Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

VIDEO-BASED GAIT ANALYSIS AND CUMULATIVE FOOT PRESSURE IMAGES FOR HUMAN IDENTIFICATION IN VIDEO SURVEILLANCE SYSTEMS

SWAPNA PAVAN G
Research Scholar, Department of Computer Applications, Dr. M. G. R Educational and Research Institute University, Chennai, India.

Dr. N. KANYA
Associate Professor, Department of Information Technology, Dr. M. G. R Educational and Research Institute, Chennai, India.

Dr. V.N RAJAVARMAN
Professor, Department of Computer Science and Engineering, Dr. M. G. R Educational and Research Institute University, Chennai, India.

Abstract

In the past few years, security has become a very important issue. Face and fingerprint recognition are used in most security systems. But there is another type of security that isn't talked about as much: gait recognition. This is the process of figuring out who someone is based on how they walk and how much pressure is put on their feet. In this paper, we'll talk about the ground reaction force, which can be used to tell the difference between how people walk. This research has also shown that using a deep learning approach can improve the rate of gait recognition. In addition to recognition based on video, we look at new modalities, such as recognition based on floor sensors, and new approaches, such as Convolutional Neural Networks (CNN). The work being proposed will be trained on the standard dataset, CASIA-D.

Index Terms: Gait, CNN, Cumulative Foot Pressure, CASIA-D, Deep learning algorithms

1. INTRODUCTION

Use of biometrics and human motion analysis in surveillance systems is a promising field of study. Better identification is achieved via the use of multi-modal biometric systems, which combine data from several biometric modalities. Increased accuracy is achieved when gait is combined with other biometrics in a multi-modal biometric system. Combining gait with face or iris or fingerprint or other biometrics may also be useful in practical applications like visual surveillance and access control. Therefore, collecting data on both gait and total foot pressure is a novel and practical approach to biometric research today.

1.1 Gait Analysis

A person's gait is a distinguishing characteristic that may be utilized in combination with other physical and behavioral characteristics. A person's gait is the manner in which they walk. Multiple studies have shown the feasibility of using gait data to distinguish between individuals. Criminal trials have even made use of gait analysis to determine guilt or innocence. Gait patterns, such as the Parkinsonian shuffle, may be utilised to diagnose and assess health conditions as well as identify individuals. Because it does not need physical contact with the individual, gait recognition is less invasive than other first-generation biometric modalities like fingerprint and iris identification. Everyone has a distinctive gait that can be identified from a biomechanical standpoint, which is the foundation of gait identification. Most people believe that a person's walking style is unique to them and hence difficult to imitate or conceal. Therefore, it serves as a crucial biometric trait for establishing individual identities. Thus, in high-security, civilian, and public settings like airports, stations, banks, and military posts, gait recognition has become a significant aspect of criminal control and detection systems. Different methods of personal identification are possible with it as well. The usage of networked wearable devices has facilitated the widespread adoption of gait recognition for use in applications ranging from gender and age prediction to cyber-physical healthcare.

1.2 Deep Learning for Gait Analysis

Including AI, supervised machine learning is a subset of machine learning. Using a set of inputs and predicted outcomes, mathematical models and algorithms are developed and "trained." The model is taught using a learning algorithm, which may use shallow or deep learning to generate a "machine" capable of doing the specified task. The user directs an investigation of the data structure based on the learned mapping function to assign the hypothesis class [1] in order to evaluate the performance of the models. Linear regression, logistic regression, decision trees, Support Vector Machine (SVM), random forests, naive Bayes, and k-nearest neighbour are examples of shallow learning approaches that depend on human produced features learned in a predetermined relationship between inputs and outputs.

Deep structured learning (or hierarchical learning) derives its concepts from the structure and function of biological brain networks. It starts with the notion of a multi-layer Artificial Neural Network in an effort to train itself how to appropriately represent data (ANN). Deep learning is the optimal strategy when the classification features, if known at all, are complicated and lack a direct mathematical connection to the raw data. Typically, "deep" refers to the number of layers in a network structure, such as a Deep Belief Network (DBN), Feed forward Deep Network (FDN), Boltzmann Machine (BM), Generative Adversarial Network (GAN), Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), or Long-Short Term Memory (LSTM), a type of RNN. This Review is not intended to give an exhaustive study of ANN and deep learning theory. The interested reader is referred to reputable sources [2] for further information. In addition, we focus on models that have practical gait applications, such as CNN and LSTM [3]. The CNN model may be used to any dimension of grid-like data [4, 5, 6].

Using a convolutional operation on the input data, the network is able to learn complicated abstractions and features from massive datasets. The network consists of

ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

convolution layers, pooling layers, and normalisation layers the majority of the time. Shared filters and weights exist between these levels. The output of the convolutional layers is an automatically created feature map from the raw data input. The pooling layers stabilise the output of the convolution layer while reducing the size of the representation. Maximal pooling and average pooling are the most prevalent kinds of pooling layers used by the CNN model. Activation functions (such as Sigmoid, Tanh, ReLU, and Leaky ReLU) use the weight and bias of a neuron to decide whether or not it should fire [5]. This applies to all convolution and pooling layers.

1.3 Cumulative Foot Pressure

The cumulative foot pressure picture provides data on the total force exerted over the course of a single gait cycle in terms of both space and time [6, 7]. Since distinguishing individuals wearing shoes is challenging, this may be a useful adaptation. Camera-based recognition systems have issues with privacy, severe conditions, and backdrops with a lot of people, making them unsuitable for use in circumstances like the security system at a prison, a bathhouse, the entry to public transit, or the front door of a Japanese home. The suggested recognition system that makes use of pictures of foot pressure may be employed in these scenarios in place of a camera-based system. Instead of using a camera to record a person's stride, as is done with gait analysis, a floor pressure sensing device [8] is used to capture a picture of the individual's total foot pressure as they walk. Therefore, neither the lighting nor the camera's perspective shift. Furthermore, as it is not dependent on appearance, cumulative foot pressure photographs may be used to locate the same individual even if they are dressed differently. It might be the best option in certain circumstances. In addition, it doesn't compromise personal privacy since it doesn't need identifying details such as a person's physical characteristics.

1.4 Dataset used

A conventional dataset (Dataset A), a multi-view gait dataset (Dataset B), an infrared gait dataset (Dataset C), and a multi-view gait dataset (Dataset D) make up the CASIA Gait Recognition Dataset (gait and its corresponding footprint dataset). The project is useful in not one but two ways. First, we created a library of photographs of gait and the pressure on the feet (a "footprint") to aid scientists in studying this topic. Gait CASIA Subdataset for Part D.

2. LITERATURE REVIEW

It was suggested by Yan et al. to use a Multilayer Perceptron (MLP) classifier in a CNN model [9]. For the automated extraction of gait characteristics, there is only one GEI that can be depended upon. The CASIA-B data set is used to assess the methods. Using multitask learning, the model is taught to generate several predictions about people.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

Each task is performed with a success rate of 95.88%, although it was observed that the adaptability to new events or views might be enhanced by training on more data.

Zebin et al. [10] suggested a system that would employ five IMU sensors worn on the lower back, thighs, and shins to detect walking-related motions. A CNN-based model is used to automatically extract the features from the raw time-series data in order to improve upon the results obtained by manually constructing the features using shallow learning.

Wan et al. provided an exhaustive summary of investigations using gait recognition (11). We investigate not just video-based identification, but also recognition based on floor sensors, radars, and accelerometers; novel methodology, such as machine learning methods; and the challenges and limits of this topic. The proposed directions for further research are also highlighted. Both seasoned researchers and novices in the field of gait detection may benefit from our review, which displays the current level of knowledge in this domain. It also includes a list of relevant books and resources as a resource for researchers interested in gait recognition.

Using geometry and image processing mathematics, Rafi et al. [12] proposed a model-based method for gait recognition. Feature matrices for gait recognition are generated using segmentation, the Hough transform, and corner detection in this manner. A collection of photographs shot at different periods might reveal a person's walking style. During this step, video sequence frames are sent to the Canny Edge detection algorithm and the Gaussian filter to determine the image's boundaries and reduce noise, respectively. Applying the Hough transform to the pre-processing output in order to retrieve the pertinent data yields a gait model. Utilizing the Harris Corner Detection methodology to find the corners and create the feature points, this method yields gait characteristics. Feature points are used to measure different characteristics of walking in order to gather and analyse gait data. Using a gait recognition interface, the attributes of an unknowing individual are compared to a database template. The proposed technique included analysing data from a database of 10 people and developing a gait recognition system based on five features. All identification criteria are unique, quantitative, and yield recognition results with an accuracy greater than 80% when the camera is oriented at the person at 90 and 270 degrees.

Based on statistical form analysis, Wang et al. [13] presented a simple and successful automatic gait identification method. Using a sophisticated background removal technique, pedestrian silhouettes are extracted from the background of each image series. The progression of the silhouettes is then shown as a collection of interrelated, complex vector configurations in a conventional reference frame. The Procrustes approach of shape analysis is utilised to estimate the median shape, which acts as a gait signature, from various vector configurations. For recognition, techniques for supervised pattern classification based on the complete Procrustes distance measure are used. The method under discussion does not do a direct gait dynamics analysis.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

Instead, it utilises the action of walking to capture biometric form indicators and structural features of gait. The system is tested on a library of 240 sequences, each of which depicts one of 20 individuals wandering outside from one of three potential angles. According on the results of the experiments, the proposed algorithm works effectively.

Wolf et al. [14] developed a three-dimensional convolutional neural network (CNN) that accepts a three-dimensional spatiotemporal tensor as its input. The first CNN channel would display a grayscale image, while the second and third channels would have optical flow. The CASIA-B dataset, the MoBo database, and the UFS database are used for model training and evaluation. Multiple aspects, such as the walkers' speed, the weather, and their apparel, were considered while evaluating the plan.

3. PROPOSED WORK

Here is the proposed research work:

3.1 Gait Recognition System

Gait biometrics has received a lot of attention as a possible solution for authentication and access control [15]. A person's gait may be captured from a great distance. On the other hand, some biometric systems need the subject to physically interact with the device in order to gather biometric data, such as via touching or coming into close proximity with the biometric data collector. This positive quality is crucial, yet it cannot be collected like other biometric traits like fingerprints. It is possible to determine a person's gait with low resolution. Other biometric technologies, such as facial recognition, may not function as effectively in low-quality movies. Gait recognition is required here. A person's gait may be measured using very simple tools. People's gaits may be monitored with the use of cameras, the accelerometers in smartphones, sensors in the ground, or even radar.

Successful applications of deep learning in the areas of safe computing [16], activity identification [17], and voice recognition [18] have emerged in recent years. Additionally, deep learning-based gait identification algorithms, such as support vector machines (SVMs), outperformed their more conventional machine learning-based counterparts [19]. Due of CNNs' superior feature extraction, numerous studies have used them to identify motion or gait [20].

3.1.1 Data Collection

The information came from the CASIA dataset.

3.1.2. Data Acquisition

Each walking sequence's gait data was matched to the correct participant and converted into a 60 by 26 three-dimensional array to satisfy the CNN's input requirements. The column displays all joint points at a given moment in time, while the

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

row displays joint data over time. The gait data were divided into two groups, one for training and one for testing, with a 50:50 split once the input was altered. To increase generality, random distribution of inputs was used.

3.2. Cumulative Foot Pressure

In this study, we propose verification as a recognition problem based on a collection of photographs of foot pressure. Figure 1 depicts an overview of how the proposed pedestrian identification system would function by using photographs of the total foot pressure. When a person walks on a pressure-sensitive floor, a succession of photographs of their feet are captured. Based on the acquired cumulative foot pressure picture collection, there are two primary steps that must be taken to complete the two suggested recognition tasks: The representation and categorization of total foot pressure

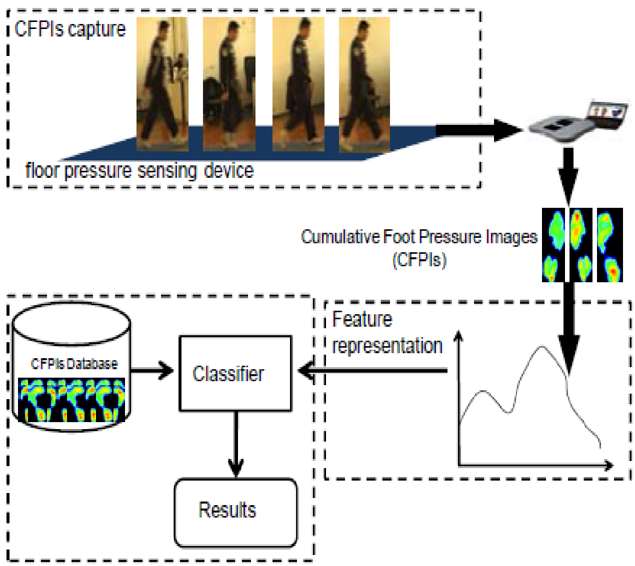

Figure 1: Overview of pedestrian identification system cumulative foot pressure images

The extracted descriptor must not only be able to tell shoes apart, but also stay the same even if shoes change or there is noise. Based on how it is shown, the classification should be strong against changes in shape. Figure 2 shows a hierarchical locality-sparse coding for a cumulative foot pressure image [21].

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

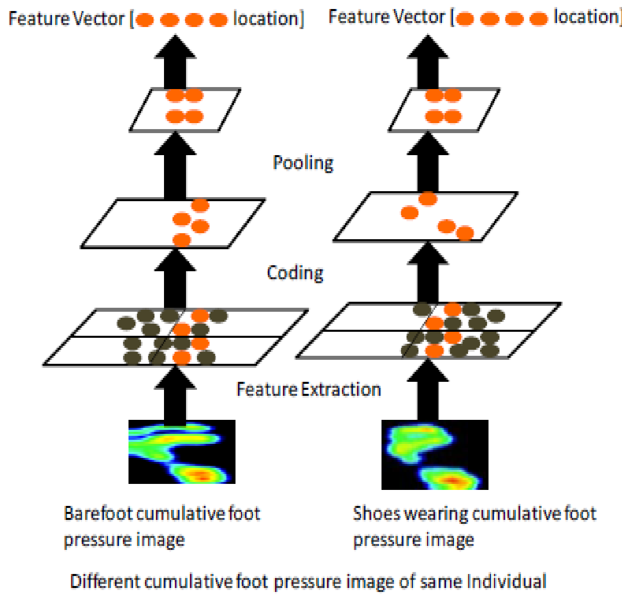

Figure 2: Hierarchical locality-constrain sparse coding scheme

Figure 3 shows a few images from the database of cumulative foot pressure images. We test how reliable the recognition algorithm is in two ways: how walking speed affects it and how different shoes affect it.

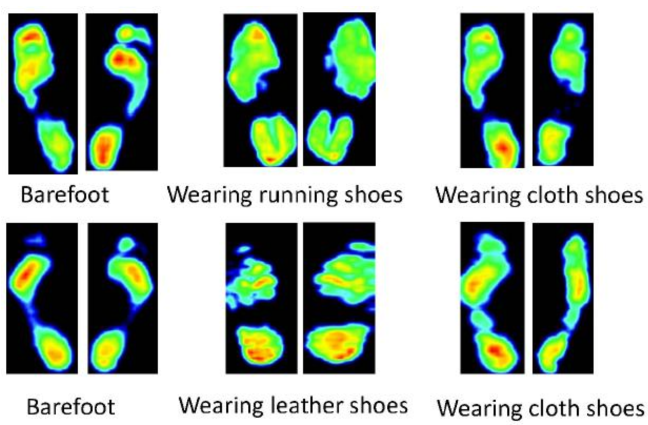

Figure 3: Cumulative foot pressure image under different wearing conditions

When a subject entered the scene, they were instructed to walk. When a subject entered the scenario, he or she was initially instructed to walk five times normally past the pressure sensor. Each individual then had 5x88 cumulative foot pressure photos obtained while walking normally. There are 20 female participants and 66 male ones. All of them are between 20 and 60 years old and are Asian. The exact data guarantee that future tests will not be confined to a single group. After the individuals had walked

ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

normally for a period of time, they were instructed to walk five times fast past the pressure sensor. In this instance, there are 880 cumulative foot pressure recordings (10 x 88). Each record has three photos of cumulative foot pressure, therefore there are a total of 2640 images to consider when considering how walking pace impacts various factors. All of these individuals were instructed to walk barefoot.

In addition, we collect data from walking people wearing diverse footwear to see how effectively the algorithms translate. For a fair comparison, 30 individuals were instructed to walk past a pressure sensor twice barefoot and four times with shoes. There are six images of each person's foot depicting the amount of pressure being applied to them. Thirty individuals were instructed to walk in running shoes, Chinese cloth shoes, and leather shoes. The collection has 30,633 photos of foot pressure in total. Six ladies and twenty-four men are topics. The majority of the individuals are between the ages of 20 and 40, and the majority of them are young.

Each image was saved as a.bmp file, which is a sort of image file. The standard size of an image is 40 by 90 pixels. The database has a large number of individuals, hence it is a large gait database. You may access the NLPR cumulative foot pressure picture dataset at location [22].

Each image was saved as a.bmp file, which is a sort of image file. The standard size of an image is 40 by 90 pixels. The database has a large number of individuals, hence it is a large gait database. The NLPR cumulative foot pressure picture dataset may be found on the website[22].

3.3 Convolutional Neural Network

In this test, a Convolutional Neural Network (CNN) was used. It has 16 hidden convolutional layers, a maximum pooling layer, and fully connected layers. Figure 4 shows a CNN in a way that makes sense. Here, the Rectified Linear Unit (ReLu) enactment work is used in both the convolution and fully connected layers, while the Soft Max initiation capacity is used in the yield layer to figure out how likely each class is. These studies use Keras, which is supported by Tensor Flow, as the framework for deep learning and scikit-learn for order and evaluation. This helps us cut down on how many standard codes we have to write. This also helps us do research faster.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

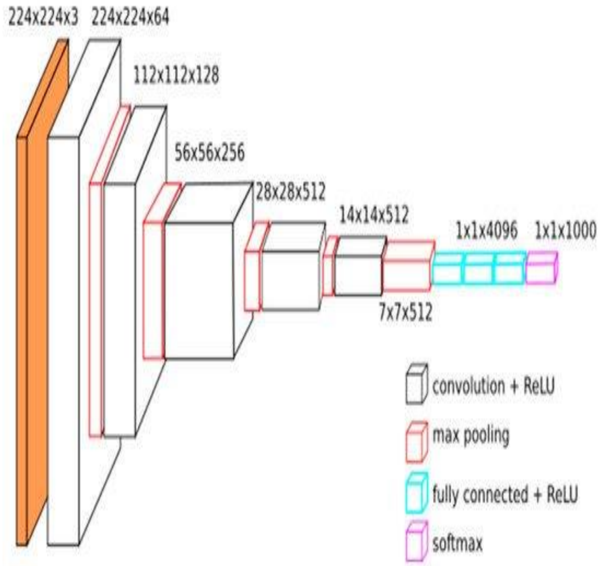

Figure 4: The structure of a Convolutional Neural Network (CNN)

The CNN network had a lot of convolutional levels that were placed in between pooling levels and normalisation levels. At every level of convolution, the convolution between inputs and groups of filters is calculated. Controlled ways of teaching were used. CNN is set up with levels called "Convolutional," "Batch Normalization," "Rectifier Linear Units," "Pooling," and "Completely Linked."

This CNN setup is similar to Alex Net, which is a CNN that has already been trained and has 5 convolutional levels and 3 levels that are completely linked. But this network only uses 4 convolutional levels and a single fully connected level. Also, batch regularisation, ReLU, and pooling levels were placed between convolutional levels.

1. Convolutional Layer

This level was used to mix together the results of descending filters both horizontally and vertically. For every filter, the result of the weights and inputs is calculated, and later, the term "bias" was added [23]. Here, 4 convolutional levels were used, each with its own set of hyper parameters. (1) And (2) show the formulas for figuring out the output dimension of the convolutional level and the number of weights per filter.

$$\text{Output size} = \frac{W_I - F + 2P}{S} + 1 \quad \text{………...(1)}$$

$$\text{Weight} = F \times F \times 3 \quad \text{……………… (2)}$$

In which the input size, the filter size, the padding size, and the number of steps are:

ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

2. Batch Normalization Layer

This level uses every input channel through mini-group plus standardizes this from subtracting mini-group average and splitting from mini-group normal aberration. Later, level moves input by learnable offset and gauges this from learnable scaling factor γ. This accelerates exercise of CNN and decreases sensitivity to network initialization [24].

$$\hat{x} = \frac{x_i - \mu_B}{\sqrt{\sigma_B^2 + \epsilon}} \qquad y_i = \gamma\hat{x} + \beta$$ …. (3)

3. Rectifier Linear Units (ReLU)

This runs the threshold process on the basic input, which changes -ve numbers to zero and keeps +ve numbers. This is also known as a neuron's activation purpose, which determines whether the neuron fires or stays quiet [25].

$$f(x) = \max(0, x)$$ ……………………… (4)

4. Pooling layer

This level functions similarly to down sampling, which splits the input into rectangular pieces for pooling and then calculates the results for each area. This also improves the appearance of the output and prevents local modifications [26]. There are three distinct methods of resource pooling: maximum, minimum, and normal. The output of the convolutional level is partitioned using the ultimate pooling level at this stage. The Supreme degree of pooling determines the greatest possible value in each sector. Formulas exist in for calculating the output dimension of the pooling level (5).

$$\text{Output size} = \frac{W_1 - P_1 + 2P}{S} + 1$$ ……… (5)

In which is input dimension, is pool dimension, is padding dimension and is quantity of strides.

5. Fully Connected Layer

This level works like the classifier level in CNN's ending process, since it catalogues previously mined features by using skilled weighted links. This adds to the input by the weight matrix, and it also adds to the bias vector. This level takes the output of the level before it and turns it into an N-dimensional vector that controls the number of divisions that can be chosen for cataloguing. Since 23 likely subjects were there to be categorised, the output's size and level of expectation were both close to 23.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

4. CONCLUSION AND DISCUSSION

In this paper, it was found that there aren't enough accurate and reliable solutions for biometric systems based on CFPI, so a problem was made. We developed a method called Cumulative Foot Pressure Images (CFPI) with the intention of achieving excellent outcomes and achieving our objectives by creating a decent model using various tools and technologies. This shows that the things that different people brought to our approach are mostly linearly separate. This research has also shown that CNN, a deep learning method that gives high accuracy, can be used to improve the rate of gait recognition. In the future, more work will be needed to put together a lot of datasets.

References

1. O. Simeone, “A very brief introduction to machine learning with applications to communication systems,” IEEE Trans. Cogn. Commun. Netw., vol. 4, no. 4, pp. 648–664, Dec. 2018. doi: 10.1109/TCCN.2018.2881442.
2. Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” Nature, vol. 521 pp. 436–444, May 2015. doi: 10.1038/nature14539.
3. Goodfellow, Y. Bengio, and A. Courville, “Deep learning,” in Deep Learning. Cambridge, MA, USA: MIT Press, 2016, ch. 2, pp. 326–415.
4. N. Ketkar, “Convolution neural network and recurrent neural networks,” in Deep Learning with Python. New York, NY, USA: Springer, 2017, ch. 5, pp. 63–79.
5. N. Yi, C. Li, X. Feng, and M. Shi, “Research and improvement of convolutional neural network,” in Proc. IEEE/ACIS 17th Int. Conf. Computer. Inf. Sci. (ICIS), Jun. 2018, pp. 637–640. doi: 10.1109/ ICIS.2018.8466474.
6. S.P.Moustakidis, J.B.Theocharis, et al, "Subject recognition based on ground reaction force measurements of gait signals." IEEE Trans. on SMC Part B 38: 1476-1485, 2008.
7. S.Zheng, K.Huang, et al, "A Cascade Two-Modality Fusion Scheme for Pedestrian Identification," ACCV, 2010, submitted.
8. D.Tan, S.Yu, et al, "Walker recognition without gait cycle estimation," ICB, 2007.
9. C. Yan, B. Zhang, and F. Coenen, “Multi-attributes gait identification by convolutional neural networks,” in Proc. 8th Int. Congr. Image Signal Process. (CISP), Shenyang, China, 2016, pp. 642–647. doi: 10.1109/CISP.2015.7407957.
10. T. Zebin, P. J. Scully, and K. B. Ozanyan, “Human activity recognition with inertial sensors using a deep learning approach,” in Proc. IEEE SENSORS, Oct./Nov. 2016, pp. 1–3. doi: 10.1109/ICSENS.2016.7808590.
11. ChangshengWan, LiWang, and Vir V. Phoha. 2018. A Survey on Gait Recognition. ACM Comput. Surv. 51, 5, Article 89 (August 2018), 35 pages. https://doi.org/10.1145/3230633
12. Mohamed Rafi, Hédi Khammari, R.S.D Wahidabanu, Yasmeen Taj, A Model Based Approach for Gait Recognition System, International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-5, November 2013 223.
13. L. Wang, T. Tan, W. Hu, and H. Ning, "Automatic Gait Recognition Based on Statistical Shape Analysis," IEEE Transactions on Image Processing, vol. 12, No.9, 2003.

Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)
ISSN：1671-5497
E-Publication Online Open Access
Vol: 41 Issue: 10-2022
DOI 10.17605/OSF.IO/NQ6P8

14. T. Wolf, M. Babaee, and G. Rigoll, “Multi-view gait recognition using 3D convolutional neural networks,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 4165–4169. doi: 10.1109/ ICIP.2016.7533144.
15. S. Sprager and M. B. Juric, “An efficient hos-based gait authentication of accelerometer data,” IEEE T-IFS, vol. 10, no. 7, pp. 1486–1498, 2015.
16. K. Ren, Q. Wang, C. Wang, Z. Qin, and X. Lin, “The security of autonomous driving: Threats, defences, and future directions,” Proceedings of the IEEE, pp. 1–14, 2019.
17. L. Zhao, Q. Wang, Q. Zou, Y. Zhang, and Y. Chen, “Privacy preserving collaborative deep learning with unreliable participants,” IEEE T-IFS, vol. 15, no. 1, pp. 1486–1500, 2020.
18. S. Ji, W. Xu, M. Yang, and K. Yu, “3d convolutional neural networks for human action recognition,” IEEE T-PAMI, vol. 35, no. 1, pp. 221– 231, 2013.
19. J. Man and B. Bhanu, “Individual recognition using gait energy image,” IEEE T-PAMI, vol. 28, no. 2, pp. 316–322, 2006.
20. Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, “A comprehensive study on cross-view gait based human identification with deep cnns,” IEEE T-PAMI, vol. 39, no. 2, pp. 209–226, 2017.
21. J. Wang et al, “Locality-constrained Linear Coding for Image Classification”, CVPR, 2010.
22. CASIA NLPR, “Casia foot pressure image database,” http://www.cbsr.ia.ac.cn/english/, 2011.
23. Math Works Inc. 2-D convolutional layer – MATLAB.https://www.mathworks.com/help/nnet/ref/nnet.cnn.layer.convoluti on2dlayer.html.
24. Ioffe, S., & Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift.
25. Math Works Inc. Rectified Linear Unit (ReLU) layer - MATLAB. https://www.mathworks.com/help/nnet/ref/nnet.cnn.layer.relulayer. html.
26. Math Works Inc. Max pooling layer - MATLAB.https://www.mathworks.com/help/nnet/ref/nnet.cnn.layer.maxpooling2dlayer.html?searchHighlight=maxpooling&s_tid=doc_srchtitle.