line with the moving direction while the Z axis the vertical, and Y forms a right-handed coordinate system together with the other two axes [43]. Therefore, all the local coordinate systems defined for each triaxial sensor could be accurately aligned [44].

2.2.2. Signature

Various gestures have been applied to user authentication systems, including hand gestures, touch gestures, finger gestures, and gaze gestures. These signals are usually captured by cameras, touch screens, or wearable devices. In this part, we mainly focus on the gesture of an individual's signature, corresponding to the wrist motion when people are signing. Current smart wrist devices like fitness trackers and smartwatches are generally embedded with accelerometers and gyroscopes, making it convenient to record user's wrist motions. Early work like MotionAuth [45] authenticates users through four simple gestures, raising the hand, lowering hand, rotation, and circle, making it easy to be imitated and attacked. On the contrary, signatures are continuous signals that obey individual patterns among users, hard to replicate, and more robust against attacks.

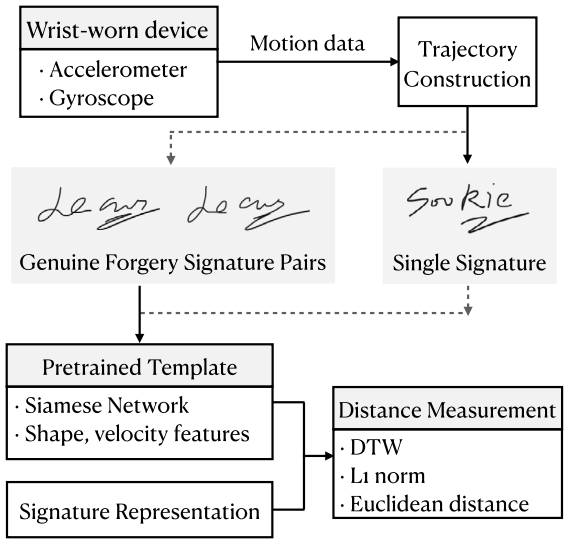

Fig. 4. The architecture of signature authentication methods.

A signature is viewed as a sequence of spatial coordinates of a motion trajectory. Template-based approaches measure the difference between the original signature template and the compared signature in the space of pre-defined gestures. The first fundamental work to address the signature verification problem is the Siamese neural network, which contains two or more subnets with shared parameters, and the dissimilarity between two vectors is calculated by the cosine distance [46,47]. Nassi et al. [48] collected data from accelerometers and gyroscopes and trained the classifier through a set of genuine and forged signatures to verify the signature. However, one user's forgery signatures are hard to acquire in practice. Therefore, Lyu et al. [49] turned to utilize interaction behavior data between users and wrist-worn devices and depict user's unique writing pattern through the proposed fine-grained writing metrics, based on more effective Savitzky-Golay filter and Dynamic Time Warping (DWT) method.

Current signature verification methods on wearable devices are generally based on the template-matching pipeline. We depict Fig. 4 to illustrate the architecture of signature authentication methods. The motion data is first collected through wrist-worn or finger-worn devices equipped with motion sensors (accelerometers or gyroscopes). A significant difference between the signature authentication system and other authentication systems lies in the trajectory construction module, which maps the motion data into signing traits. There are two methods for signature verification based on the type of training data. One is signature pairs of genuine and forgery signatures, commonly used for Siamese network-based methods, aiming at calculating the similarity between input pairs (genuine and genuine signature pairs or genuine and forgery signature pairs); the other is only the single signatures considering the forgery signatures are rare. The pre-trained template is usually constructed with fiducial features like shape and velocity.

In-air handwriting authentication is another potential application scenario needs on wearables, which excludes the extra touch screen to record handwritten images [50]. Kinematic theory characterizes rapid human movement by strokes and virtual points, corresponding to the motion trajectory in signing, which provides the fundamental model for signature verification. By observing the fact the authentication accuracy is not stable due to the inaccuracy in gravity accelerations, Huang et al. [51] proposed only to use the gyroscope sensor to capture the signing motions, avoiding the impact of gravity. They developed an accurate trajectory construction method that maps wrist motion to the signing traits.

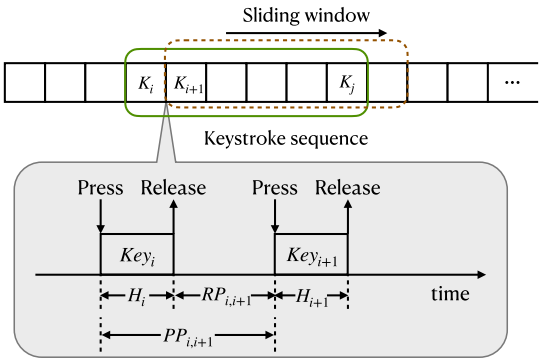

Fig. 5. Illustration of continuous keystroke sequence authentication.

Raw signature data collected from accelerometers and gyroscopes are first smoothed through a lowpass filter to reduce the noise. Magnitude features of three-dimensional components from the accelerometer and orientation data from the gyroscope are gathered. Besides, basic statistics, gradients, and eigenvectors of each measured component are also calculated as features [52].

2.2.3. Keystroke

Keystroke behavior records an individual's unique keyboard typing style, which is an emerging trend in the user authentication system. The main advantage of key-stroke-based authentication is low cost. Since other biometric signals like fingerprints or face IDs need to be collected through expensive biosensors, keystroke data could be acquired through only a keyboard without any special equipment. Besides, each individual has a unique typing pattern that will not change over time; typing pattern could be reflected on the typing speed, typing time intervals among each letter, commonly used typing keys, etc.

From previous literature, keystroke authentication technologies could be broadly identified as two stages, from Keystroke Static Authentication (KSA) to Keystroke Continuous Authentication (KCA). Static in KSA suggests that the text being typed is fixed and predefined, mostly the passwords and PINs used in the user login process. However, fixed texts contain too little information to represent the user's typing pattern and suffer low scalability. KSA employs template-matching architectures, which compares the current keystroke sample to a recorded typing template expressed as feature vectors [53,54].

Later, researchers started to focus on continuous user authentication based on free text. As depicted in Fig. 5, continuous authentication processes a keystroke sequence through a sliding window at the time scope. As the sliding window moves forward, the authentication system could continuously check the new inputs in the keystroke sequence, achieving the goal of continuous identity-checking as the user typing in a period of time. Besides, the system records the keystroke timing information as keystroke identification features. H_i indicates the holding time of Key_i, from key pressing timestamp to key releasing timestamp; $PP_{i,i+1}$ represents the time interval of nearby key pressing times and $RP_{i,i+1}$ is the time interval of former key releasing time and later key pressing time.

Several public benchmark databases and algorithms are listed as follows. Sun et al. [55] released the Buffalo keystroke dataset, a large publicly accessible dataset for long text, and tested the dataset with existing Gaussian mixture model, which assumes the digraph patterns in keystroke data are multiple distributions via Gaussian Mixture Model (GMM) rather than a single Gaussian distribution [56]. Vural et al. [57] obtained a new keystroke dataset including short pass-phrases, free text, and fixed text (transcrip-

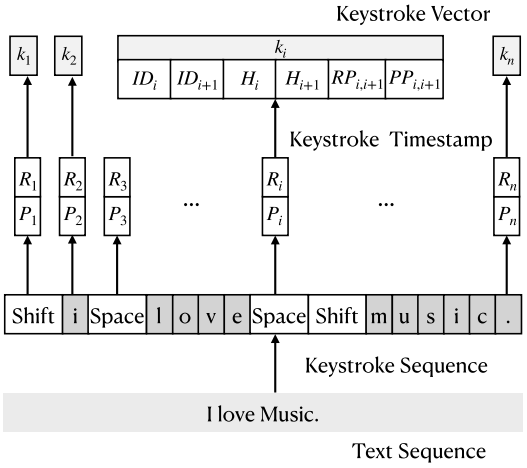

Fig. 6. Illustration of keystroke vector generation process from a text sequence.

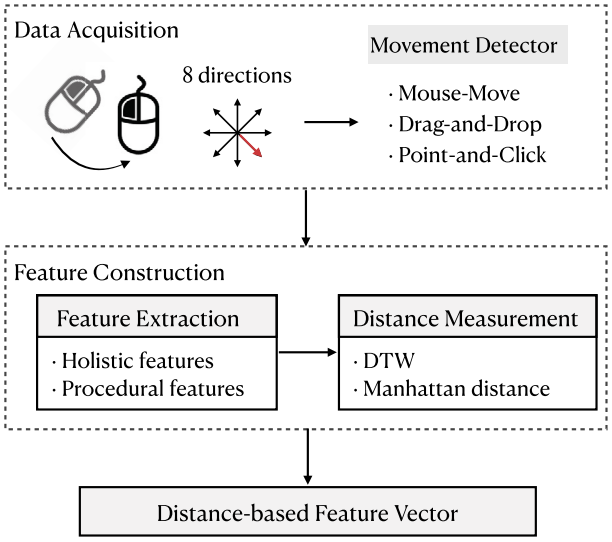

Fig. 7. Mouse dynamics feature vector generation process.

tion of long proses), and provided the baseline results of current state-of-the-art method in free text keystroke, Gunetti & Picardi's algorithm [58], as a testbed for future improvements. Murphy et al. [59] provided a large keystroke dataset including keyboard operations, mouse dynamics, and activate programs, captured in completely natural and uncontrolled settings.

Huang et al. [60] proposed a larger dataset of users' normal computing interaction behavior with the personal laptops and introduced a Kernel Density Estimation (KDE) algorithm which estimates the probability density of a digraph in the reference profile and test samples. We depict Fig. 6 to illustrate the keystroke generation process from a text sequence. The keystroke sequence and the keystroke timestamps (pressing timestamp P_i and releasing timestamp R_i) of a corresponding text sequence is recorded to generate keystroke vector k_i for each keystroke.

Accurate feature extraction from the keystroke of long free text is an essential part of the keystroke authentication system. Some statistical features utilized text sub-word frequency. For instance, one selected typical keystroke feature is the average PP-intervals, calculated based on the most frequently occurred digraphs and trigraphs. R-measure and A-measure are also used as standard features in previous studies. R-measure is the consistency of relative typing speed order for common syllables between two users, while A-measure is the absolute typing speed similarity for common syllables.

Another type of keystroke feature is generated from the keyboard layout, where keys are divided into three groups L group (keys typed by the left hand), R group (keys typed by the right hand), and S group (only the space key). Thus each digraph in keystroke data could be considered as an action pair selected from eight categories of L-L, L-R, L-S, R-L, R-R, R-S, S-L, S-R, and the keystroke data transforms into an eight-dimensional vector with the P-P interval value of the action pair.

Compared with the keyboard layout, typing speed of a digraph could be a more robust feature of typing behavior. Thus the typing-speed rank is obtained as a keystroke feature, ordered from the fastest-typed digraph to the slowest-typed digraph [61].

2.2.4. Mouse dynamics

Similar to keystroke dynamics, mouse dynamics is another behavioral biometric for continuous user authentication, which automatically and passively monitors user mouse behaviors without extra expensive sensor devices. We depict Fig. 7 to illustrate the mouse dynamics feature vector generation process. Mouse dynamics is acquired by a mouse movement detector, which records various movement characteristics, including movement type, movement speed, movement direction, traveled distance, etc. Movement type could be classified by the conception described in [62], which includes Mouse-Movement (MM) action, Drag-and-Drop (DD) action, and Point-and-Click (PC) action. As for movement direction, it could be specified by a continuous variable of movement angle or a simpler discrete definition of eight angles averagely divided from 0-degree to 360-degree. Each 45-degree angle interval is sequentially numbered from one to eight [63].

Shen et al. [64] mapped the combination of traditional holistic features and user procedural features into distance-based features and then identify user identity through a support vector machine classifier. Shen et al. [65] conveyed comprehensive performance evaluation methods on an equal basis. Monda et al. [63] introduced several techniques to improve performance on the same dataset, including Multi Classifier Fusion (MCF) to combine multiple classifiers, score boosting algorithm, weighted fusion scheme, and static/dynamic trust model.

Apart from traditional machine learning-based approaches, recent studies incorporated more novel technologies, such as deep neural networks, multi-biometric fusion techniques, and different training methods. Kasprowski et al. [66] employed a fused feature analysis on mouse dynamics and eye movement biometrics for the first time, which defined a robust authentication model with the comparable result by using shorter mouse recordings. While Monda et al. [67] proposed Pairwise User Coupling (PUC) of the combination of keystroke and mouse dynamics. Other studies focus on eliminating manual feature design by automatically extracting features through neural networks, such as convolutional neural network [68], recurrent neural network [69], and the combination of this two networks [70].

Mouse dynamics record the mouse operations in each timestamp. The fiducial features could be generally divided into holistic features and procedural features. Holistic features depict the overall static properties of mouse dynamics, such as single-click statics, double-click statics, mouse movement offset, mouse movement elapsed time, etc. Procedural features are the property at the operation moment, including movement direction, movement type, movement speed, etc. Since holistic features and procedural features do not contain enough information to reveal an individual's typing pattern, distance measurement methods are applied to generate a distance-based feature vector. The distance feature vector is obtained by calculating the distance between all pairs of raw training features.

3. Signal processing for feature extraction

In this section, we summarize the signal processing methods, which contain signal preprocessing methods and feature extraction methods. Signal preprocessing methods provide a clear base signal for the feature extraction module. Data preprocessing module first filters high-frequency noise and segments signals into periodic cycles. Next, the key points and key intervals are detected in the signal as fiducial features. The feature extraction module aims to extract useful features to represent the signal. The features could be divided into two categories, statistical features, and signal transform domain features. Statistical features are calculated to analyze data distribution in both time and frequency domain. Transform domain features are extracted from various kinds of widespread feature extraction approaches in the signal-processing field.

3.1. Signal preprocessing

The raw biometric signal consists of various inevitable high-frequency noise during the measurement process, including frequency interference, baseline drift, muscle noise, and motor artifacts [71]. Preprocessing step is essential for feature engineering since it could suppress noise and observe signal pick clearly, providing stable and usable signals to commit accurate features. The biometric signal preprocessing process mainly includes signal filtering, peak point detection, signal segmentation, and feature selection. The first step in data preprocessing is data filtering to remove the noise. Based on related studies, many methods have been proposed to filter noise in biometric signals. One way uses a band-pass Butterworth filter to remove a given band of frequency from the signal without degrading signal quality. For ECG signals, a Band-pass Butterworth filter between 40 Hz–100 Hz can be applied to eliminate motion artifacts. Baseline wanders and a 50Hz notch filter can be used to eliminate line noise [72].

Another method for biometric signal preprocessing is wavelet denoising. Features with large wavelet coefficients are reserved, while others are considered as noise to be removed. Moreover, the adaptive filter is another method for biometric noise removal. The least mean squares (LMS)-based adaptive filter is mostly adopted due to its computational simplicity [73]. LMS filter algorithm aims to learn a target filter by minimizing the least mean square of the error signal. Some LMS algorithm variants like Normalized LMS (NLMS) and Delayed Error Normalized LMS (DENLMS) adaptive filter are further adopted for low latency and less computational consumption.

The filtered biometric signal is then segmented and aligned into cardiac cycles. Key points and intervals are detected as fiducial features. As for the ECG signal illustrated in Fig. 8, five typical peak points (P, Q, R, S, and T point), boundaries, and intervals between them are detected in each cycle. Every R-R interval, duration between two peak points in ECG signal, is usually considered a heartbeat. Major fiducial features such as QRS interval, QRS amplitude, R-R interval, and ECG wave slope are utilized as classification features. After signal filter and segmentation, the signal is finally normalized for further process.

3.2. Signal processing for feature extraction

Features for biometric signals could be obtained through both statistical methods and various signal processing algorithms. Statistical methods are the first techniques to be executed to form a comprehensive insight into the structure and distribution of the signals. Then more specific signal processing methods would be applied to extract more profound features.

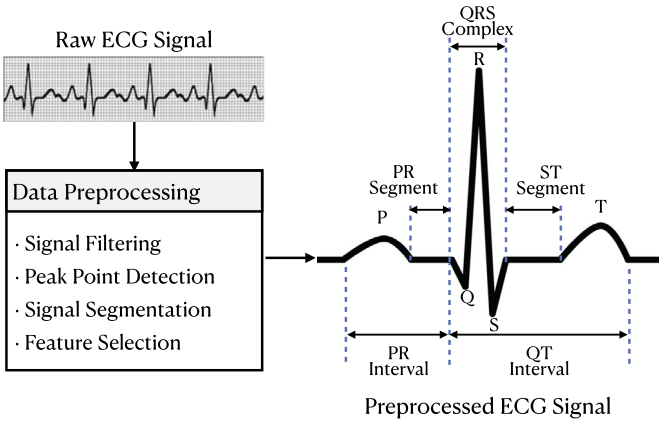

Fig. 8. ECG signal preprocessing.

3.2.1. Direct-domain signal processing for feature extraction

Statistical features such as mean, maximum, minimum, variance, skewness, and Shannon Energy are first analyzed [74]. When the time-domain signal is transformed into the frequency domain, other basic statistical features in the frequency domain for waveform signal are also calculated. Raw biometric signals are analyzed based on time, frequency, and energy. Some useful statistical features are described as follows:

Variance. Variance measures the average squared differences of a random variable from its average value, indicating how far the numbers are distributed from the average. A higher variance represents a larger amplitude variation in the given ECG signal, which could serve as a distribution feature to distinguish signals. The mathematical equation of variation is defined as

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}[x_i - \bar{x}]^2 \tag{1}$$

Where N is the sample number of the signal and $\bar{X}$ is the mean value.

Skewness. Another distribution feature, Skewness, measures the asymmetry of the probability distribution of a random variable about its mean value. It indicates how symmetric the samples are spread out around the peak point. Consider the input signal with N samples. Skewness is mathematically expressed as the average cubed deviation from the mean $\bar{X}$ divided by the cubed standard deviation S,

$$Skewness = \frac{\sum_{i=1}^{N}(x_i - \bar{x})^3/N}{S^3} \tag{2}$$

Shannon Energy. Shannon energy measures the energy of the local spectrum for each sampled biometric signal, indicating how much the information content is conveyed within a trial. Besides, Shannon energy transfers the high elements into lower elements, leading to the advantage of retaining the capacity to emphasize medium compared with other classic energy. Shannon energy is the product of the preprocessed biometric signal value x_i and its logarithm value, where i is the index of the sampled signal.

$$SE = -|x_i|\log(|x_i|) \tag{3}$$

Zero Crossing Rate (ZCR). ZCR corresponds to the times that biometric waveform changes its sign (cross zero), normalized by

the window length N in samples. This feature is considered in breath signal, since it contains both inhalation and exhalation process [30].

$$ZCR = \frac{1}{N}\sum_{i=1}^{N}\mathbb{I}(x_i x_{i-1}) < 0 \tag{4}$$

Short Time Energy (STE). STE specifies the signal amplitude of a certain signal point over a period of time, which is commonly used in audio processing to distinct voiced and unvoiced sounds from silence [75]. STE can be calculated as:

$$STE = \frac{1}{N}\sum_{i=1}^{N}x^2[i] \tag{5}$$

Power Spectral Density (PSD). PSD is a measure of signal power versus frequency, which is the magnitude squared of the Fourier Transform. Consider the collected data $X(t)$ with autocorrelation function of Fourier transform $R_X(\tau)$, PSD written in $S_X(f)$ is defined as

$$S_X(f) = \mathcal{F}R_X(\tau) = \int_{-\infty}^{\infty} R_X(\tau)e^{-2j\pi f\tau}d\tau \tag{6}$$

where $j = \sqrt{-1}$.

Other features. Some other features are also utilized to discriminate a group of different subjects. Occupied bandwidth is a bandwidth scope that contains 99% the integrated power of the signal. Median frequency represents the midpoint (lines 50% of the total power) of the power distribution. Spectral edge frequency is similar to occupied bandwidth, which contains 95% of the total power and serves as an upper bound of the power spectrum. Spectral slope reflects the intensity of the harmonics, which is computed through Discrete Fourier Transformation at frequencies of $\frac{\pi}{2}$ and π.

3.2.2. Transform-domain signal processing for feature extraction

Since the continuously collected waveform signal is in the time domain, various signal processing methods are applied to obtain more frequency domain features. Signal processing methods are widely used as feature extraction algorithms in speech recognition, so the signal processing-based methods are utilized as pivotal features for biometric signals such as ECG signal [71], breath signal [76], gait signal [77], keystroke signal [78], gesture signal [79], etc.

Fast Fourier Transform (FFT). FFT transfers the time-domain signal into a frequency-domain signal with individual spectral components; it is an optimized version of Discrete Fourier Transformation with less computation complexity, which is defined as

$$X_k = \sum_{i=1}^{N}x_i e^{\frac{-i2\pi ki}{N}}, \quad k = 0, 1, .., N-1 \tag{7}$$

Discrete Cosine Transform (DCT). DCT could be regarded as a special case related to Fourier transform on real number space. DCT is widely applied in lossy compression due to the energy compaction property of the cosine function, where signal information is mainly concentrated in low-frequency. As illustrated in Equation (8), DCT transforms real numbers $x_i \in (x_1, x_2, ..., x_N)$ into corresponding real numbers $f_m \in (f_1, f_2, ..., f_N)$, the m_{th} transformed discrete number is calculated as:

$$f_m = \sum_{i=1}^{N}x_i \cos\left[\frac{\pi}{N}m\left(i+\frac{1}{2}\right)\right] \tag{8}$$

Discrete Wavelet Transform (DWT). DWT is a commonly used mathematical tool in signal processing, which decomposes a signal into a set of the orthogonal waveform in time and frequency domain through a Low Pass Filter (LPF) $g[\cdot]$ and a High Pass Filter $h[\cdot]$. Each level of decomposition follows the following two equations, which separate signals into sub-bands of different frequencies and resolutions.

$$y_{low}[k] = \sum_n x[n]h[2k-n] \rightarrow y_{low} = (x * h)\downarrow 2 \tag{9}$$

$$y_{high}[k] = \sum_n x[n]g[2k-n] \rightarrow y_{high} = (x * g)\downarrow 2 \tag{10}$$

Mother wavelet function and decomposition level would primarily affect signal analysis, which varies among different tasks and signals. As for ECG signals, Haar and Symlet order 7 function at level 6 decomposition are frequently applied to extract subbands information [71].

Mel-Frequency Cepstral Coefficient (MFCC). MFCC is a linear representation of a cosine transform of a period of the logarithmic power spectrum of the speech signal on a nonlinear scale Mel frequency [80], containing a pipeline of framing, windowing, Fast Fourier Transform (FFT), Mel Filter Bank, and Discrete Cosine Transform (DCT).

MFCC is positively related to the window size; a larger time window could enhance the frequency domain's frequency resolution. Therefore, the rectangle window for ECG data is centered at the R peak point. According to the data validation result, the window size is selected, demonstrating that this window could capture whole essential markers in a cardiac cycle. Unlike the QRS complex feature, MFCC window size is usually set larger than the QRS complex, involving the P and T waves. The framing step divides the signal into equal length frames of 10–30 ms. Then Hamming window is usually applied to maintain signal continuity as equation below, where each frame contains N samples and $W[n]$ is the n_{th} Hamming window coefficient:

$$W[n] = 0.54 - 0.4\cos\left[\frac{2\pi n}{N-1}\right] \tag{11}$$

$$Y[n] = X[n] \times W[n] \tag{12}$$

Later FFT transfers each frame in the time-domain signal into a frequency-domain signal as equation below, and Mel scaling the signal through Mel Filter Bank $H[m]$ with Q filters.

$$Y[m] = \frac{1}{N}\sum_{n=0}^{N-1}Y[n]e^{\frac{-2\pi nm}{N}} \tag{13}$$

Finally, MFCC coefficient $C[k]$ is calculated through Discrete Cosine Transform (DCT) as follows:

$$C[k] = \sum_{i=0}^{Q-1}\ln\Big(\sum_{m=0}^{Q-1}|Y[m]|^2 H_i[m]\Big)\cos\Big(\frac{\pi k(i-0.5)}{Q}\Big) \tag{14}$$

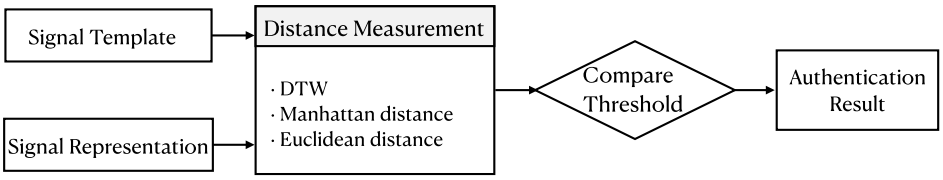

Fig. 9. The architecture of template matching method.

4. Classification

After feature extraction, the system performs feature selection algorithms to reduce feature dimension. Features that have a high correlation with other features are removed through dimensionality reduction algorithms like Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA).

The selected features are then fed into classifiers to obtain output probability. The user authentication system is mostly a binary classification problem to output whether the obtained signal belongs to the current user. It is a multi-classification problem for user identification to identify the obtained signal belongs to which user. There are three major classification methods, namely template matching methods, traditional machine learning methods, and deep learning methods. Template matching methods first construct a signal template using the extracted features and return an authentication result based on a threshold. However, template matching methods require careful distance measurement selection and threshold setting. Therefore, classification models utilize machine learning algorithms or deep neural networks to classify the given biometric signals are widely utilized. blue Machine learning methods and deep learning methods both can learn a model for classification. However, machine learning methods usually take some certain features as input, while deep learning methods just take the original data as the input and extract useful features automatically. Also, generally, deep learning methods require more data than machine learning methods.

4.1. Template matching methods

The main idea of the template matching method is illustrated in Fig. 9. The template matching method compares the distance between a signal template constructed from the acquired database and a signal representation constructed by features in the collected signal. The signal template is regarded as the average value of detected signal cycles. Signal representation is extracted through Linear Discriminant Analysis (LDA) [81], Canonical Correlation Analysis (CCA) [82,83], Independent Component Analysis (ICA) [84,85] or Deep Neural Networks (DNN) [86,87]. The similarity between the signal template and the signal representation is calculated through Euclidean distance or some other metrics [88]. Finally, the authentication result is then determined by the comparison result between the calculated distance value and the threshold.

We take the gait template construction as an example, which contains step length estimation and gait cycle extraction [81]. Accelerometers detect three-dimensional accelerations (x, y, z) respectively from the x-axis, y-axis, and z-axis. Thus the raw data could be represented as

$$data = \sqrt{x^2 + y^2 + z^2} \tag{15}$$

Then, step length is approximately estimated as the quotient of acceleration sampled frequency f_{sample} and the step frequency f_{step} as in Equation (16). The step frequency is represented as the maximum point of the spectral density, which is obtained by transforming signal into frequency domain through Fast Fourier Transform:

$$L_{step} = \frac{f_{sample}}{f_{step}} \tag{16}$$

The gait cycle is a periodic segment measuring the time interval of the same foot repeatedly touches the ground again. It is an essential component for gait data segmentation. However, gait cycle extraction is harrowing due to an individual's changing walking speed or irregular body movement, leading to the accelerometer readings turn distorted [36].

Considering a gait cycle, its start point sp is the first minima point in gait raw data sequence d and the endpoint ep can be set by Equation (17), where L_{step} is the estimated step length and d is a parameter to measure the floating range between estimated step length and actual step length considering the changing gait speed, controlled by the deviation coefficient β:

$$sp + L_{step} - d < ep < sp + L_{step} + d \tag{17}$$

where $d = \beta L_{step}$.

Finally, an individual's gait template could be represented as the mean value of N gait cycles as the estimation above. The gait template is represented as the normalized gait cycle. Since each gait cycle has a different step length, so a linear interpolation is performed to map inconsistent gait cycle into a normalized cycle with step length m; thus each gait template could be defined as:

$$T_n = \{X_{sp_n}, X_{sp_n+1}, ..., X_{sp_n+m-1}\}, n = 1, 2, ..., N \tag{18}$$

After the construction of the signal template and signal representation, the distance measurement method is executed to compare similarities.

Manhattan distance. Considering two vectors $X = (x_1, x_2, ..., x_n)$, $X' = (x'_1, x'_2, ..., x'_n)$, Each element of the vector corresponds to a point in high-dimensional space. Manhattan distance computes the summation of absolute differences between two vectors, it is widely used for low computation demands:

$$d_{Manhattan}(x_i, x'_i) = \sum_{i=1}^{n} |x_i - x'_i| \tag{19}$$

Euclidean Distance. Euclidean distance measure the length of a line segment between two vectors in Euclidean space:

$$d_{Euclidean}(x_i, x'_i) = \sqrt{\sum_{i=1}^{n} (x_i - x'_i)^2} \tag{20}$$

Discrete Time Warping (DTW). DTW measures the similarity of two temporal sequences of different lengths and rhythms through dynamic programming. DTW aims to calculate an optimal match from the first sequence to the second sequence under a series of restrictions and rules. For each matched point in two sequences,

the optimal match corresponds to the minimal loss of the sum of absolute differences between paired values.

Considering two sequences with different length $X = (x_1, x_2, ..., x_n)$, $X' = (x'_1, x'_2, ..., x'_m)$. To align two sequences using DTW, a distance matrix $M \in R_{m \times n}$ is first constructed, where each element corresponds to the Euclidean distance between the values of two sequences $d(x_i, x'_j)$:

$$M_{i,j} = d(x_i, x'_j) = (x_i - x'_j)^2 \tag{21}$$

Combined with the continuity and monotonic constraints in warping path, directions of next step in warping path could be $(i+1, j)$, $(i, j+1)$ or $(i+1, j+1)$ when the current step passes the point (i, j). Therefore, the accumulated distance of two sequences is calculated using a recursive formula:

$$\begin{aligned} d_{DTW}(i, j) = \min[&d_{DTW}(i-1, j-1), d_{DTW}(i-1, j), \\ &d_{DTW}(i, j-1)] + d(x_i, x'_j) \end{aligned} \tag{22}$$

Zhang et al. [89] utilized DTW distance-based dynamic template-matching approach for gait modeling and identifying, obtaining about 92% in recognition accuracy. De Marsico et al. [90] enhanced the DTW distance-based dynamic template matching approach by applying a new segmentation algorithm for the gait signal. It reaches about 93% of Recognition Rate (RR) on the ZJU-gaitacc dataset and an Equal Error Rate (EER) of 0.09 in the verification result.

As for ECG signal-related user authentication, Shdefat et al. [91] used a template matching technique for the authentication process and achieved the accuracy rate 97.2% with a false acceptance rate of 1.21%. Will et al. [92] proposed an advanced template matching (ATM) algorithm containing multiple heterogeneous templates to enhance the performance regarding instantaneous heartbeat detection. It reduced the root mean square error (RMSE) of the interbeat intervals in ECG signals from 68.2 ms to 18.0 ms compared to a standard template matching algorithm.

4.2. Traditional machine learning methods

Various kinds of machine learning methods-based classifiers are widely employed for user authentication systems, involving Decision Tree [64,65], Support Vector Machine, Hidden Markov Model, etc.

Decision Tree. Decision tree is a common machine learning algorithm based on a tree structure, an iterative algorithm following the "divide and conquer" strategy. At each step, the optimal attribute a^* is selected for more information gain $Gain(D, a)$ based on information entropy $Ent(D)$:

$$a^* = \arg\max_{a \in A} Gain(D, a) \tag{23}$$

Each discrete attribute a has V values $a^1, a^2, ..., a^V$, thus V branch nodes would be generated given attribute a. Due to each branch node contains a different sample number D^V, a sample-number related weight is attached to the information gain:

$$Gain(D, a) = Ent(D) - \sum_{v=1}^{V} \frac{|D^V|}{|D|} Ent(D^V) \tag{24}$$

Considering dataset D contains $|y|$ classes, and the sample proportion of class k is p_k, thus the information entropy $Ent(D)$ is defined as:

$$Ent(D) = -\sum_{k=1}^{|y|} p_k \log_2 p_k \tag{25}$$

Additionally, pre-pruning and post-pruning techniques are usually applied to a decision tree to prevent over-fitting, and an advanced oblique decision tree is used in complex classification boundary situations by employing the linear combination of attributes.

Alex Santos et al. [93] use a decision tree based on random forest to process the ECG signal for continuous authentication. Based on the QRS complex features of the ECG signal, the decision tree can achieve a true positive rate of 95.8% for five users identification on a dataset from the Physionet Database.

Support Vector Machine (SVM). SVM aims to find a hyperplane to divide the dataset with maximum margin. Consider the hyperplane is defined as $w^T x + b = 0$, and the support vectors are the closest points to the hyperplane to classify the dataset. Thus the objective function could be written as:

$$\begin{aligned} &\min_{w,b} \frac{1}{2} \|w\|^2, \\ &\text{s.t. } y_i(w^T x_i + n) \geq 1, \quad i = 1, 2, ..., m \end{aligned} \tag{26}$$

This is the fundamental SVM version used for the linearly separable dataset, while for the dataset that a hyperplane could not separate, the kernel function is applied to map samples to a higher dimensional feature space.

Sugondo et al. [94] use a series of support vector machine methods based on the ECG signal and related features obtained by processing the original signals with Hjorth Descriptor and Sample Entropy for user authentication. The result shows that the Gaussian SVM achieves the highest accuracy value of 93.8% and 86.2% on Hjorth Descriptor features and Sample Entropy features.

Besides the ECG signal and the metric of accuracy, SVM [95] is also used in biometric user authentication based on the keystroke with the metric of an average equal error rate of 8.6% with a standard deviation of 0.0627.

Hidden Markov Model (HMM). HMM is a primary dynamic Bayesian network, a probabilistic graphical model that depicts probability relationships through graph structure, which is commonly used for formulating time series data. A Markov chain describes the dependency relationship between states y_i and observations x_i, following the assumption that current state y_t is only determined by the previous state y_{t-1}, thus the joint probability distribution can be written as:

$$\begin{aligned} &P(x_1, y_1, ..., x_n, y_n) \\ &\quad = P(y_1)P(x_1|y_1) \prod_{i=2}^{n} P(y_i|y_{i-1})P(x_i|y_i) \end{aligned} \tag{27}$$

Thus the complex computation could be solved by calculating $P(y_1)$, $P(y_i|y_{i-1})$ and $P(x_i|y_i)$, which are respectively determined by initial probability distribution π, state transition probability matrix A and output observation likelihood B.

Feriel et al. [96] use the Hidden Markov Model-Universal Background Model (HMM-UBM) model for continuous authentication based on the position, the accelerometer, and the gravity of the device. This method can achieve the Equal Error Rate (EER) value of 14.8% on HMOG, a public dataset.

4.3. Deep learning methods

Since traditional authentication models generally employ extracted features, which are insufficient to represent complicated typing patterns, deep learning methods are automatically applied

to learn time serials. Machine learning methods depend largely on extracted features, which need plenty of engineering work. Deep learning methods, on the contrary, could automatically extract features through Neural Network (NN).

The basic neural network structure is Multi-Layer Perceptron (MLP). Initial attempts at data through deep learning methods are usually based on MLP. Sequence-processing approaches commonly take a sequence of timing signals as input, which is suitable to address through Recurrent Neural Network (RNN) or Long Short Term Memory (LSTM). RNN/LSTM can preserve history information and maintain the dependency relationship in sequential data by calculating current output based on both current input and previous information [97]. Besides, data forms like gait images use the variation of LSTM, Convolutional LSTM (Conv-LSTM) to first extract image features through Convolutional Neural Network (CNN) [98]. Xiaofeng et al. [99] proposed a novel algorithm that extracted individual keystroke features of keystroke vector through CNN and RNN. Therefore, We will first briefly introduce the principle of neural networks and then introduce some recurrent neural networks that are intuitive choices for tackling sequential data.

MLP. MLP is a hierarchical structure containing multiple layers of an input layer, hidden layers, and output layer. Each layer consists of multiple inter-connected neurons controlled by connection weight. Consider input vector x_i combined with the corresponding connection weight w_i are passed to the neuron, the weighted sum is calculated and compared with threshold θ and then processed by an activation function $f(\cdot)$. Thus the neuron output can be represented as:

$$y = f(\sum_{i=1}^{n} w_i x_i - \theta) \tag{28}$$

After obtaining the model output through model forward, model parameters are learned through error back propagation based on the gradient descent strategy.

Adam et al. [100] utilize the deep neural network in an embedded ECG pattern recognition system for biometric authentication. Just taking the original ECG signals as input, this deep neural network-based system achieves 99.93% accuracy, 99.85% sensitivity, 99.96% specificity, an Equal Error Rate (EER) of 0.0582% on ECG-ID, a public database.

Chu-Hsing Lin et al. [101] use the deep neural network for biometric authentication based on keystroke dynamics. The result shows that the deep neural network can achieve 90% accuracy and 100% the false rejection rate (FRR) on their self-collected data.

CNN. Since MLP ignores the spatial relationship between elements in the raw data, CNN uses a local feature extractor in the signal to construct local spatial features. CNN is a hierarchical neural network with a series of convolutional layers, sub-sampling layers, and MLP classifier. The Convolutional layer utilizes several kernels to literately pass a sliding window over the data to get a feature map. As for sequential signals collected by wearable devices, the 1D version of CNN is commonly used for automatic classification.

Joao et al. [102] use CNN for ECG-based biometric authentication and achieves the 7.86% and 15.37% equal error rate (EER) on UofTDB and CYBHi datasets, respectively, and attained 9.06% EER on the PTB on-the-person database. Another work [103] using CNN combines ECG and fingerprint and achieves the 99.68% and 99.74% accuracy, 0.3% and 0.3% false acceptance rate, and 0.4% and 0.2% false rejection rate on MDB1 and MDB2 datasets, respectively.

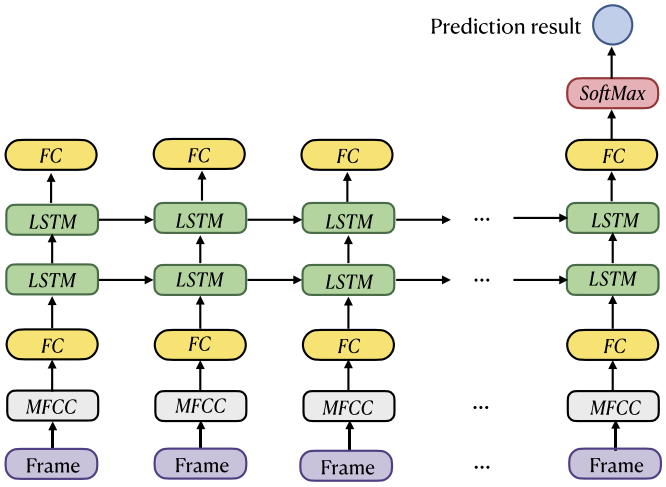

Fig. 10. Architecture of an end-to-end LSTM-based breath signal authentication method.

RNN. RNN is able to model sequential data through its internal state. Thus current input is considered to be related to the previous input and the history input information is preserved, making it suitable for processing continuous biometric signals. RNN performs same function for each input data and the current state h_t considers both current input x_t and the hidden state at last timestamp h_{t-1}:

$$h_t = f(h_{t-1}, x_t) \tag{29}$$

Ka-Wing Tse et al. [104] use RNN based the fusion of keystroke and swipe dynamics for biometric user authentication. By combining temporal features, spatial features, and swipe features together, RNN achieves the highest accuracy and F1 score value of 94.26% and 93.19%.

LSTM. LSTM is a modified version of RNN to solve the gradient vanishing and exploding problems in RNN. LSTM could better remember past information through the introduced gates. Input gate decides which input value should be taken to modify the memory, forget gate decides what information needs to be discarded, and the output gate decides the output given the input and the memory information in this block.

Farnaz et al. [105] use LSTM for biometric authentication for dementia patients. Two features based on PPG and ECG signals are provided to the LSTM network for training. Finally, this model achieves the accuracy of the PPG, and ECG-based identifications reached 100% and 88.9%, and F1 scores reached 1.00 and 0.86 respectively on a ten-users dataset.

We take an end-to-end LSTM-based breath signal authentication method as an example [106]. It aims to recognize whether a breath sample belongs to the user in the collected database. We depict Fig. 10 to illustrate the model architecture. The prediction model takes the segmented breath frame as input and outputs the prediction result of user identity. The MFCC features are first extracted from the data frame, and then a Fully Connected (FC) layer is performed to embed the initial features. The major part of the module is a two-layer LSTM. Each LSTm cell takes the input feature at the current time step and the previous hidden state at the last time step as input. The LSTM output is further embedded to a user authentication score through a FC layer, whose output size is the number of users. Finally, the authentication score is transformed into a probability through the SoftMax layer and output the final prediction result.

However, previously mentioned methods suffer inevitable shortcomings: template-matching methods could not retain temporal

Table 1
The evaluation result of the ECG signal authentication method.

	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
ECG Signal	PTB [109]	98.8	99.7	98.8	99.2
	CYBHi [109]	99.2	99.6	99.4	99.4

Table 2
The evaluation result of the gait signal authentication method.

	Dataset	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)
Gait Signal	UPCV	95.3	94.4	94.0	93.3
	GaitBiometry	98.1	98.0	98.3	97.8

information while sequence-processing methods require data to keep strict sequential order. Therefore, an unordered set-based gait recognition method, inspired by the promising method PointNet [107] in the field of computer vision, is then proposed. A sequential signal is considered as a set of independent frames [108], which is immune to the permutation of frames. It first transforms the frame-level features into set-level features and then learns the discriminative representations of each user's distribution.

5. Evaluation

Performance evaluation is an essential part of evaluating the proposed classification models through various kinds of metrics such as accuracy. Using only one metric is insufficient to judge a model. Thus several performance evaluation results should be analyzed simultaneously.

Accuracy. Classification accuracy is the most commonly used metric, it is the ratio of correct predictions number to the total samples number:

$$acc(f; D) = \frac{1}{m}\sum_{i=1}^{m} \mathbb{I}(f(x_i) = y_i) \tag{30}$$

While its opposite metric error rate is defined as:

$$E(f; D) = 1 - acc(f; D) \tag{31}$$

Confusion Matrix. Confusion matrix indicates the model complete performance, involving four important terms:

- True Positives (TP): cases where prediction is the same as ground truth of positive.
- True Negatives (TN): cases where prediction is the same as ground truth of negative.
- False Positives (FP): cases where prediction is positive but the ground truth is negative.
- False Negatives (FN): cases where prediction is negative but the ground truth is positive.

Precision. Precision is the ratio of the number of correctly predicted points to the number of all positive prediction results by the classifier:

$$P = \frac{TP}{TP + FP} \tag{32}$$

Recall. Recall is the ratio of the number of correctly predicted points to the number of all related points with positive label:

$$R = \frac{TP}{TP + FN} \tag{33}$$

F1 Score. F1 score is the Harmonic Mean between precision and recall, seeking the balance between both precision and robustness. A higher F1 score indicates a better model performance:

$$F1 = \frac{2PR}{P + R} \tag{34}$$

Precision and recall is a pair of opposite metrics, high precision often corresponds to low recall. Since we may have different emphasis in different scenarios, a more general F_β score is proposed:

$$F_\beta = \frac{(1 + \beta^2) \times P \times R}{(\beta^2 \times P) + R} \tag{35}$$

We respectively show the high-performance evaluation results of physiological biometric and behavioral biometrics based authentication methods as below. Since not all related references paper compare the metrics of accuracy, precision, recall, and F1 score, we first only list two tables containing the evaluation results of above four metrics in ECG signal and gait signal authentication methods as an example.

As for ECG signal illustrated in Table 1, a ResNet-Attention model [109] achieves better performance than other existing methods. It first extracted ECG features through ResNet and weighted different signals through attention mechanism. The experiment is conducted on two ECG datasets Physikalisch-Technische Bundesanstalt (PTB) and Check Your Bio-signals Here initiative (CYBHi), evaluating the metrics of accuracy, precision, recall, and F1 score.

As for the gait signal illustrated in Table 2, a deep neural network model with new geometric features [110] performs better than previous methods. The experiment is conducted on two public gait datasets recorded with the Microsoft Kinect sensor, UPCV gait dataset, and Kinect GaitBiometry dataset. The proposed model is evaluated through the metrics of accuracy, precision, recall, and F1 score.

Then we focus on the "Accuracy" metric and respectively demonstrate the state-of-the-art results of both physiological biometrics and behavioral biometrics in past five years. See Tables 3 and 4.

6. Conclusion

Wearable technology represents a hastily growing set of product categories and a dynamic and evolving field. From the security perspective, they offer a possibility to re-envision and adapt conventional authentication schemes to the brand new contexts and interactive methods of wearable gadgets. A large number of user authentication systems have been proposed and designed in the past decade. In this paper, we have reviewed and summarized recent solutions. We provided representative examples of authentication interfaces with a variety of wearable form factors such as a smartwatch, smart glass. While in this survey, we have reviewed more than 120 papers; the list of current structures is by no means comprehensive but includes much of the recent advancements and directions. We hope this survey can help researchers interested in this area identify research gaps and easily find research directions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Table 3
Accuracy result of Physiological Biometrics authentication methods.

Physiological Biometrics	No.	Author and Year	Methods	Dataset	Accuracy (%)
ECG (Heartbeat)	1	Tan et al., 2016 [111]	Wavelet Analysis, Probabilistic Random Forest	ECG-ID	98.79
				MIT-BIH	99.43
	2	Zhang, 2017 [112]	Multiresolution 1-D-CNN	MIT-BIH	93.5
	3	Salloum & Kuo, 2017 [113]	LSTM	ECG-ID	100
				MIT-BIH	100
	4	Bassiouoni et al., 2018 [114]	MLP, SVM	ECG-ID	99
				MIT-BIH	100
	5	Wang et al., 2019 [115]	PCANet, SVM	ECG-ID	97.75
				MIT-BIH	100
	6	Ihsanto et al., 2020 [116]	Residual Depthwise Separable CNN	ECG-ID	100
				MIT-BIH	100
Breath Signal	1	Chauhan et al., 2017 [28]	DCT Feature, Gaussian Mixtures Model	Off-the-shelf Hardware, 10 users, 7 days	94
	2	Lu et al., 2017 [29]	MFCC Feature, Template Matching	50 Smartphone Users	99.6
	3	Islam et al, 2019 [117]	FFT Feature, SVM	Doppler Rader System	100
	4	Leem et al., 2020 [118]	CNN, SVM	Ultra-wideband Radio Receiver	96.7
Muscle Signal	1	Shang & Wu, 2019 [119]	Local Outlier Factor model	Smartwatches	96.31
	2	Shin et al., 2021 [120]	Fiducial Features, SVM	Wearable Electromyogram System	87.1
	3	Chen et al., 2021 [121]	Variational Autoencoer, CNN	Electrical Muscle Stimulation System	97.6

Table 4
Accuracy result of Behavioral Biometrics authentication methods.

Behavioral Biometrics	No.	Author and Year	Methods	Dataset	Accuracy (%)
Gait Signal	1	Zhang et al. [122], 2014	Signature Points Feature, Sparse-code Collection Classifier	ZJU-GaitAcc Public Dataset	95.8
	2	Giorgi et al. [123], 2017	CNN	ZJU-GaitAcc Public Dataset	92
	3	Sun et al. [36], 2018	Adaptive Gait Cycle Extraction, Template Matching	ZJU-GaitAcc Public Dataset	91.75
	4	Sun et al. [124], 2019	Template Matching, Decision Fusion	OU_ISIR Public Dataset	96.7
	5	Qin et al. [125], 2019	LSTM, Extreme value Analysis	ZJU-GaitAcc Public Dataset	98.4
Signature Signal	1	Lai et al. [126], 2017	Length-Normalized Path Signature Feature, RNN	SVC2004 Task2	91
				MCYT-100	97.63
	2	Al-Hmouz et al. [127], 2019	Probabilistic Dynamic Time Warping	SVC2004 Task2	98.2
				MCYT-100	98.1
	3	Okawa, M [128], 2020	Mean Template Matching, Weighted DTW Distances	SVC2004 Task1	95.74
				SVC2004 Task2	98.2
				MCYT-100	98.72
Keystroke Signal	1	Kim et al. [61], 2018	User-adaptive Keystroke Feature, SVM	150 users, 13000 keystrokes	95.6
	2	Ali et al. [95], 2018	Partially Observable HMM, SVM	CMU Keystroke Dataset	91.4
	3	Tse & Hung [104], 2020	Mult-stream RNN, Feature Fusion	31 users, 50 times	94.26
Mouse Dynamics	1	Chong et al. [68], 2018	2D-CNN	Balabit	90
				TWOS	87
	2	Almalki et al. [129], 2019	Decision Tree	Point and Click Action Data	87.6
			K-Nearest Neighbors		99.3
			Random Forest		89.9
	3	Fu et al. [70], 2020	CNN-RNN Combined Model	15 users, 300 trials each	96.84

Acknowledgments

This work was supported in part by the Hong Kong RGC grant ECS 21212419 and ECS 21201420, the Guangdong Basic and Applied Basic Research Foundation under Key Project 2019B1515120032, the City University of Hong Kong SRG-Fd Grant 7005561, APRC grant 9610485, and Start-up grant 7200642. The work described in this paper was partially supported by a grant from Chow Sang Sang Group Research Fund sponsored by Chow Sang Sang Holdings International Limited (Project No. 9229062).

References

[1] Wearable technology market global opportunity analysis and industry forecast, http://www.prnewswire.com/news-releases/, 2017.

[2] W. Xu, C. Javali, G. Revadigar, C. Luo, N. Bergmann, W. Hu, Gait-key: a gait-based shared secret key generation protocol for wearable devices, ACM Trans. Sens. Netw. 13 (1) (2017) 1–27.

[3] I. Deutschmann, P. Nordström, L. Nilsson, Continuous authentication using behavioral biometrics, IT Prof. 15 (4) (2013) 12–15.

[4] J. Xu, L. Song, J.Y. Xu, G.J. Pottie, M. Van Der Schaar, Personalized active learning for activity classification using wireless wearable sensors, IEEE J. Sel. Top. Signal Process. 10 (5) (2016) 865–876.

[5] M.O. Derawi, Smartphones and biometrics: gait and activity recognition, 2012.

[6] T. Matsumoto, H. Matsumoto, K. Yamada, S. Hoshino, Impact of artificial "gummy" fingers on fingerprint systems, in: Optical Security and Counterfeit Deterrence Techniques IV, vol. 4677, International Society for Optics and Photonics, 2002, pp. 275–289.

[7] J. Galbally, S. Marcel, J. Fierrez, Image quality assessment for fake biometric detection: application to iris, fingerprint, and face recognition, IEEE Trans. Image Process. 23 (2) (2013) 710–724.

[8] R. Tolosana, M. Gomez-Barrero, C. Busch, J. Ortega-Garcia, Biometric presentation attack detection: beyond the visible spectrum, IEEE Trans. Inf. Forensics Secur. 15 (2019) 1261–1275.
[9] T. Chugh, A.K. Jain, Fingerprint presentation attack detection: generalization and efficiency, in: 2019 International Conference on Biometrics, ICB, IEEE, 2019, pp. 1–8.
[10] R. Gonçalves Pires, A. Nilceu Marana, J. Paulo Papa, et al., Deep features extraction for robust fingerprint spoofing attack detection, J. Artif. Intell. Soft Comput. Res. 9 (2019).
[11] A. George, Z. Mostaani, D. Geissenbuhler, O. Nikisins, A. Anjos, S. Marcel, Biometric face presentation attack detection with multi-channel convolutional neural network, IEEE Trans. Inf. Forensics Secur. 15 (2019) 42–55.
[12] J.C. Neves, R. Tolosana, R. Vera-Rodriguez, V. Lopes, H. Proença, J. Fierrez, Ganprintr: improved fakes and evaluation of the state of the art in face manipulation detection, IEEE J. Sel. Top. Signal Process. 14 (5) (2020) 1038–1048.
[13] A.N. Kataria, D.M. Adhyaru, A.K. Sharma, T.H. Zaveri, A survey of automated biometric authentication techniques, in: 2013 Nirma University International Conference on Engineering, NUiCONE, IEEE, 2013, pp. 1–6.
[14] N. Yusuf, K.A. Marafa, K.L. Shehu, H. Mamman, M. Maidawa, A survey of biometric approaches of authentication, Int. J. Adv. Comput. Res. 10 (47) (2020) 96–104.
[15] A. Mahfouz, T.M. Mahmoud, A.S. Eldin, A survey on behavioral biometric authentication on smartphones, J. Inf. Secur. Appl. 37 (2017) 28–37.
[16] Z. Rui, Z. Yan, A survey on biometric authentication: toward secure and privacy-preserving identification, IEEE Access 7 (2018) 5994–6009.
[17] S. Bhatt, T. Santhanam, Keystroke dynamics for biometric authentication—a survey, in: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, IEEE, 2013, pp. 17–23.
[18] S.P. Banerjee, D.L. Woodard, Biometric authentication and identification using keystroke dynamics: a survey, J. Pattern Recognit. Res. 7 (1) (2012) 116–139.
[19] V. Chandrashekhar, P. Singh, M. Paralkar, O.K. Tonguz, Pulse id: the case for robustness of ecg as a biometric identifier, in: 2020 IEEE 30th International Workshop on Machine Learning for Signal Processing, MLSP, IEEE, 2020, pp. 1–6.
[20] M. Abdar, W. Ksiażek, U.R. Acharya, R.-S. Tan, V. Makarenkov, P. Pławiak, A new machine learning technique for an accurate diagnosis of coronary artery disease, Comput. Methods Programs Biomed. 179 (2019) 104992.
[21] U.R. Acharya, H. Fujita, S.L. Oh, Y. Hagiwara, J.H. Tan, M. Adam, R. San Tan, Deep convolutional neural network for the automated diagnosis of congestive heart failure using ECG signals, Appl. Intell. 49 (1) (2019) 16–27.
[22] K. Georgiou, A.V. Larentzakis, N.N. Khamis, G.I. Alsuhaibani, Y.A. Alaska, E.J. Giallafos, Can wearable devices accurately measure heart rate variability? A systematic review, Folia Med. 60 (1) (2018) 7–20.
[23] H.-Y. Jan, M.-F. Chen, T.-C. Fu, W.-C. Lin, C.-L. Tsai, K.-P. Lin, Evaluation of coherence between ECG and PPG derived parameters on heart rate variability and respiration in healthy volunteers with/without controlled breathing, J. Med. Biol. Eng. 39 (5) (2019) 783–795.
[24] D.B. Saadi, G. Tanev, M. Flintrup, A. Osmanagic, K. Egstrup, K. Hoppe, P. Jennum, J.L. Jeppesen, H.K. Iversen, H.B. Sorensen, Automatic real-time embedded qrs complex detection for a novel patch-type electrocardiogram recorder, IEEE J. Transl. Eng. Heal. Medicine 3 (2015) 1–12.
[25] F. Sun, C. Yi, W. Li, Y. Li, A wearable h-shirt for exercise ECG monitoring and individual lactate threshold computing, Comput. Ind. 92 (2017) 1–11.
[26] C. Beach, S. Krachunov, J. Pope, X. Fafoutis, R.J. Piechocki, I. Craddock, A.J. Casson, An ultra low power personalizable wrist worn ECG monitor integrated with iot infrastructure, IEEE Access 6 (2018) 44010–44021.
[27] M. Sarkar, I. Madabhavi, N. Niranjan, M. Dogra, Auscultation of the respiratory system, Ann. Thorac. Medicine 10 (3) (2015) 158.
[28] J. Chauhan, Y. Hu, S. Seneviratne, A. Misra, A. Seneviratne, Y. Lee, Breathprint: breathing acoustics-based user authentication, in: Proceedings of the 15th Annual International Conference on Mobile Systems, Applications, and Services, 2017, pp. 278–291.
[29] L. Lu, L. Liu, M.J. Hussain, Y. Liu, I sense you by breath: speaker recognition via breath biometrics, IEEE Trans. Dependable Secure Comput. (2017).
[30] D. Ruinskiy, Y. Lavner, An effective algorithm for automatic detection and exact demarcation of breath sounds in speech and song signals, IEEE Trans. Audio Speech Lang. Process. 15 (3) (2007) 838–850.
[31] S.H. Dumpala, K.R. Alluri, An algorithm for detection of breath sounds in spontaneous speech with application to speaker recognition, in: International Conference on Speech and Computer, Springer, 2017, pp. 98–108.
[32] X. Jiang, K. Xu, X. Liu, C. Dai, D.A. Clifton, E.A. Clancy, M. Akay, W. Chen, Neuromuscular password-based user authentication, IEEE Trans. Ind. Inform. 17 (4) (2021) 2641–2652.
[33] User authentication via electrical muscle stimulation [EB/OL], http://sandlab.cs.uchicago.edu/electricauth/electricauth.pdf.
[34] J. Frank, S. Mannor, D. Precup, Activity and gait recognition with time-delay embeddings, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 24, no. 1, 2010.
[35] M.O. Derawi, C. Nickel, P. Bours, C. Busch, Unobtrusive user-authentication on mobile phones using biometric gait recognition, in: 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing, IEEE, 2010, pp. 306–311.
[36] F. Sun, C. Mao, X. Fan, Y. Li, Accelerometer-based speed-adaptive gait authentication method for wearable iot devices, IEEE Int. Things J. 6 (1) (2018) 820–830.
[37] Q. Zou, Y. Wang, Q. Wang, Y. Zhao, Q. Li, Deep learning-based gait recognition using smartphones in the wild, IEEE Trans. Inf. Forensics Secur. 15 (2020) 3197–3212.
[38] W. Xu, Y. Shen, C. Luo, J. Li, W. Li, A.Y. Zomaya, Gait-watch: a gait-based context-aware authentication system for smart watch via sparse coding, Ad Hoc Netw. 107 (2020) 102218.
[39] C. Luo, J. Wu, J. Li, J. Wang, W. Xu, Z. Ming, B. Wei, W. Li, A.Y. Zomaya, Gait recognition as a service for unobtrusive user identification in smart spaces, ACM? Trans. Internet Thing 1 (1) (2020) 1–21.
[40] Y. Shen, H. Wen, C. Luo, W. Xu, T. Zhang, W. Hu, D. Rus, Gaitlock: protect virtual and augmented reality headsets using gait, IEEE Trans. Dependable Secure Comput. 16 (3) (2018) 484–497.
[41] W. Xu, G. Lan, Q. Lin, S. Khalifa, M. Hassan, N. Bergmann, W. Hu, Keh-gait: using kinetic energy harvesting for gait-based user authentication systems, IEEE Trans. Mob. Comput. 18 (1) (2018) 139–152.
[42] W. Xu, G. Lan, Q. Lin, S. Khalifa, N. Bergmann, M. Hassan, W. Hu, Keh-gait: towards a mobile healthcare user authentication system by kinetic energy harvesting, in: NDSS, no. 0.2, 2017, pp. 0–4.
[43] G. Yang, W. Tan, H. Jin, T. Zhao, L. Tu, Review wearable sensing system for gait recognition, Clust. Comput. 22 (2) (2019) 3021–3029.
[44] T. Liu, Y. Inoue, K. Shibata, A wearable ground reaction force sensor system and its application to the measurement of extrinsic gait variability, Sensors 10 (11) (2010) 10240–10255.
[45] J. Yang, Y. Li, M. Xie, Motionauth: motion-based authentication for wrist worn smart devices, in: 2015 IEEE International Conference on Pervasive Computing and Communication Workshops, PerCom Workshops, IEEE, 2015, pp. 550–555.
[46] J. Bromley, J.W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Säckinger, R. Shah, Signature verification using a "Siamese" time delay neural network, Int. J. Pattern Recognit. Artif. Intell. 7 (04) (1993) 669–688.
[47] G. Li, H. Sato, Handwritten signature authentication using smartwatch motion sensors, in: 2020 IEEE 44th Annual Computers, Software, and Applications Conference, COMPSAC, IEEE, 2020, pp. 1589–1596.
[48] B. Nassi, A. Levy, Y. Elovici, E. Shmueli, Handwritten signature verification using hand-worn devices, preprint, arXiv:1612.06305, 2016.
[49] Q. Lyu, Z. Kong, C. Shen, T. Yue, Wristauthen: a dynamic time wrapping approach for user authentication by hand-interaction through wrist-worn devices, preprint, arXiv:1710.07941, 2017.
[50] W. Xu, J. Tian, Y. Cao, S. Wang, Challenge-response authentication using in-air handwriting style verification, IEEE Trans. Dependable Secure Comput. 17 (1) (2017) 51–64.
[51] C. Huang, Z. Yang, H. Chen, Q. Zhang, Signing in the air w/o constraints: robust gesture-based authentication for wrist wearables, in: GLOBECOM 2017-2017 IEEE Global Communications Conference, IEEE, 2017, pp. 1–6.
[52] A. Levy, B. Nassi, Y. Elovici, E. Shmueli, Handwritten signature verification using wrist-worn devices, in: Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, vol. 2, no. 3, 2018, pp. 1–26.
[53] R. Giot, B. Dorizzi, C. Rosenberger, A review on the public benchmark databases for static keystroke dynamics, Comput. Secur. 55 (2015) 46–61.
[54] R. Giot, A. Rocha, Siamese networks for static keystroke dynamics authentication, in: 2019 IEEE International Workshop on Information Forensics and Security, WIFS, IEEE, 2019, pp. 1–6.
[55] Y. Sun, H. Ceker, S. Upadhyaya, Shared keystroke dataset for continuous authentication, in: 2016 IEEE International Workshop on Information Forensics and Security, WIFS, IEEE, 2016, pp. 1–6.
[56] H. Ceker, S. Upadhyaya, Enhanced recognition of keystroke dynamics using gaussian mixture models, in: MILCOM 2015-2015 IEEE Military Communications Conference, IEEE, 2015, pp. 1305–1310.
[57] E. Vural, J. Huang, D. Hou, S. Schuckers, Shared research dataset to support development of keystroke authentication, in: IEEE International Joint Conference on Biometrics, IEEE, 2014, pp. 1–8.
[58] D. Gunetti, C. Picardi, Keystroke analysis of free text, ACM Trans. Inf. Syst. Secur. 8 (3) (2005) 312–347.
[59] C. Murphy, J. Huang, D. Hou, S. Schuckers, Shared dataset on natural human-computer interaction to support continuous authentication research, in: 2017 IEEE International Joint Conference on Biometrics, IJCB, IEEE, 2017, pp. 525–530.
[60] J. Huang, D. Hou, S. Schuckers, T. Law, A. Sherwin, Benchmarking keystroke authentication algorithms, in: 2017 IEEE Workshop on Information Forensics and Security, WIFS, IEEE, 2017, pp. 1–6.
[61] J. Kim, H. Kim, P. Kang, Keystroke dynamics-based user authentication using freely typed text based on user-adaptive feature extraction and novelty detection, Appl. Soft Comput. 62 (2018) 1077–1087.
[62] L. Wang, X. Geng, Behavioral Biometrics for Human Identification: Intelligent Applications: Intelligent Applications, IGI Global, 2009.
[63] S. Mondal, P. Bours, A computational approach to the continuous authentication biometric system, Inf. Sci. 304 (2015) 28–53.

[64] C. Shen, Z. Cai, X. Guan, Y. Du, R.A. Maxion, User authentication through mouse dynamics, IEEE Trans. Inf. Forensics Secur. 8 (1) (2012) 16–30.
[65] C. Shen, Z. Cai, X. Guan, R. Maxion, Performance evaluation of anomaly-detection algorithms for mouse dynamics, Comput. Secur. 45 (2014) 156–171.
[66] P. Kasprowski, K. Harezlak, Fusion of eye movement and mouse dynamics for reliable behavioral biometrics, Pattern Anal. Appl. 21 (1) (2018) 91–103.
[67] S. Mondal, P. Bours, Combining keystroke and mouse dynamics for continuous user authentication and identification, in: 2016 IEEE International Conference on Identity, Security and Behavior Analysis, ISBA, IEEE, 2016, pp. 1–8.
[68] P. Chong, Y.X.M. Tan, J. Guarnizo, Y. Elovici, A. Binder, Mouse authentication without the temporal aspect–what does a 2d-cnn learn?, in: 2018 IEEE Security and Privacy Workshops, SPW, IEEE, 2018, pp. 15–21.
[69] P. Chong, Y. Elovici, A. Binder, User authentication based on mouse dynamics using deep neural networks: a comprehensive study, IEEE Trans. Inf. Forensics Secur. 15 (2019) 1086–1101.
[70] S. Fu, D. Qin, D. Qiao, G.T. Amariucai, Rumba-mouse: rapid user mouse-behavior authentication using a cnn-rnn approach, in: 2020 IEEE Conference on Communications and Network Security, CNS, IEEE, 2020, pp. 1–9.
[71] S.A.A. Yusuf, R. Hidayat, MFCC feature extraction and KNN classification in ECG signals, in: 2019 6th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE, IEEE, 2019, pp. 1–5.
[72] V. Vimala, K. Ramar, M. Ettappan, An intelligent sleep apnea classification system based on EEG signals, J. Med. Syst. 43 (2) (2019) 36.
[73] K. Prajna, C. Mukhopadhyay, Fractional Fourier transform based adaptive filtering techniques for acoustic emission signal enhancement, J. Nondestruct. Eval. 39 (1) (2020) 1–15.
[74] S. Aziz, M.U. Khan, Z.A. Choudhry, A. Aymin, A. Usman, ECG-based biometric authentication using empirical mode decomposition and support vector machines, in: 2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference, IEMCON, IEEE, 2019, pp. 0906–0912.
[75] M. Jalil, F.A. Butt, A. Malik, Short-time energy, magnitude, zero crossing rate and autocorrelation measurement for discriminating voiced and unvoiced segments of speech signals, in: 2013 the International Conference on Technological Advances in Electrical, Electronics and Computer Engineering, TAEECE, IEEE, 2013, pp. 208–212.
[76] A. Abushakra, M. Faezipour, Acoustic signal classification of breathing movements to virtually aid breath regulation, IEEE J. Biomed. Health Inform. 17 (2) (2013) 493–500.
[77] C. Nickel, H. Brandt, C. Busch, Classification of acceleration data for biometric gait recognition on mobile devices, in: BIOSIG 2011–Proceedings of the Biometrics Special Interest Group, 2011.
[78] M. Pleva, E. Kiktova, J. Juhar, P. Bours, Acoustical user identification based on mfcc analysis of keystrokes, Adv. Electr. Electron. Eng. 13 (4) (2015) 309–313.
[79] N. Baranwal, G.C. Nandi, An efficient gesture based humanoid learning using wavelet descriptor and MFCC techniques, Int. J. Mach. Learn. Cybern. 8 (4) (2017) 1369–1388.
[80] M. Boussaa, I. Atouf, M. Atibi, A. Bennis, ECG signals classification using MFCC coefficients and ANN classifier, in: 2016 International Conference on Electrical and Information Technologies, ICEIT, IEEE, 2016, pp. 480–484.
[81] C. Mao, Y. Li, F. Sun, Accelerometer-based gait recognition using PCA & LDA algorithms, in: 2018 IEEE 23rd International Conference on Digital Signal Processing, DSP, IEEE, 2018, pp. 1–4.
[82] X. Xing, K. Wang, T. Yan, Z. Lv, Complete canonical correlation analysis with application to multi-view gait recognition, Pattern Recognit. 50 (2016) 107–117.
[83] W. Xu, C. Luo, A. Ji, C. Zhu, Coupled locality preserving projections for cross-view gait recognition, Neurocomputing 224 (2017) 37–44.
[84] G.R. Naik, S.E. Selvan, S.P. Arjunan, A. Acharyya, D.K. Kumar, A. Ramanujam, H.T. Nguyen, An ICA-EBM-based SEMG classifier for recognizing lower limb movements in individuals with and without knee pathology, IEEE Trans. Neural Syst. Rehabil. Eng. 26 (3) (2018) 675–686.
[85] M. Sharif, M. Attique, M.Z. Tahir, M. Yasmim, T. Saba, U.J. Tanik, A machine learning method with threshold based parallel feature fusion and feature selection for automated gait recognition, J. Organ. End. User Comput. 32 (2) (2020) 67–92.
[86] Y. He, J. Zhang, H. Shan, L. Wang, Multi-task gans for view-specific feature learning in gait recognition, IEEE Trans. Inf. Forensics Secur. 14 (1) (2018) 102–113.
[87] X. Bai, Y. Hui, L. Wang, F. Zhou, Radar-based human gait recognition using dual-channel deep convolutional neural network, IEEE Trans. Geosci. Remote Sens. 57 (12) (2019) 9767–9778.
[88] H. Wang, Y. Fan, B. Fang, S. Dai, Generalized linear discriminant analysis based on euclidean norm for gait recognition, Int. J. Mach. Learn. Cybern. 9 (4) (2018) 569–576.
[89] H. Zhang, Z. Liu, H. Zhao, Gait modeling and identifying based on dynamic template matching, J. Comput. Inf. Syst. 7 (4) (2011) 1155–1162.
[90] M. De Marsico, A. Mecca, Biometric walk recognizer, Multimed. Tools Appl. 76 (4) (2017) 4713–4745.
[91] A.Y. Shdefat, M.-I. Joo, S.-H. Choi, K. Hee-Cheol, Utilizing ECG waveform features as new biometric authentication method, Int. J. Electr. Comput. Eng. 8 (2) (2018) 658.
[92] C. Will, K. Shi, R. Weigel, A. Koelpin, Advanced template matching algorithm for instantaneous heartbeat detection using continuous wave radar systems, in: 2017 First IEEE MTT-S International Microwave Bio Conference, IMBIOC, IEEE, 2017, pp. 1–4.
[93] A. Santos, I. Medeiros, P. Resque, D. Rosário, M. Nogueira, A. Santos, E. Cerqueira, K.R. Chowdhury, ECG-based user authentication and identification method on vanets, in: Proceedings of the 10th Latin America Networking Conference, 2018, pp. 119–122.
[94] S. Hadiyoso, S. Aulia, A. Rizal, One-lead electrocardiogram for biometric authentication using time series analysis and support vector machine, Int. J. Adv. Comput. Sci. Appl. 10 (2) (2019) 276–283.
[95] M.L. Ali, C.C. Tappert, Pohmm/svm: a hybrid approach for keystroke biometric user authentication, in: 2018 IEEE International Conference on Real-Time Computing and Robotics, RCAR, IEEE, 2018, pp. 612–617.
[96] F. Cherifi, M. Omar, K. Amroun, An efficient biometric-based continuous authentication scheme with HMM prehensile movements modeling, J. Inf. Secur. Appl. 57 (2021) 102739.
[97] H.X. Tan, N.N. Aung, J. Tian, M.C.H. Chua, Y.O. Yang, Time series classification using a modified LSTM approach from accelerometer-based data: a comparative study for gait cycle detection, Gait Posture 74 (2019) 128–134.
[98] X. Wang, W.Q. Yan, Human gait recognition based on frame-by-frame gait energy images and convolutional long short-term memory, Int. J. Neural Syst. 30 (01) (2020) 1950027.
[99] L. Xiaofeng, Z. Shengfei, Y. Shengwei, Continuous authentication by free-text keystroke based on CNN plus RNN, Proc. Comput. Sci. 147 (2019) 314–318.
[100] A. Page, A. Kulkarni, T. Mohsenin, Utilizing deep neural nets for an embedded ecg-based biometric authentication system, in: 2015 IEEE Biomedical Circuits and Systems Conference, BioCAS, IEEE, 2015, pp. 1–4.
[101] C.-H. Lin, J.-C. Liu, K.-Y. Lee, On neural networks for biometric authentication based on keystroke dynamics, Sens. Mater. 30 (3) (2018) 385–396.
[102] J.R. Pinto, J.S. Cardoso, An end-to-end convolutional neural network for ECG-based biometric authentication, in: 2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems, BTAS, IEEE, 2019, pp. 1–8.
[103] M. Hammad, Y. Liu, K. Wang, Multimodal biometric authentication systems using convolution neural network based on different level fusion of ECG and fingerprint, IEEE Access 7 (2018) 26527–26542.
[104] K.-W. Tse, K. Hung, User behavioral biometrics identification on mobile platform using multimodal fusion of keystroke and swipe dynamics and recurrent neural network, in: 2020 IEEE 10th Symposium on Computer Applications & Industrial Electronics, ISCAIE, IEEE, 2020, pp. 262–267.
[105] F. Farid, F. Ahamed, Biometric authentication for dementia patients with recurrent neural network, in: 2019 International Conference on Electrical Engineering Research & Practice, ICEERP, IEEE, 2019, pp. 1–6.
[106] J. Chauhan, J. Rajasegaran, S. Seneviratne, A. Misra, A. Seneviratne, Y. Lee, Performance characterization of deep learning models for breathing-based authentication on resource-constrained devices, Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol. 2 (4) (2018) 1–24.
[107] C.R. Qi, H. Su, K. Mo, L.J. Guibas, Pointnet: deep learning on point sets for 3d classification and segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 652–660.
[108] H. Chao, Y. He, J. Zhang, J. Feng, Gaitset: regarding gait as a set for cross-view gait recognition, in: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 33, 2019, pp. 8126–8133.
[109] M. Hammad, P. Pławiak, K. Wang, U.R. Acharya, Resnet-attention model for human authentication using ECG signals, Expert Syst. (2020) e12547.
[110] A.H. Bari, M.L. Gavrilova, Artificial neural network based gait recognition using kinect sensor, IEEE Access 7 (2019) 162708–162722.
[111] R. Tan, M. Perkowski, ECG biometric identification using wavelet analysis coupled with probabilistic random forest, in: 2016 15th IEEE International Conference on Machine Learning and Applications, ICMLA, IEEE, 2016, pp. 182–187.
[112] Q. Zhang, D. Zhou, X. Zeng, Heartid: a multiresolution convolutional neural network for ECG-based biometric human identification in smart health applications, IEEE Access 5 (2017) 11805–11816.
[113] R. Salloum, C.-C.J. Kuo, ECG-based biometrics using recurrent neural networks, in: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, IEEE, 2017, pp. 2062–2066.
[114] M.M. Bassiouni, E.-S.A. El-Dahshan, W. Khalefa, A.M. Salem, Intelligent hybrid approaches for human ECG signals identification, Signal Image Video Process. 12 (5) (2018) 941–949.
[115] D. Wang, Y. Si, W. Yang, G. Zhang, T. Liu, A novel heart rate robust method for short-term electrocardiogram biometric identification, Appl. Sci. 9 (1) (2019) 201.
[116] E. Ihsanto, K. Ramli, D. Sudiana, T.S. Gunawan, Fast and accurate algorithm for ecg authentication using residual depthwise separable convolutional neural networks, Appl. Sci. 10 (9) (2020) 3304.
[117] S.M. Islam, A. Rahman, N. Prasad, O. Boric-Lubecke, V.M. Lubecke, Identity authentication system using a support vector machine (svm) on radar respiration measurements, in: 2019 93rd ARFTG Microwave Measurement Conference, ARFTG, IEEE, 2019, pp. 1–5.

[118] S.K. Leem, F. Khan, S.H. Cho, Remote authentication using an ultra-wideband radio frequency transceiver, in: 2020 IEEE 17th Annual Consumer Communications & Networking Conference, CCNC, IEEE, 2020, pp. 1–8.
[119] J. Shang, J. Wu, A usable authentication system using wrist-worn photoplethysmography sensors on smartwatches, in: 2019 IEEE Conference on Communications and Network Security, CNS, IEEE, 2019, pp. 1–9.
[120] S. Shin, M. Kang, J. Jung, Y.T. Kim, Development of miniaturized wearable wristband type surface EMG measurement system for biometric authentication, Electronics 10 (8) (2021) 923.
[121] Y. Chen, Z. Yang, R. Abbou, P. Lopes, B.Y. Zhao, H. Zheng, User authentication via electrical muscle stimulation, 2021.
[122] Y. Zhang, G. Pan, K. Jia, M. Lu, Y. Wang, Z. Wu, Accelerometer-based gait recognition by sparse representation of signature points with clusters, IEEE Trans. Cybern. 45 (9) (2014) 1864–1875.
[123] G. Giorgi, F. Martinelli, A. Saracino, M. Sheikhalishahi, Try walking in my shoes, if you can: accurate gait recognition through deep learning, in: International Conference on Computer Safety, Reliability, and Security, Springer, 2017, pp. 384–395.
[124] F. Sun, W. Zang, R. Gravina, G. Fortino, Y. Li, Gait-based identification for elderly users in wearable healthcare systems, Inf. Fusion 53 (2020) 134–144.
[125] Z. Qin, G. Huang, H. Xiong, Z. Qin, K.-K.R. Choo, A fuzzy authentication system based on neural network learning and extreme value statistics, IEEE Trans. Fuzzy Syst. (2019).
[126] S. Lai, L. Jin, W. Yang, Online signature verification using recurrent neural network and length-normalized path signature descriptor, in: 2017 14th IAPR International Conference on Document Analysis and Recognition, ICDAR, vol. 1, IEEE, 2017, pp. 400–405.
[127] R. Al-Hmouz, W. Pedrycz, K. Daqrouq, A. Morfeq, A. Al-Hmouz, Quantifying dynamic time warping distance using probabilistic model in verification of dynamic signatures, Soft Comput. 23 (2) (2019) 407–418.
[128] M. Okawa, Online signature verification using single-template matching with time-series averaging and gradient boosting, Pattern Recognit. 102 (2020) 107227.
[129] S. Almalki, P. Chatterjee, K. Roy, Continuous authentication using mouse clickstream data analysis, in: International Conference on Security, Privacy and Anonymity in Computation, Communication and Storage, Springer, 2019, pp. 76–85.

Shuqi Liu received the B.S. degree and the M.S. degree from the Northeastern University, Shenyang, China, in 2017 and 2020. She is currently working toward the Ph.D. degree at the Department of Computer Science, City University of Hong Kong. Her research interests lie in the machine learning, data science and natural language processing.

Wei Shao received the B.S. degree from the Peking University, Beijing, China, in 2020. He is currently working toward the Ph.D. degree at the Department of Computer Science, City University of Hong Kong. His research interests include the machine learning, natural language processing and information theory.

Tan Li received the B.S. degree from the Central South University, Changsha, China, in 2016, and the M.S. degree from the University of Science and Technology of China, Hefei, China, in 2019. She is currently working toward the Ph.D. degree at the Department of Computer Science, City University of Hong Kong. Her research interests lie in the edge computing, distributed computing and machine learning for wireless communication.

Weitao Xu is an Assistant Professor at the Department of Computer Science at City University of Hong Kong. Before that, he was a Postdoctoral Research Associate at the School of Computer Science and Engineering (CSE) at UNSW from June 2017 to August 2019. He obtained his PhD degree from the University of Queensland in 2017 (advised by Prof. Neil Bergmann and Dr. Wen Hu). He received his B.E. degree in Communication Engineering and M.E. degree in Communication and Information System (advised by Prof. Dongfeng Yuan) both from the School of Information Science and Engineering, Shandong University (SDU), China, in 2010 and 2013, respectively. His research areas include mobile computing, sensor network and IoT.

Linqi Song is currently an Assistant Professor with the Department of Computer Science, City University of Hong Kong. He received the Ph.D. degree in electrical engineering from the University of California, Los Angeles (UCLA), USA and the B.S. and M.S. degrees from Tsinghua University, China. He was a Postdoctoral Scholar with the Department of Electrical and Computer Engineering, UCLA. His research interests include information theory, machine learning, and big data. He has received the Hong Kong RGC Early Career Scheme, in 2019, and the Best Paper Award from IEEE MIPR 2020.