# An Overarching Study on Gait Recognition – Appearance Based Approach

Induja P[1], Sujata B M[2], Sathyabama S[3]

*HKBK College of Engineering, Bengaluru*

*induja.r.s@gmail.com[1], sujaya.bm@gmail.com[2], ssb.bama@gmail.com[3]*

***Abstract:*** *Gait Recognition is becoming a topic of continual research interest in biometrics research community. It has turn out to be a vital biometric feature for human identification. It is main advantage is its non – invasive nature. Gait Recognition is a type of pattern recognition based on the walking styles of a person. Machine and Deep learning technologies have modernized the field of gait recognition by aiding computers to automatically learn and extract complex patterns. These techniques analyze video recordings of a person to determine key features in each person's gait, and these features are used to identify the person. The primary goal of this paper is to provide an elucidate survey of the progressive ideas in appearance-based gait recognition techniques, mentioning their applications, strengths and limitations. Through the analysis of the techniques discussed in the recent years, enlighten the significant advance, draw attention to the drawbacks that have been faced, and identify areas of prospective future research and advances in technology of gait recognition. Its hopeful that this survey analysis in appearance-based gait recognition will be a helpful resource for the researchers on this dynamic field.*

**Keywords:** *Gait Recognition, Feature Extraction, Neural Networks, Classification techniques.*

## 1. Introduction

To identify subjects / persons features like iris, face, fingerprint, etc., have been used previously [1], [2]. These features can work well but it's in the need of cooperation of the subjects under investigation and so they have limitations in use. Gait features can be easily detected at a long distance that has become a mainstream technology for identification [3], [4]. By comparing other identification systems, the classification and identification of human gait offers several advantages. Human gait may be considered as a biometric that allows people to be identified by the way they walk. Unlike the other biometrics, gait can be recorded from a distance and without the individual's active participation. Gait is useful for video surveillance-based applications because of these advantages. Gait recognition has potential uses in security and surveillance, including the identification of people in crowd public places and the tracking of suspects [5]. One major complexity in gait-based identification is the sensitive towards the circumstances such as background variations, lighting changes, and objects carried [6]. One important attribute in capturing gait patterns accurately is camera's viewing angle. When a person walks, the shape of their body changes over time so that, Gait is a time variant behavior. For getting the human gait identification to be efficient, the methods incorporating for the same should be considered for the view invariant and shape invariant supported. Many deep learning-based methods have been proposed [7], [8], [9] to address this problem.

The general steps included in a gait recognition system is detailed as follows: [10].

1) ***Data Collection:*** It is essential to collect the gait patterns of the individuals. Video recordings, pressure, floor and motion capture sensors, can be used to collect the data.

2) ***Feature Extraction:*** It is important to extract unique features of their walking pattern, like stride length, walking speed and foot angle. The features used in gait-based user recognition can be broadly classified into static and dynamic information in which static information includes the person's physical appearance like height, limb length, and stride length and dynamic information can be obtained from their movements.

3) ***Dimension Reduction:*** Features extracted from the gait data cannot be used for classification because the size of those features will be more than the number of training samples. Therefore, dimensionality reduction should be adhered.

4) ***Classification:*** Based on the gait features obtained in the previous steps, either deep learning or a machine learning algorithm.

In the computer vision, the gait recognition can be approached in two ways: (1) Model based and (2) Appearance based. In model-based approach, the mathematical models to represent the walking motion of a person. The geometry of motion of joint angles are modeled when the person walk is utilized for this approach. In appearance-based approach, the features are extracted from the physical appearance of a person's gait, such as body shape and limb movements. In this approach, from the silhouettes of the person's gait, the features are extracted. One advantage of this appearance-based method is that it does not require any sensors, or subject consent since it depends on the data capturing from security cameras. This makes it useful for real world applications. The former method has benefits such as detailed motion information and modeling skeletal systems and has the disadvantage that the key point estimation will not be accurate. In addition, among the model based and appearance-based gait recognition approaches, the former approach results in lower performance in recognition [11-13]. Hence, the motive of this paper is giving survey on appearance-based gait recognition approach.

In the next section, the detailed view of gait recognition techniques has been discussed.

## 2. Gait Recognition

The conceptual framework of the traditional gait recognition is shown in Fig.1. It includes the blocks of data collection and machine learning techniques like Feature extraction, Feature Representation, Dimensionality Reduction, Classifier. This method is suitable for appearance-based gait recognition.

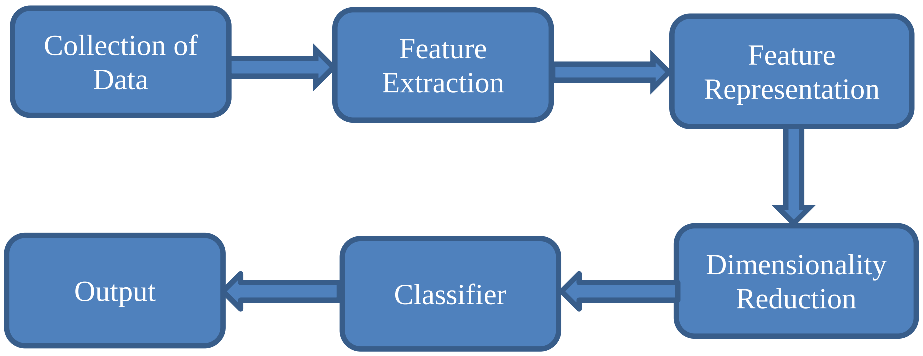


**Figure 1: Framework of Traditional Gait Recognition System**

The deep learning-based gait recognition consists of data collection, training blocks includes neural network that does feature extraction and classification. It takes less steps than traditional since in deep learning-based gait recognition the feature extraction and classification has been done in a single step shown in Fig. 2. This method suits for model-based gait recognition. It will improve efficiency and reduce the likelihood of errors but it needs additional data and computational resources for training and testing along with that deep learning skills.

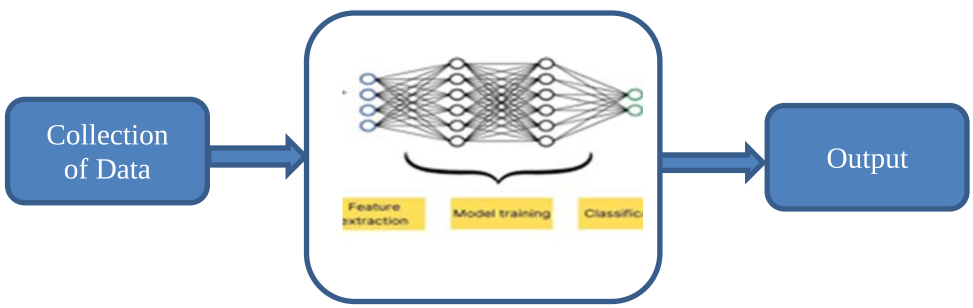


**Figure 2: Framework of Deep learning-based Gait Recognition System**

The following section explains the steps involved in the gait recognition process and some of the machine learning and deep learning techniques in detail:

### 2.1 Data Collection

The initial stage in the gait recognition framework includes data collection to identify individual's gait patterns. Recognition of gait can be performed using various input data like RGB image, silhouette, GEI (Gait Energy Image), optical flow image, body skeleton, etc., In addition to that wearable sensor can also give the movement data and pressure data that can also be used for gait recognition. The examples of the different input data type obtained from different gait dataset are shown in Fig. 3 [14 – 17].

(a) (b) (c)

**Figure 3: Examples of input data types for gait recognition a) RGB image b) Silhouette c) GEI**

### 2.2 Machine Learning Techniques

#### 2.2.1 Feature Extraction and Representation

The methods for extracting the features for the gait recognition requires that to describe the unique characteristics of individuals and to have the adoptability for changing environmental conditions, clothing conditions, etc., As mentioned earlier, there are two approaches as model based and appearance based in gait recognition. The difference between these two approaches is in how the features are extracted and the type of the data used for recognition. In the former approach, features are extracted from a physical model of the human body that predicts like joint angles, etc., during walking. But in the latter approach, the features are extracted by considering the entire movement pattern of the walking individual's body. This method contains more invariant features [18].

#### 2.2.2 Dimensionality Reduction:

The main aim of dimensionality reduction is to reduce the dimension of the feature vector that represents the gait patterns. The extracted feature vector is high dimensional and comprises a huge number of variables. This issue makes the gait recognition as a complex one. This dimensionality reduction technique goals to address this issue without losing the essential information. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two main reduction techniques. Its brief is as follows:

***PCA:*** transforms the feature vector into a set of orthogonal principal components, each of that is a linear combination of the original variables. This includes firstly primary components that are kept and the other components are discarded [19].

***LDA:*** It increases the distance between the means of different classes, while reducing the variance within each class. It does the reduction technique by maximizing the difference between the various classes [20].

#### 2.2.3 Classification:

The features extracted in the previous step are to be turned as the feature vector representing the gait. The similarity between the features is measured by a vector similarity metric such as Cosine similarity, Euclidean Distance, Manhattan Distance, etc., This will be the input to a classification stage in which the gait sequence is labeled or classified by using classification algorithms like SVM, HMM, etc., Finally a label is assigned to each image by this classifier.

The two main SVM and HMM classifier algorithm are discussed below briefly:

***(a) Support Vector Machine (SVM):***

This algorithm has the two phases namely training and testing. It finds the hyperplane that differentiates the data points of the different individuals. This hyperplane is selected in such a way that the distance between the hyperplane and the nearest data points from each class. As the SVM model has been trained, the new gait patterns are classified by the features extracted. The new gait pattern will fall on one side of the hyperplane. Based on that the gait pattern is classified as particular individual trained on that side of the hyperplane. Basically, SVM is a binary classification algorithm that classifies among only two classes. To make it suitable for the gait recognition, SVM multiclass technique has to be considered since this recognition is among multiple individuals [21].

***(b) Hidden Markov Model (HMM):***

This method is also having two phases as training and testing. The model parameters of HMM are estimated by using a training gait data. The temporal properties of gait sequences are represented by using Hidden Markov Model (HMM). Each different gait patterns are described as several states. The possibility of transitioning gait patterns is shown by transition probabilities between the states. These parameters include the transition and observation probability [22].

### 2.3 Deep Learning Techniques

#### 2.3.1 Feature Extraction

This method is automatically learning to identify individuals directly from the data based on the unique gait patterns. Using of this layered technique gives the complex feature extraction from unprocessed data and reducing the necessity of manual process of extracting features. This automatic technique learns both spatial and temporal features across the frame sequences.

#### 2.3.2 Dimensionality Reduction:

Coming to dimensionality reduction techniques in the case of deep learning, the following are some of the dimensionality reduction techniques.

***Pooling:*** It is applied for a set of values arranged in a grid-like format. To get the single value as the output, this pooling process divides the entire grid structure into non-overlapping or overlapping subregions. The aggregate function is applied to the values in each subregion. There are two main pooling functions. They are Maximum pooling and average pooling. The former pooling takes the max value in each subregion and the latter involves the average value [23].

***Autoencoders (AEs):*** It consists of an encoder that is responsible for the input dimensions and a decoder that reconstructs the input data from the reduced representation. The middle layer or the code layer has the lowest dimensionality and the reduced representation of the input data [24].

### 2.3.3 Classification:

For classifying the gait data into the particular class, this deep learning model is using the learned features. An activation layer in the output layer is doing this classification. This method is having two phases as training and testing. The labeled dataset is used to train the model. Backpropagation method that is based on the difference between the predicted and original labels, to improve the classification accuracy.

Some Neural Network techniques used for gait recognition is described as follows.

i) ***Convolutional Neural Networks (CNN):*** CNN is one of the common neural network types used in the gait recognition. It is having the layers of many interconnected nodes like convolutional layers, pooling layers and fully connected layers [25]. The convolutional layer is responsible for the learning and extracting the features from the input data. Pooling layer is doing the dimensionality reduction operation. The fully connected layers are responsible for the classification of outputs from the previous layers into different gait patterns of individuals. This CNN is highly effective for spatial features. In addition to that, CNN can be updated to extract temporal features.

ii) ***Recurrent Neural Network: (RNN):*** RNN which is having multiple neurons and is a well suitable method for evaluating the temporal changes of human gait patterns. It is working as internal states (hidden) are storing the previous inputs, that is updated while new data points are processed [26]. This is having the recurrent behavior, that enables the layer to receive the sequence of inputs and output of a sequence. It's not suitable for the long-term sequences of gait patterns. For that Long Short- term Memory (LSTM) is developed to solve this issue.

iii) ***Generative Adversarial Networks (GAN):*** GAN is making this gait recognition approach as the more robust due to their adaptability under various conditions. GAN is having two neural networks namely the generator and the discriminator, that are to be trained simultaneously [27]. It is very useful for cross view gait recognition. But it is having various challenges like stability in training, etc., that cause result in low-quality recognition.

iv) ***3D Convolutional Neural Network (3D CNN):*** 3D CNN is enhancing the abilities of the conventional CNN so that it can process the volumetric data directly. It examines a sequence of frames as a single input so that it can extract features that store both the shape and the movement of the subject [28]. Though it's having more benefits, its high computational cost of processing 3D data for training the long sequence data is the reason of using these techniques as less.

v) ***Hybrid Models:*** Instead of making the single neural network for feature extraction and classification, the combination of more neural network technique will be efficient for the gait recognition. For example, spatial features such as shape, posture are extracted from CNN's and temporal features are extracted by using RNN's or LSTM's. This hybrid technique makes the accuracy on gait recognition as the best one.

## 3. Appearance Based Approach on Gait Recognition

This section reviews the appearance-based Gait Recognition approaches based on the publication in recent years. This approach considers the structure and motion of the entire human body. It also depends on the silhouette shape and the dynamic information.

i) In Zhang Y, et. Al., [29], a new loss function for cross view gait recognition and an LSTM model are proposed that is used to learn spatial – temporal features. Dividing of Gait silhouettes into four horizontal parts and each of them are given as an input to separate CNN. To average frame level features, attention weights of each part has been calculated. In training phase, the various features are treated separately in different loss functions. In testing phase, all the weighted features are concatenated to a feature vector. Local gait features are obtained by the several independent CNN's that have the input as each local part. For temporal features, LSTM – based temporal attention model is used. The datasets of CASIA – B, OULP, and OUMVLP with the accuracy of 96.0%, 99.3%, 89.0% respectively in this research. The features obtained by this method gives the increased accuracy of the model since the different parts of the silhouettes assigned with certain weight values. It is brought the disadvantage of high computational cost and the feature dimensionality problem.

ii) Fan et al. [30] used a deep learning-based solution consists of Frame – Level Part feature Extractor (FPFE) that learns the part – level features and Micro – Motion capture module (MCM) that aims to learn spatiotemporal features. This method used the CASIA – B and OUMVLP datasets with the accuracy of 96.2% and 88.7% has been achieved. The detailed modeling of temporal features has been addressed in this paper.

iii) Hou S, Cao C, et al. [31] proposed the Gait Lateral Network (GLN) in which the discriminative features from deep CNN's have been extracted to integrate silhouette level and set level features, lateral connections are used. It is having a compact block to decrease the dimension of the gait representation. With the datasets of CASIA – B and OUMVLP with the accuracy of 96.8% and 89.1% respectively has been achieved. The lateral connections are aggregating more visual details so that the accuracy of the gait recognition has been increased. This proposed method is outperformed in terms of accuracy and dimension.

iv) Hou S, et al. [32] in this research paper, a Set Residual Network (SRN) is used for the recognition of the gait patterns. Set Residual Block (SR Block) is the fundamental block in which the features are learnt from silhouettes. It is divided into two parallel branches namely Silhouette branch and Set branch. The resultant features from these two branches are concatenated by using Residual connection and Leaky ReLU. This paper also mentioned about the Dual Feature Pyramid Approach for feature learning using shallow layer features. It is tested on the CASIA – B and OUMVLP datasets and achieved the accuracy of 97.1 % and 89.1 %. Compared to GLN methods mentioned in the previous paper [31], due to shallow layer features, the accuracy is better than previous GLN methods.

v) Gul S, Malik MI, et al. [33] used 3D convolutional Neural Networks (3D CNN) to extract the spatiotemporal features of the gait sequence. Two sets of convolutional layers namely pooling layer and two fully connected layers are there in this network. GEI (Gait Energy Image) and the 3D CNN model was used for feature extraction and gait recognition. Next optimization of these network parameters is done by Bayesian Optimization. The datasets of CASIA – B and OULP are used and the achieved accuracy is 98.3 % and 93.1 % respectively.

vi) Huang X, et al. [34] proposed a method which uses the Context – Sensitive – Temporal Feature Learning (CSTL) network and salient Spatial Feature Learning (SSFL) module. To combine multiscale temporal features, CSTL is used and the SSFL is suggested to fix the feature corruption problems. CASIA – B and OUMVLP datasets were used and achieved the accuracy of 98.5 % and 98.7 % respectively.

vii) Hou S, et al. [35] described a method named Gait Quality Aware Network (GQAN). It is made up of two blocks namely the Frame Quality (FQBlock) block and the part quality (PQBlock) block. FQblock measures the quality of each frame and evaluates the features of each frame separately. It gives the weights across different silhoutees that makes this method to give more accurate results. The same CASIA – B and OUMVLP data sets are used and the accuracy of 98.5 % and 89.7 % are achieved respectively.

viii) Huang X, et al. [36] introduced STAR – Spatio-Temporal Augmented Relation Network) approach that has two modules namely MDFG – Multi Branch Diverse – Region Feature Generator and STAI - Spatio-Temporal Augmented Interactor. The former module is able to find the body features within separate regions and the later module connects the regions of the frames and forms the intra and inter relation models. The same CASIA – B and OUMVLP data sets are used and the accuracy of 97.3 % and 89.7 % are achieved respectively.

The summary of these reviews is given in Table 1.

**Table 1. Summary of the review of Appearance based Gait Recognition approach research papers**

| Reference | Method | Methodology | Datasets | Accuracy (%) |
|---|---|---|---|---|
| *Zhang Y, et. Al., [29]* | ACLGait | CNN + LSTM | CASIA B | 96.0 |
| | | | OU – LP | 99.3 |
| | | | OU - MVLP | 89.0 |
| *Fan et al. [30]* | Gait Part | CNN + Attention + HP | CASIA B | 96.2 |
| | | | OU - MVLP | 88.7 |
| *Hou S, Cao C, et al. [31]* | GLN | CNN + Lateral Feature Pyramid + HPM | CASIA B | 96.8 |
| | | | OU - MVLP | 89.1 |
| *Hou S, Cao C, et al. [32]* | SRN | RNN + Dual Feature Pyramid + HPM | CASIA B | 97.1 |
| | | | OU - MVLP | 89.1 |
| *Gul S, Malik MI, et al. [33]* | Multi view gait Recognition System | 3D CNN | CASIA B | 98.3 |
| | | | OU – LP | 93.1 |
| *Huang X, et* | CSTL | CNN + MSTE + | CASIA B | 98.5 |

| | | | | |
|---|---|---|---|---|
| *al. [34]* | | ATA | OU – MVLP | 98.7 |
| *Hou, et al. [35]* | GQAN | GQAN + FQBlock + PQBlock | CASIA B | 98.5 |
| | | | OU – MVLP | 89.7 |
| *Huang et al. [36]* | STAR | CNN + MDFG + STAI | CASIA B | 97.3 |
| | | | OU – MVLP | 89.7 |

# 4. Conclusion

This review paper provides an extensive examination of appearance-based methods for human gait recognition. It is still in need for the further improvement. There are still many challenges that include variance in walking patterns, camera view angle, bearing some things, Clothing, etc., the accuracy can be improved by considering these challenges. As a whole, appearance-based gait recognition is having the application in surveillance, health and biometrics, etc.,

## REFERENCES:


1. Huang et al., "Joint segmentation and identification feature learning for occlusion face recognition," IEEE Trans. Neural Networks. Learn. Syst., early access, May 13, 2022, DOI: 10.1109/TNNLS.2022.3171604.

2. W. Zhang, Y. Chen, W. Yang, G. Wang, J.-H. Xue, and Q. Liao, "Class variant margin normalized SoftMax loss for deep face recognition," IEEE Trans. Neural Networks. Learn. Syst., vol. 32, no. 10, pp. 4742–4747, Oct. 2021.

3. Wang L, Ning H, Tan T, Hu W (2004) Fusion of static and dynamic body biometrics for gait recognition. IEEE Trans Circuits Syst Video Technol 14:149–158. https://doi.org/10.1109/TCSVT.2003. 821972.

4. Wu Z, Huang Y, Wang L et al (2017) A comprehensive study on cross-view gait based human identification with deep CNNs. IEEE Trans Pattern Anal Mach Intel 39:209–226. https://doi.org/10. 1109/TPAMI.2016.2545669.

5. Chen J (2014) Gait correlation analysis based human identification. Sci World J 2014:1–8. https://doi. org/ 10. 1155/ 2014/ 168275.

6. Juyoung KIM, Beomseong KIM, heesung Lee, "Temporally Deformable Convolution for Gait Recognition"," IEEE Systems, MAN and Cybernetics Society Section., vol. 13, pp. 6475-6486, January 2025.

7. X. Li, L. Liu, and X. Lu, "Person Re-identification based on elastic projections," IEEE Trans. Neural Networks. Learn. Syst., vol. 29, no. 4, pp. 1314–1327, Apr. 2018.

8. Y. Liu, Z. Li, S. Pan, C. Gong, C. Zhou, and G. Karypis, "Anomaly detection on attributed networks via contrastive self-supervised learning," IEEE Trans. Neural Networks. Learn. Syst., vol. 33, no. 6, pp. 2378–2392, Jun. 2022.

9. M. Yang, Z. Huang, P. Hu, T. Li, J. Lv, and X. Peng, "Learning with twin noisy labels for visible-infrared person re-identification," in Proc. IEEE/CVF Conf. Computer Vis. Pattern Recognition. (CVPR), Jun. 2022, pp. 14288–14297.

10. Wan C, Wang L, Phoha VV (2019) A survey on gait recognition. ACM Computer Surv 51:1–35. https:// DOI. org/ 10. 1145/ 3230633.

11. Huang X, Wang X, He B et al (2023) STAR: spatiotemporal augmented relation network for gait recognition. IEEE Trans Biometric Behavioral

Identity Sci 5:115–125. https:// DOI. org/ 10. 1109/ TBIOM. 2022.3211843

12. Li H, Qiu Y, Zhao H et al (2022) Gait Slice: a gait recognition model based on spatiotemporal slice features. Pattern Recognition 124:108453. https:// DOI. org/ 10. 1016/j. patcog. 2021. 108453

13. Hou S, Liu X, Cao C, Huang Y (2021) “Set residual network for silhouette-based gait recognition”, IEEE Trans Biometric Behavioral Identity Sci 3:384–393. https:// doi. org/ 10. 1109/ TBIOM. 2021. 3074963.

14. wama H, Okumura M, Makihara Y, Yagi Y (2012) The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Trans Inform Forensic Security 7:1511–1521. https://doi.org/10.1109/TIFS.2012.2204253

15. Takemura N, Makihara Y, Muramatsu D et al (2018) “multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. IPSJ Trans Computer Vision Appl 10:4. https:// doi.org/10.1186/s41074-018-0039-6

16. Zhu Z, Guo X, Yang T, et al. (2022). Gait Recognition in the Wild: A Benchmark. IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 14769–14779.

17. Zheng J, Liu X, Liu W, et al. (2022). Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 20228–20237).

18. Rani V, Kumar M (2023) Human gait recognition: a systematic review. Multimedia Tools Appl. https://doi.org/10.1007/s11042-023-15079-5

19. Pearson K (1901) LIII. On lines and planes of closest ft to systems of points in space. Lond Edinburgh Dublin Philosophical Mag J Sci 2:559–572. https://doi.org/10.1080/14786440109462720.

20. Fisher RA (1936) The use of multiple measurements in taxonomic problems. Ann Eugen 7:179– 188. https://doi.org/10.1111/j.1469-1809.1936.tb02137.x

21. Begg RK, Palani swami M, Owen B (2005) Support vector machines for automated gait classification. IEEE Trans Biomed Eng 52:828–838. https://doi.org/10.1109/TBME.2005.845241

22. Rabiner LR (1989) A tutorial on hidden Markov models and selected applications in speech recognition. Proc IEEE 77:257–286. https://doi.org/10.1109/5.18626.

23. Zafar A, Aamir M, Mohd Nawi N et al (2022) A comparison of pooling methods for convolutional neural networks. Appl Sci 12:8643. https://doi.org/10.3390/app12178643

24. Hinton GE, Salakhutdinov RR (2006) Reducing the dimensionality of data with neural networks. Science 313(5786):504–507.

25. Lecun Y, Bottou L, Bengio Y, Hafner P (1998) Gradient-based learning

applied to document recognition. Proc IEEE 86:2278–2324. https://doi.org/10.1109/5.726791

26. Medsker LR, Jain L (2001) Recurrent neural networks. Design Appl 5(64–67):2

27. Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. (2014). Generative adversarial networks. arXiv: 1406.2661.

28. Tran D, Bourdev L, Fergus R, et al. (2015). Learning spatiotemporal features with 3d convolutional networks. In: Proceedings of the IEEE International Conference on Computer Vision

29. Zhang Y, Huang Y, Yu S, Wang L (2020) Cross-view gait recognition by discriminative feature learning. IEEE Trans on Image Process 29:1001–1015. https://doi.org/10.1109/TIP.2019.2926208

30. Fan C, Peng Y, Cao C, et al. (2020). Gait Part: Temporal Part-Based Model for Gait Recognition. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, Seattle, WA, USA, pp 14213–14221

31. Hou S, Cao C, Liu X, Huang Y (2020) Gait lateral network: learning discriminative and compact representations for gait recognition. In: Vedaldi A, Bischof H, Brox T, Frahm J-M (eds) Computer vision— ECCV 2020. Springer International Publishing, Cham, pp 382–398

32. Hou S, Liu X, Cao C, Huang Y (2021) Set residual network for silhouette-based gait recognition. IEEE Trans Biometric Behavioral Identity Sci 3:384–393. https://doi.org/10.1109/TBIOM.2021.3074963.

33. Gul S, Malik MI, Khan GM, Shafait F (2021) Multi-view gait recognition system using spatiotemporal features and deep learning. Expert Syst Appl 179:115057. https://doi.org/10.1016/j.eswa. 2021.115057

34. Huang X, Zhu D, Wang X, et al. (2022). Context-Sensitive Temporal Feature Learning for Gait Recognition. In: Proceedings of the IEEE/CVF International Conference on Computer Vision (pp. 12909–12918).

35. Hou S, Liu X, Cao C, Huang Y (2022) Gait quality aware network: toward the interpretability of silhouette-based gait recognition. IEEE Trans Neural Networks Learn Syst. https://doi.org/10.1109/ TNNLS.2022.3154723.

36. Huang X, Wang X, He B et al (2023) STAR: spatio-temporal augmented relation network for gait recognition. IEEE Trans Biometric Behavioral Identity Sci 5:115–125. https://doi.org/10.1109/TBIOM.2022. 3211843