

A survey on gait recognition in IoT applications

Wang-Yang Yu¹, Shui-Hua Wang^{1,*}, Yu-Dong Zhang^{1,†}

¹School of Computing and Mathematical Sciences, University of Leicester, Leicester, LE1 7RH, UK

Abstract

In IoT applications, identity recognition is a basic and critical requirement. In recent years, IoT technology has developed rapidly, and IoT devices such as wearable devices, environmental sensors, and WiFi devices have been popularized and developed. The unobtrusive, low-cost, continuous advanced identity recognition methods are needed in IoT applications. Gait recognition not only has high performance but also is not easy to be forged or hidden. It has excellent potential in IoT intelligent identity recognition. This review discusses the implementation of gait analysis, representative datasets, and algorithms. Finally, we also discussed the challenges of gait analysis in IoT applications.

Received on 08 March 2022; accepted on 26 March 2022; published on 12 April 2022

Keywords: Gait recognition, IoT, Pervasive Computing, WiFi, Sensor

Copyright © 2022 Wang-Yang Yu *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eetiot.v7i28.446

1. Introduction

Pervasive Computing was proposed [1] as early as 1991, emphasizing the integration of the physical and digital worlds. In Pervasive Computing, one of the most basic needs is identity recognition. However, due to the limitation of perception ability, this technology has not been applied and developed. In recent years, with the rapid development of IoT (Internet of Things) technology based on comprehensive perception [2, 3], identity recognition has obtained new development opportunities. With the great development of perception ability, we can expect better performance such as unobtrusive, cost-effective or continuous sensing to meet up the specific requirement of IoT application.

IoT emphasizes the interconnection of all things and pursues perception-based computing, communication, and control to integrate the physical and digital worlds [4]. In recent years, the rapid popularity of Internet of things devices such as cameras, smartphones, wearables, WiFi device etc., has laid the foundation for a privacy-safe, accurate and unobtrusive identity detection methods.

The IoT has changed people's lives greatly. A typical example of IoT application is smart spaces [5], which integrates perception technology into the environment

and comprehensively senses the target information to provide more accurate and personalized services. On the basis of IoT infrastructure and smart spaces, smart homes, smart buildings, smart cities, and other applications can be achieved. The sketch map of gait recognition is shown in Figure 1.

This widespread use makes user identification accuracy more critical because identity recognition is often related to authentication [6]. At present, there are many solutions for user identification, which are mainly divided into four categories: object-based, knowledge-based and physiological biometric-based, and behavioural biometric-based solutions [7, 8]. The object-based method uses a particular object for identification, such as the keys or the access card to open the door. The knowledge-based method requires users to enter the password, PIN (Personal Identification Number) or graphic PIN, to complete identity confirmation. This kind of method is based on the specific knowledge of people. This method can also be combined with object-based methods. For example, a credit card needs to use both a bank card and pin code. The process based on biological information completes the identification by collecting the user's physical characteristics (e.g. facial images, fingerprint, voiceprint and iris). The behaviour biometric-based method recognizes the user's gait information, keypress and other behaviours. The latter two ways need to be classified by machine learning or deep learning.

*Corresponding author. Email: shuihuawang@ieee.org

†Corresponding author. Email: yudongzhang@ieee.org

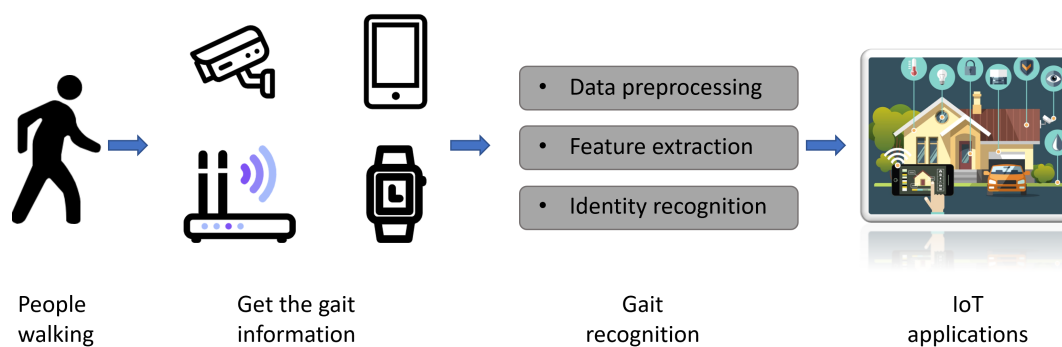


Figure 1. Sketch map of gait recognition

In a specific application scenario, it is a many-to-many mapping task, which requires an authentication scheme with high complexity and high robustness [9]. However, the present challenge is that gait recognition is affected by many factors, such as different walking speeds [25–27], different clothes [30], different sensor viewing points [28, 29, 31] and etc. [18, 32], which will make the recognition process more difficult and greatly affect the recognition results.

Although Many recognition methods have achieved good user recognition accuracy, there are still deficiencies in specific IoT applications [10]. The password and PIN can be obtained through various attacks. For example, shoulder surfing and recording attacks [11–13], thermal attack [14, 15] and smudge attacks [16, 17]. Also, it has managing problems since people can only remember a few passwords. A survey [18] shows that 49% of the users save their passwords somewhere, and 67% of the users never change their passwords. Image recognition can be passed by using face photos. Fingerprints can be conquered by using fake fingerprints made by conductive printing or silicon like glue [19] or by using the method of deep learning [20]. Voice-based identification is also challenged by many malicious spoofing attacks [21, 22].

Existing methods require users to cooperate actively, which are obtrusive. For example, the user needs to walk to a specific location and look at the camera in the facial identification, which is unnatural. We should reduce the obtrusiveness in recognition and complete the recognition without the user's extra attention. The image-based method and the voice-based method can not protect users' privacy, and the user's information may be leaked during the transmission process. Besides, the recognition method based on image and video has high recognition cost and needs a lot of computing resources.

Gait recognition is an emerging technology that uses people's walking patterns for recognition. It is very suitable to solve the problem of identity recognition in IoT applications. Firstly, the most

prominent feature of gait recognition is that it can achieve unobtrusive measurement [23] because it use sensors embedded in ambient environment which can minimize the disturbance in recognition and give the users the best experience [24]. People can be recognized without cooperative action or even imperceptible. Gait recognition has a long recognition distance. Even when it is far from the user, gait has good accuracy when the resolution of the target is low. Compared with other methods, because gait recognition is based on dynamic biological information rather than static, it is more difficult to attack [11–17] and will not be forgotten, stolen, hidden or forged. Wireless-based gait recognition have excellent privacy because it does not transmit biological information. In real complex scenes, identification should not be a one-off task, and continuous recognition has become a basic requirement, which can also be achieved by sensor-based gait recognition. Gait recognition can use IoT devices and public devices (WiFi, camera), so it is cost-effective and easy to deploy. These characteristics make gait recognition a competitive method to solve identity recognition in IoT.

At present, there have been many reviews in the field of gait analysis, including wireless-based [74], vision-based [68, 77] and sensor-based [75]. There is also a review of identity authentication in IoT. In this reviews, Liang et al. claimed that IoT authentication need to be continuous, unobtrusive, and cost-effective and discussed the challenges in the future work. However, this review focused on behavioral biometrics information rather than just gait information. Also, there have been new breakthroughs in the field of gait recognition in recent years. At present, there is no review just focusing on IoT gait recognition. With the rapid development of IoT and gait recognition, and the importance of identity recognition, a paper focusing on IoT gait recognition is needed.

In this paper, the second part introduces the wireless method, discusses how to implement it in four steps, and introduces several related works. The third part

introduces the types of sensor-based methods, focusing on the acceleration sensor. The fourth part presents the method based on vision, introduces the implementation method according to the feature extraction method [68] and raises the related work. In the fifth part, we discuss the current challenges in IOT gait recognition. The sixth part concludes this paper.

2. Wireless-based method

Currently, Wireless-based methods are mainly based on WiFi. WiFi devices are prevalent in daily life. Using infrastructure such as WiFi for gait recognition has many advantages in cost, ease of use, privacy and universality. It can be widely used in personal identification in smart homes, offices, and public places. It is also the most studied direction at present. Recently, some studies have realised gait recognition based on RFID [33] and millimetre wave [34]. In this section, we will mainly discuss the implementation of WiFi-based sensing technology.

At present, the research of using WiFi signals for identity recognition is mainly based on action recognition, which mostly senses the identity by analysing the disturbance characteristics of gait action on the signal. The experiment includes explicitly four steps: signal acquisition, preprocessing, feature extraction and identity recognition. Firstly, the signal data reflecting the action characteristics of the human body is obtained from the specific WiFi signal receiving device. Secondly, preprocess the collected data to reduce the signal noise. Then, the effective segments containing action information are separated by the algorithm, and then the features are extracted. Finally, the trained classifier is used for identity recognition. The following mainly introduces the concepts that will be used in the methods mentioned in this section by steps:

2.1. Signal Acquisition

Signal acquisition is the first step of sensing signals. Whether the collected data is effective or not is directly related to the effect of identity recognition. The acquisition device generally comprises a transmitting end and a receiving end. The transmitting end is usually a commercial WiFi device, and the receiving end is typically a computer with a wireless network card. At present, there are two main ways of WiFi sensing. One is that the received signal from the MAC layer is RSSI (Received Signal Strength Indication), and the other is CSI (Channel State Information) from the physical layer. RSSI signal is easy to be disturbed and has low stability, the complexity of indoor environment will make RSSI seriously disturbed by multipath effect [35]. CSI information is more

sensitive to RSS information and can provide more fine-grained changes and richer information, such as the scattering, fading, Doppler shift and other information of the transmission signal [36]. CSI is also less susceptible to multipath effects because the computer receiver uses OFDM (Orthogonal Frequency Division Multiplexing) technology for modulation. RSS was used to realise identification in early research, but all of the gait recognition based on WiFi proposed recently choose to use CSI because of the CSI's superiority.

After the CSI is generated, the CSI information will not be obtained directly. Because of this situation, in 2010, CSI tools [37] were proposed to get CSI information directly from commercial equipment. This set of tools is modified from the driver of the Intel WiFi link 5300 network card and provides the reading interface of CSI. After that, CSI information has been widely used in WiFi perception. In 2015, Atheros CSI Tool have been proposed [38] which make CSI more widely used, but the current WiFi-based gait recognition is mainly based on the CSI tool.

2.2. Pretreatment

After getting the signal, whether RSS or CSI signal, due to the interference caused by the equipment or the environment, it contains a lot of noise and needs preprocessing. Preprocessing methods include outlier removal, filtering, PCA and so on.

Due to the internal changes of the equipment, such as the changes in transmission power and transmission rate, these will lead to apparent differences in the signal. These obvious abnormal points are outliers, which need to be removed [39].

Filtering, human actions are generally concentrated in low frequency, so high-frequency components and DC components need to be filtered out. According to the frequency of different human activities, the range of filtering is also diverse (the human gait frequency in 0.3hz-2hz is noticeably high [40]).

2.3. Signal Processing

After obtaining effective data, it is necessary to extract the feature of the signal. In WiFi sensing, the commonly used features include statistical features in the time domain and frequency domain, Doppler frequency shift features, wavelet transform features and time-frequency map features. Statistical time-domain features are extracted directly from the waveform, and statistical frequency-domain features are extracted after the FFT transformation of the original waveform.

The calculation of statistical features is intuitive, and the processing is relatively simple, but it is easy to ignore the effective feature information. The characteristics of the Doppler frequency shift have good discrimination. Wavelet transform can realise signal

Table 1. METHODS AND RELVENT INFORMATIONS

Type	Category	Signal	Pretreatment	Processing	Classification	User	Result
WiFi-based	Wiwho [40]	Amplitude of CSI	Multipath elimination, Bandpass filtering	Statistica features	Decision tree	20	92%-80% for 2-6 people
	WiFi-id [48]	Amplitude of CSI	Bandpass filtering,CWT	Statistical features	SAC	20	93%-77% for 2-6 people
	Freeseense [42]	Amplitude and phase of CSI	Low pass filtering, PCA,DWT	Sliding window Segmentation approach,DTW	KNN	9	94.5%-88.9% for 2-6 people
	WiFiU [41]	Amplitude of CSI	PCA,STFT	Features including footstep length and walking speed	SVM	50	Top1,Top2,Top3 accuracies: 79.28%, 89.52%, 93.05%
	WiDIGR [49]	Amplitude of CSI	Bandpass filtering, PCA	STFT	SVM	60	92.83%-78.28% for 3-6 people
	TransferSense [50]	Amplitude and phase of CSI	Butterworth filtering	CNN	CNN	44	97%;77% for 10 new users
	AGait [46]	Amplitude of CSI	Butterworth filtering, 3-level wavelet shrinkage denoising	AE	AE	22	97.32%-89.77% for 4-10 people
Other	GRaaS [33]	RSSI and phase of CSI	Multi-tag solution	AT-LSTM , TSNet	AT-LSTM, TSNet	18	96.3%
	mmGaitNet [34]	3D point cloud	CFAR,DBscan	Hungarian algorithm	CNN	95	90%-88% for 1-5 people

analysis on multiple frequency scales and extract more sophisticated features. The time-frequency map feature converts the action signal into a time-frequency map through STFT, which can have richer information and be displayed visually.

2.4. Classification and Identification

After extracting features and establishing a database, it is necessary to recognise the signal. For human identification, there are already works done by conventional classification such as SVM [41], KNN [42], decision tree [40] and by deep learning-based classification including CNN [43], RNN [44], MLP [45] and AE [46]. The conventional classification relies on feature engineering using the hand-crafting feature, which can be time-consuming. At the same time, the biometric feature is expected to change when the user behaviour changes [] and the machine learning-based method cannot be adjusted when the user behaviour changes. To solve this problem, many researcher [41] reduced dimension to get better performance for not all hand-crafted features are noticeably contributive. They found that variability reduction in feature engineering could enhance the authentication mechanisms [47].

And many researcher turn to use deep learning methods, in order to increase the scale, accuracy and universality of sensing.

2.5. Related Works based on WiFi

WiWho [40] is the first research on WiFi-Based identity recognition, proposed in 2016. This method can identify a person from a small group of people through WiFi. WiWho uses multipath elimination and band-pass filtering to remove noise and conduct walking detection in the frequency domain. When a person starts walking, it carries out gait recognition After getting the information, including the step and walk features, Wiwho uses the decision tree to classify. WiWho carries out experimental evaluation in multiple places. In most cases, a walking distance of 2-3 meters is enough to identify. The experimental results show that WiWho can recognise one person among 2-6 people with an average accuracy of 92%-80%.

WiFi-ID [48] uses CSI amplitude information for identification, which mainly uses CSI data on the LOS path because this part of the data contains the most significant features of human gait. WiFi-ID removes high-frequency noise through Butterworth and uses CWT (Continuous Wavelet Transform) to select the frequency band signal of 20-80hz. The effects of human gait are most pronounced in the 20-80 Hz frequency bands for 5GHz WiFi, according to the empirical measurements. Seven time-domain features and three frequency-domain features are extracted in feature extraction. For there are too many dimensions of features and not all features are useful, the ReliefF feature selection algorithm is used to calculate various time-domain and frequency-domain features.

The experimental results show that the accuracy of WiFi-ID for single person identification in 2-6 people is 93%-77%.

Freeseense [42] method is also based on identity recognition on Los path and uses PCA and DWT (Discrete Wavelet Transform) to extract the shape features of Los path waveform for identity recognition through the nearest neighbor classifier. The experimental results show that the accuracy of Freeseense for single person identification in 2-6 people is 94.5%-88.9%.

WiFiU [41] is also a human gait recognition system based on WiFi. WiFiU obtains CSI signals from commercial WiFi devices, uses PCA analysis to denoise, and uses STFT (Short Time Fourier Transform) to transform them into the time-frequency joint domain. Use the SVM to classify the characteristics of gait extracted, such as gait speed and step length. For the gait data of 50 people in a 50 m² room, the recognition accuracy is 79.82%.

Crosssense [45] is a system that extends WiFi awareness to new scenes and a wider range. The existing application can be translated and utilised across sites and solve a larger sensing problem with CrossSense. Crosssense used MLP for generating the virtual samples. It used a feed-forward fully connected network with seven hidden layers used data from two domains to learn the mapping relation between them.

AGait [46] is a cycle-independent human gait recognition and walking direction estimation system using the attention-based RNN. AGait is implemented in three different indoor environments on commercial WiFi devices and achieves average F1 scores of 97.32%-89.77% in 4-10 people.

2.6. Related Works based on other passive sensing

Luo et al. [33] realised RFID-based gait recognition by monitoring the interruption of RFID signal when the user passes between transmitter and receivers [53]. By designing a set of edge cloud service architecture (GRaas), realising the fast response and low cost, achieving a robust recognition with 96.3%.

Zhen et al. [34] Realised the gait recognition based on millimeter waves. Firstly, they established a millimeter wave gait dataset, which contains the gait data of 95 volunteers in two cases. Based on this database, mmGaitNet based on deep learning is proposed. mmGaitNet achieved 90% accuracy for single person scenarios and 88% accuracy for five co-existing persons scenarios.

All the methods and its details are listed in Table 1. Wiwho requires subjects to move in a fixed direction within the specified area, and the accuracy decreases rapidly when the number of people increases. WiFi-ID does not take non-Los paths into consideration, and perform bad when there are a large group of

people. The innovation of WiFiU is mainly converting the signal into a spectrum, which can be analysed by image processing. The disadvantage is that it is not ideal in large groups. Freeseense has strong robustness and improves the recognition accuracy of large groups. WiDIGR [49] mainly achieved direction-independent recognition, which can be recognized no matter which direction people walk. In the latest works, Transersense [50] performs well in large-scale groups and has strong transfer ability. AGait achieves the best performance in both small and large groups by using attention-based RNN. As for other based algorithms, GRaas achieves both accurate recognition and low cost. Because of different signal carriers, mmGaitNet has more advantages for large-scale recognition.

3. Sensor-based method

Among sensor devices, there are accelerometer and gyroscope-based sensor methods, pressure sensor-based method, Electromyograph (EMG)-based approaches and ground reaction force (GRF) measurements. With the development of WIoT (Wearable Internet of Things) technology, various sensor technologies have been integrated into wearable Internet of things devices, such as smartwatches, smartphones, smart glasses, smart headphones, and so on.

Among them, accelerometers and gyroscope sensors are widely used. Unobtrusive acceleration gait recognition gets new development opportunities, which is also the current focus of sensor-based gait recognition research. Take the mobile phone as an example, the phone is fixed on the part of people's bodies during the experiment. The acceleration obtained by the sensor and the acceleration mode on the three axes become the features, and the gyroscope can get the rotation angle on the three axes. Pressure and GRF sensors measure the force applied to the sensor and has already been used for gait recognition [51, 52]. The EMG sensor measures either voluntary or involuntary muscle contract. The signal is obtained by the electrode on the skin surface, which can obtain different gait information. The acceleration sensor-based method and GRF measurement are suitable for IoT applications because of their unobtrusiveness and low cost for using the existing IoT devices.

There are few public datasets of accelerometers at present [53, 54], and it is difficult to make meaningful comparisons. Researchers generally establish their own database. This is because the experiment settings are quite different. The actual fixed place is different. Moreover, the number of volunteers in each database also varies greatly, so there is no way to compare the performance directly.

According to medical research [56], walking is performed according to the nerve output from the brain and then to muscles, joints, etc. For any kind of exercise,

it is necessary to change the brain's ideas into specific muscle patterns for walking [55]. The reaction force generated by the force applied to the ground supports the body to walk. The periodic movement of legs during walking is the essence of gait periodicity. Each gait cycle can be divided into two stages and eight configurations [56], as shown in Figure 2.

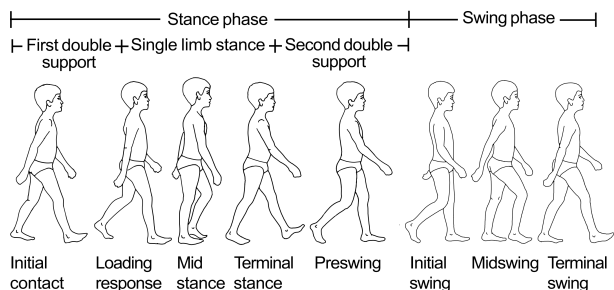


Figure 2. Walk cycle dynamics [56]

According to a study, the basic movement pattern is constant over a fixed range of motion speeds. Gait characteristics are often unique and naturally stereotyped [57]. For example, differences in pelvic sizes, age, and muscle will affect gait information greatly. In the exercise process, it is necessary to coordinate the muscles of the whole body to save energy as much as possible under the condition of maintaining stability. The final movement patterns often have significant differences, which can be used for identity recognition.

3.1. Related Works

Maesico et al. [58] proposed a mobile accelerometer gait recognition method based on a single consumer, using a new segmentation method that segments the gait signal into periods and steps. It achieves a 93% of recognition rate on the ZJU-Gaitacc dataset.

The accuracy of accelerometer gait recognition will decline when the speed changes. To solve the problem, Sun et al. [59] proposed a speed adaptive gait cycle segmentation method and individualized matching threshold generation method. And it achieved the average gait recognition and user authentication rates of 96.9% and 91.75%, respectively, in the ZJU-GaitAcc dataset and the self-collected dataset sampled at various walking speeds.

However, traditional methods also have many challenges. The above two test results are based on closed set identification, is achieved using the most constrained method, like limiting the walks in the gallery and in the probe to have a similar number of steps. All the methods do not perform well in the open set, and there are many influencing factors in practical application, which are difficult to apply to the

application. For example, people always fix the mobile phone on their bodies in the experiment, but in reality, people will put it in their pocket or hold it in their hands. Zou et al. [60] proposed a deep learning method based on gait data for learning and modelling gait features to solve this problem. In the test of establishing the dataset, the subjects were not constrained and did not limit the direction and speed. Finally, 93.5% higher than and 93.7% accuracy in person identification and authentication, respectively.

Terrier et al. [52] proposed a gait recognition method based on GRF. 30 people is asked to wear non-high heels and walk for 30 minutes on the force platform of the treadmill without prompting. The collected data are used to train CNN network. The data of the other six people are used for migration learning. Using several small samples for fine-tuning and the accuracy achieves 100%. It proved that as long as there is a pretrained CNN, only a few steps are enough to learn and recognize a new user's gait.

4. Vision-based method

Vision-based methods are different from WiFi-based and sensor-based methods. Public datasets are widely used in vision-based methods. Evaluation protocols are generally divided into subject-dependent and subject-independent. In the subject-independent test protocol, the data are divided into test sets and train sets, which are disjoint. After using the model to extract the features, the classifier compares the test features and train features, and the two highest similar features will be marked as the same person. At present, both protocols are widely used in gait recognition.

At present, there are a large number of datasets with different parameters, including clothes, observation angle, subjects' self-occlusion or ambient brightness. Generally speaking, we hope that the more samples and parameters in the dataset, the better. A better dataset can enable us to obtain a model with stronger performance. The current datasets mainly include ISIR datasets (MVLP [69], LP [64], LP-bag [66]) by the Osaka University, CASIA-B [72] dataset by the Chinese Academy of Sciences and USF [67] dataset by the University of South Florida. As shown in Table 2.

4.1. Related datasets

CASIA-B is currently the most widely used gait dataset, providing RGB and silhouette images from 124 subjects. This dataset captured and obtained data from eleven viewpoints and considered three different walking patterns, normal walking (NM), walking with a coat (CL), and walking with a bag (BG). The dataset was trained with the data of 74 people and tested with the data of the remaining 50 people.

Table 2. EXISTING DATASETS

Dataset	# of Subjects & Sequences	# of Views	Variations
OU-MVLP [69]	10,307 / 259,013	14	NW
OU-ISIR LP [64]	4,007 / 31,368	4	NW
OU-ISIR LP Bag [66]	62,528 / 187,584	1	NW,BG
CASIA-B [72]	124 / 13,680	11	NW,CL,BG
USF HumanID [67]	122 / 1,870	2	NW,CL,BG

OU-MVLP is the largest dataset of gait sequences at present (259,013). The data were obtained from fourteen observation angles of 0° - 90° , 180° - 270° . The data of 5153 people were designated as the training set and the rest as the test set.

In addition to these two datasets, there are datasets taking many other parameters into account. Including outdoor gait information [67], different walking speeds [70], multiple clothing combinations [63], carrying different items [66], and human bone information [61]. In 2015, after the deep neural network was widely used, the neural network began to be used for gait recognition. In recent years, the classification method of gait recognition has an apparent trend from using non-deep learning to deep learning.

4.2. Related works

According to a deep gait recognition focused taxonomy [68], classification methods can be divided into four dimensions: body representation, temporal representation, feature representation and neural architectures.

- (i) **Body representation.** Body representation is mainly based on people's body contour or bone information, of which the most commonly used is the silhouette. To obtain the silhouette, we need to detect, segment the pedestrians, remove them from the background image and binarize it. This method can focus on gait information, but it is not ideal when the appearance changes. In the pose-estimation method, because the skeleton information can be obtained, it is more adaptive to the change of viewing angle, but it is very sensitive to occlusion.
- (ii) **Temporal Representation.** This dimension is based on time information. It summarizes the walking information of a sequence of silhouettes in a single map. As for deep gait recognition architectures, gait silhouettes can be aggregated in the initial layer of a network (Figure 3.a), also known as temporal templates. Gait silhouettes can alternatively be aggregated in an intermediate layer of the network after several convolution

and pooling layers (Figure 3.b), also known as convolutional template. A typical and widely used example of temporal template is GEI [73], which averages gait silhouettes over one period. A convolutional template example is GCEM [76], which average convolutional maps obtained by several convolution and pooling layers. The frame rate does not affect this type of method.

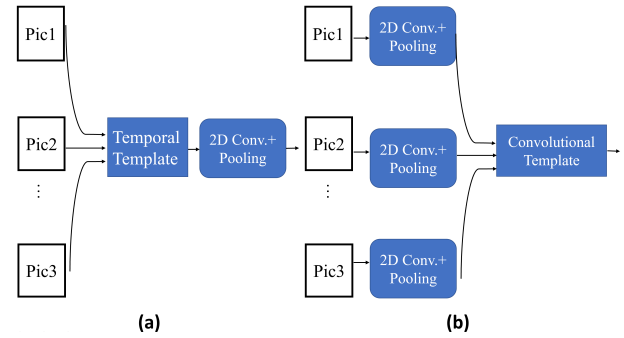


Figure 3. Deep gait recognition architectures. Generating templates in: (a) the initial layer of a deep network; (b) an intermediate layer of the network after several convolution and pooling layers

- (iii) **Feature Representation.** There are two features: global feature and local feature used in this dimension. The global feature learns the overall contour and skeleton, and the local representation should be divided into local regions for further processing. Global-based methods are more sensitive to appearance changes and occlusion. Local features often make different contributions to the results, and the relationship between features can also be learned. It can retain the location information and has higher robustness to the change of viewing angle.
- (iv) **Neural Architectures.** This dimension uses the deep neural network to obtain high-level abstract features. CNN is one of the most widely used methods in gait recognition. CNN have many different kind of layers including convolutional layer, pooling layer and fully-connected layer. Convolutional layer use input image and learned filters to create activation feature maps that capture features with different levels of detail. Convolutional layer also have the activation function to increase the non-linearity in the output. Pooling layers then reduce the spatial size of the feature maps by using nonlinear down-sampling strategies, such as average pooling or maximum pooling, thus decreasing the complexity of the network. Fully connected layers are finally used to learn the resulting 2D feature maps into 1D feature maps for further processing. Besides, many

other kind of neural network models have been applied, including DBN, RNN, DAE, GAN, CapsNet, 3D CNNs, GCN and hybrid deep networks.

GaitSet [62] uses the set perspective, regards gait as a group of independent frames, which are not affected by frame permutation, and can integrate frames of different videos in different scenes. GaitSet achieved 95% (NW), 87.2% (BG) and 70.4% (CL) accuracy rates on CASIA-B. And only a few frames are needed to achieve good accuracy. On CASIA-B, the accuracy of 7 frames is 82.5%.

3DCNNGait [65] proposed a novel multiple-temporal-scale gait recognition, which uses both frame and interval fusion information. Realized the interval-level representation by a local transformation module and applied 3D CNN to the temporal scales. This model achieved 96.7% (NW), 93.0% and 81.5% on CASIA-B.

EV-GATE-3DGRAP [71] uses dynamic vision sensors (event cameras), which has ultra-low resource consumption, extremely high sampling rate, greater resolution and dynamic range. The information obtained is an asynchronous event and event flow. In this work, based on the graph and image-like representation of event flow, a new event-based gait recognition method using GCN and CNN is proposed. EV-GATE-3DGRAP achieved 94.9% in the self-built DVS128-Gait dataset.

5. Challenges

With the development of IoT technology, gait recognition has new development opportunities, but it comes with higher requirements and challenges in IoT applications. In vision-based methods, although the accuracy of NM has been high, the problems of BG and CL recognizing have not been well solved. (the SOTA algorithm achieves 93.0% in BG and 81.5% in CL) In reality, more factors will affect the results, such as occlusion or cross angle parameters. Moreover, mood, drinking alcohol [78] and health status can also affect people's gait information.

Although the method based on deep learning achieves higher accuracy, it needs higher computational cost. Even the best algorithm at present may not achieve the continuous gait recognition for some edge computing devices with weak computing performance, such as mobile phones and wearable devices. The deep learning model needs to be compressed, including tensor decomposition [79], pruning [82] and parameter sharing [80]. At the same time, the deep learning method relies heavily on data, and the current dataset is not big enough. Models based on the sketched distribution are not robust. IoT systems are also faced with numerous security threats on physical, protocol, communication, and application layers [82].

6. Conclusion

In IoT applications, identity recognition is a fundamental and critical requirement. With the development of IOT technologies such as wearable devices and environmental sensors, it can meet the new and higher requirements put forward in practical IoT applications. This paper reviewed the implementation of gait recognition according to three methods: wireless-based, sensor-based and vision-based method. We also reviewed the representative and latest datasets, algorithms, and compared the performance of algorithms. Finally, we discussed the current challenges, pointing out several promising future research directions in IOT gait recognition. We expect that this survey provides insights into comprehensive landscape of IOT gait recognition guiding researchers in advancing future research.

References

- [1] Weiser M. He Computer for the 21 St Centy. SCIENTIFIC AMERICAN. 1991:13.
- [2] Chen D, Hu Y, Wang L, Zomaya AY, Li X. H-PARAFAC: Hierarchical Parallel Factor Analysis of Multidimensional Big Data. IEEE Transactions on Parallel and Distributed Systems. 2016;28(4):1091–1104.
- [3] Chen D, Tang Y, Zhang H, Wang L, Li X. Incremental Factorization of Big Time Series Data with Blind Factor Approximation. IEEE Transactions on Knowledge and Data Engineering. 2019;33(2):569–584.
- [4] Perera C, Zaslavsky A, Christen P, Georgakopoulos D. Sensing as a Service Model for Smart Cities Supported by Internet of Things. Transactions on emerging telecommunications technologies. 2014;25(1):81–93.
- [5] Xiaohang Wang, Jin Song Dong, ChungYau Chin, Hettiarachchi S, Daqing Zhang. Semantic Space: An Infrastructure for Smart Spaces. IEEE Pervasive Computing. 2004 Jul;03(03):32–39.
- [6] Gupta BB, Sheng QZ. Machine Learning for Computer and Cyber Security: Principle, Algorithms, and Practices. CRC Press; 2019.
- [7] Jain AK, Ross A, Pankanti S. Biometrics: A Tool for Information Security. IEEE transactions on information forensics and security. 2006;1(2):125–143.
- [8] Ometov A, Petrov V, Bezzateev S, Andreev S, Koucheryav Y, Gerla M. Challenges of Multi-Factor Authentication for Securing Advanced IoT Applications. IEEE Network. 2019;33(2):82–88.
- [9] Liu H, Zhang P, Pu G, Yang T, Maharjan S, Zhang Y. Blockchain Empowered Cooperative Authentication with Data Traceability in Vehicular Edge Computing. IEEE Transactions on Vehicular Technology. 2020;69(4):4221–4232.
- [10] Unar J, Seng WC, Abbasi A. A Review of Biometric Technology along with Trends and Prospects. Pattern recognition. 2014;47(8):2673–2688.
- [11] Čagalj M, Perković T, Bugarić M. Timing Attacks on Cognitive Authentication Schemes. IEEE Transactions

- on Information Forensics and Security. 2014;10(3):584–596.
- [12] Chen T, Farcasin M, Chan-Tin E. Smartphone Passcode Prediction. *IET Information Security*. 2018;12(5):431–437.
- [13] Kwon T, Hong J. Analysis and Improvement of a Pin-Entry Method Resilient to Shoulder-Surfing and Recording Attacks. *Ieee transactions on information forensics and security*. 2014;10(2):278–292.
- [14] Abdelrahman Y, Khamis M, Schneegass S, Alt F. Stay Cool! Understanding Thermal Attacks on Mobile-Based User Authentication. In: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*; 2017. p. 3751–3763.
- [15] Li D, Zhang XP, Hu M, Zhai G, Yang X. Physical Password Breaking via Thermal Sequence Analysis. *IEEE Transactions on Information Forensics and Security*. 2018;14(5):1142–1154.
- [16] Aviv AJ, Gibson K, Mossop E, Blaze M, Smith JM. Smudge Attacks on Smartphone Touch Screens. In: *4th USENIX Workshop on Offensive Technologies (WOOT 10)*; 2010. .
- [17] Schneegass S, Steimle F, Bulling A, Alt F, Schmidt A. Smudgesafe: Geometric Image Transformations for Smudge-Resistant User Authentication. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*; 2014. p. 775–786.
- [18] Gafurov D. A Survey of Biometric Gait Recognition: Approaches, Security and Challenges. In: *Annual Norwegian Computer Science Conference*. Annual Norwegian Computer Science Conference Norway; 2007. p. 19–21.
- [19] Kim S, Lee H, Kwon T. Poster: Rethinking Fingerprint Identification on Smartphones. In: *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*; 2017. p. 2515–2517.
- [20] Roy A, Memon N, Ross A. Masterprint: Exploring the Vulnerability of Partial Fingerprint-Based Authentication Systems. *IEEE Transactions on Information Forensics and Security*. 2017;12(9):2013–2025.
- [21] Ren Y, Fang Z, Liu D, Chen C. Replay Attack Detection Based on Distortion by Loudspeaker for Voice Authentication. *Multimedia Tools and Applications*. 2019;78(7):8383–8396.
- [22] Wang Y, Cai W, Gu T, Shao W, Li Y, Yu Y. Secure Your Voice: An Oral Airflow-Based Continuous Liveness Detection for Voice Assistants. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*. 2019;3(4):1–28.
- [23] Meng W, Wong DS, Furnell S, Zhou J. Surveying the Development of Biometric User Authentication on Mobile Phones. *IEEE Communications Surveys & Tutorials*. 2014;17(3):1268–1293.
- [24] Weiser M. The Computer for the 21st Century. *IEEE pervasive computing*. 2002;1(1):19–25.
- [25] Aqmar MR, Shinoda K, Furui S. Robust Gait Recognition against Speed Variation. In: *2010 20th International Conference on Pattern Recognition*. IEEE; 2010. p. 2190–2193.
- [26] Guan Y, Li CT. A Robust Speed-Invariant Gait Recognition System for Walker and Runner Identification. In: *2013 International Conference on Biometrics (ICB)*. IEEE; 2013. p. 1–8.
- [27] Mansur A, Makihara Y, Aqmar R, Yagi Y. Gait Recognition under Speed Transition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2014. p. 2521–2528.
- [28] Chen Q, Wang Y, Liu Z, Liu Q, Huang D. Feature Map Pooling for Cross-View Gait Recognition Based on Silhouette Sequence Images. In: *2017 IEEE International Joint Conference on Biometrics (IJCB)*. IEEE; 2017. p. 54–61.
- [29] Muramatsu D, Makihara Y, Yagi Y. View Transformation Model Incorporating Quality Measures for Cross-View Gait Recognition. *IEEE transactions on cybernetics*. 2015;46(7):1602–1615.
- [30] Yu S, Tan D, Tan T. A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. In: *18th International Conference on Pattern Recognition (ICPR'06)*. vol. 4. IEEE; 2006. p. 441–444.
- [31] Zhang Z, Chen J, Wu Q, Shao L. GII Representation-Based Cross-View Gait Recognition by Discriminative Projection with List-Wise Constraints. *IEEE transactions on cybernetics*. 2017;48(10):2935–2947.
- [32] Cham R, Redfern MS. Changes in Gait When Anticipating Slippery Floors. *Gait & posture*. 2002;15(2):159–171.
- [33] Luo C, Wu J, Li J, Wang J, Xu W, Ming Z, et al. Gait Recognition as a Service for Unobtrusive User Identification in Smart Spaces. *ACM Transactions on Internet of Things*. 2020 Mar;1(1):1–21.
- [34] Meng Z, Fu S, Yan J, Liang H, Zhou A, Zhu S, et al. Gait Recognition for Co-Existing Multiple People Using Millimeter Wave Sensing;8.
- [35] Xiao W, Song B, Yu X, Chen P. Nonlinear Optimization-Based Device-Free Localization with Outlier Link Rejection. *Sensors*. 2015;15(4):8072–8087.
- [36] Ohara K, Maekawa T, Matsushita Y. Detecting State Changes of Indoor Everyday Objects Using Wi-Fi Channel State Information. *Proceedings of the ACM on interactive, mobile, wearable and ubiquitous technologies*. 2017;1(3):1–28.
- [37] Halperin D, Hu W, Sheth A, Wetherall D. Tool Release: Gathering 802.11 n Traces with Channel State Information. *ACM SIGCOMM computer communication review*. 2011;41(1):53–53.
- [38] Xie Y, Li Z, Li M. Precise Power Delay Profiling with Commodity Wi-Fi. *IEEE Transactions on Mobile Computing*. 2018;18(6):1342–1355.
- [39] Davies L, Gather U. The Identification of Multiple Outliers. *Journal of the American Statistical Association*. 1993;88(423):782–792.
- [40] Zeng Y, Pathak PH, Mohapatra P. WiWho: WiFi-based Person Identification in Smart Spaces. In: *2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN)*. IEEE; 2016. p. 1–12.
- [41] Wang W, Liu AX, Shahzad M. Gait Recognition Using Wifi Signals. In: *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. Heidelberg Germany: ACM; 2016. p. 363–373.

- [42] Xin T, Guo B, Wang Z, Li M, Yu Z, Zhou X. Freesense: Indoor Human Identification with Wi-Fi Signals. In: 2016 IEEE Global Communications Conference (GLOBECOM). IEEE; 2016. p. 1–7.
- [43] Batchuluun G, Naqvi RA, Kim W, Park KR. Body-Movement-Based Human Identification Using Convolutional Neural Network. *Expert Systems with Applications*. 2018;101:56–77.
- [44] Ming X, Feng H, Bu Q, Zhang J, Yang G, Zhang T. HumanFi: WiFi-based Human Identification Using Recurrent Neural Network. In: 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI). IEEE; 2019. p. 640–647.
- [45] Zhang J, Tang Z, Li M, Fang D, Nurmi P, Wang Z. CrossSense: Towards Cross-Site and Large-Scale WiFi Sensing. In: Proceedings of the 24th Annual International Conference on Mobile Computing and Networking. New Delhi India: ACM; 2018. p. 305–320.
- [46] Xu Y, Yang W, Chen M, Chen S, Huang L. Attention-Based Gait Recognition and Walking Direction Estimation in Wi-Fi Networks. *IEEE Transactions on Mobile Computing*. 2022 Feb;21(2):465–479.
- [47] Cai Z, Shen C, Guan X. Mitigating Behavioral Variability for Mouse Dynamics: A Dimensionality-Reduction-Based Approach. *IEEE Transactions on Human-Machine Systems*. 2014;44(2):244–255.
- [48] Zhang J, Wei B, Hu W, Kanhere SS. WiFi-ID: Human Identification Using WiFi Signal. In: 2016 International Conference on Distributed Computing in Sensor Systems (DCOSS). Washington, DC, USA: IEEE; 2016. p. 75–82.
- [49] Zhang L, Wang C, Ma M, Zhang D. WiDIGR: Direction-independent Gait Recognition System Using Commercial Wi-Fi Devices. *IEEE Internet of Things Journal*. 2019;7(2):1178–1191.
- [50] Bu Q, Ming X, Hu J, Zhang T, Feng J, Zhang J. TransferSense: Towards Environment Independent and One-Shot Wifi Sensing. *Personal and Ubiquitous Computing*. 2021:1–19.
- [51] Meharia P, Agrawal DP. The Human Key: Identification and Authentication in Wearable Devices Using Gait. *Journal of Information Privacy and Security*. 2015;11(2):80–96.
- [52] Terrier P. Gait Recognition via Deep Learning of the Center-of-Pressure Trajectory. *Applied Sciences*. 2020;10(3):774.
- [53] Ngo TT, Makihara Y, Nagahara H, Mukaigawa Y, Yagi Y. The Largest Inertial Sensor-Based Gait Database and Performance Evaluation of Gait-Based Personal Authentication. *Pattern Recognition*. 2014;47(1):228–237.
- [54] Zhang Y, Pan G, Jia K, Lu M, Wang Y, Wu Z. Accelerometer-Based Gait Recognition by Sparse Representation of Signature Points with Clusters. *IEEE transactions on cybernetics*. 2014;45(9):1864–1875.
- [55] Enoka RM. *Neuromechanical Basis of Kinesiology*. ERIC; 1988.
- [56] Vaughan CL, Davis BL, O’connor JC. Dynamics of Human Gait. *Human Kinetics*; 1992.
- [57] Borghese NA, Bianchi L, Lacquaniti F. Kinematic Determinants of Human Locomotion. *The Journal of physiology*. 1996;494(3):863–879.
- [58] Marsico MD, Mecca A. Biometric Walk Recognizer. *Multimedia Tools and Applications*. 2017;76(4):4713–4745.
- [59] Sun F, Mao C, Fan X, Li Y. Accelerometer-Based Speed-Adaptive Gait Authentication Method for Wearable IoT Devices. *IEEE Internet of Things Journal*. 2018;6(1):820–830.
- [60] Zou Q, Wang Y, Wang Q, Zhao Y, Li Q. Deep Learning-Based Gait Recognition Using Smartphones in the Wild. *IEEE Transactions on Information Forensics and Security*. 2020;15:3197–3212.
- [61] An W, Yu S, Makihara Y, Wu X, Xu C, Yu Y, et al. Performance Evaluation of Model-Based Gait on Multi-View Very Large Population Database with Pose Sequences. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. 2020;2(4):421–430.
- [62] Chao H, He Y, Zhang J, Feng J. GaitSet: Regarding Gait as a Set for Cross-View Gait Recognition. *Proceedings of the AAAI Conference on Artificial Intelligence*. 2019 Jul;33:8126–8133.
- [63] Hossain MA, Makihara Y, Wang J, Yagi Y. Clothing-Invariant Gait Identification Using Part-Based Clothing Categorization and Adaptive Weight Control. *Pattern Recognition*. 2010;43(6):2281–2291.
- [64] Iwama H, Okumura M, Makihara Y, Yagi Y. The OU-ISIR Gait Database Comprising the Large Population Dataset and Performance Evaluation of Gait Recognition. *IEEE Transactions on Information Forensics and Security*. 2012;7(5):1511–1521.
- [65] Lin B, Zhang S, Bao F. Gait Recognition with Multiple-Temporal-Scale 3D Convolutional Neural Network. In: Proceedings of the 28th ACM International Conference on Multimedia. Seattle WA USA: ACM; 2020. p. 3054–3062.
- [66] Uddin MZ, Ngo TT, Makihara Y, Takemura N, Li X, Muramatsu D, et al. The Ou-Isir Large Population Gait Database with Real-Life Carried Object and Its Performance Evaluation. *IPSN Transactions on Computer Vision and Applications*. 2018;10(1):1–11.
- [67] Sarkar S, Phillips PJ, Liu Z, Vega IR, Grother P, Bowyer KW. The Humanid Gait Challenge Problem: Data Sets, Performance, and Analysis. *IEEE transactions on pattern analysis and machine intelligence*. 2005;27(2):162–177.
- [68] Sepas-Moghaddam A, Etemad A. Deep Gait Recognition: A Survey. *arXiv:210209546 [cs]*. 2022 Feb.
- [69] Takemura N, Makihara Y, Muramatsu D, Echigo T, Yagi Y. Multi-View Large Population Gait Dataset and Its Performance Evaluation for Cross-View Gait Recognition. *IPSN Transactions on Computer Vision and Applications*. 2018;10(1):1–14.
- [70] Tsuji A, Makihara Y, Yagi Y. Silhouette Transformation Based on Walking Speed for Gait Identification. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE; 2010. p. 717–722.
- [71] Wang Y, Zhang X, Shen Y, Du B, Zhao G, Cui Lizhen LC, et al. Event-Stream Representation for Human

- Gaits Identification Using Deep Neural Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2021:1–1.
- [72] Yu S, Tan D, Tan T. A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. In: 18th International Conference on Pattern Recognition (ICPR'06). vol. 4. IEEE; 2006. p. 441–444.
- [73] Han J, Bhanu B. Individual Recognition Using Gait Energy Image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 2006 Feb;28(2):316–322.
- [74] Liang Y, Samtani S, Guo B, Yu Z. Behavioral Biometrics for Continuous Authentication in the Internet-of-Things Era: An Artificial Intelligence Perspective. *IEEE Internet of Things Journal*. 2020 Sep;7(9):9128–9143.
- [75] Marsico MD, Mecca A. A Survey on Gait Recognition via Wearable Sensors. *ACM Computing Surveys*. 2020 Jul;52(4):1–39.
- [76] Sepas-Moghaddam A, Etemad A. View-Invariant Gait Recognition With Attentive Recurrent Learning of Partial Representations. *IEEE Transactions on Biometrics, Behavior, and Identity Science*. 2021 Jan;3(1):124–137.
- [77] Singh JP, Jain S, Arora S, Singh UP. Vision-Based Gait Recognition: A Survey. *IEEE Access*. 2018;6:70497–70527.
- [78] Fein G, Greenstein D. Gait and Balance Deficits in Chronic Alcoholics: No Improvement from 10 Weeks through 1 Year Abstinence. *Alcoholism: Clinical and Experimental Research*. 2013;37(1):86–95.
- [79] Lane ND, Bhattacharya S, Georgiev P, Forlivesi C, Jiao L, Qendro L, et al. Deepx: A Software Accelerator for Low-Power Deep Learning Inference on Mobile Devices. In: 2016 15th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN). IEEE; 2016. p. 1–12.
- [80] Lu Z, Sindhwani V, Sainath TN. Learning Compact Recurrent Neural Networks. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2016. p. 5960–5964.
- [81] Mosenia A, Jha NK. A Comprehensive Study of Security of Internet-of-Things. *IEEE Transactions on emerging topics in computing*. 2016;5(4):586–602.
- [82] Yang TJ, Chen YH, Sze V. Designing Energy-Efficient Convolutional Neural Networks Using Energy-Aware Pruning. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2017. p. 5687–5695.