International Journal of Science, Environment and Technology, Vol. 9, No 1, 2020, 1 – 14

ISSN 2278-3687 (O)
2277-663X (P)

GAIT RECOGNITION BASED ON DENSENET TRANSFER LEARNING

Xiaofang Wu[1], Tianqi Yang[2] and Zhongchao Xia[3]
School of Information Science and Technology, Jinan University, Guangzhou 510632, China
E-mail: [1]739327953@qq.com, [2]y_tq@163.com, [3]512801539@qq.com

Abstract: Gait recognition is a new biometric recognition technology, which plays an important role in people's life. This paper proposes a gait recognition method which employing Densely connected neural network as the basic algorithm for transfer learning, referred to as DenseNet-based transfer learning. Firstly, this method introduces spatial information of gait through inputting Gait Energy Image (GEI), then it extracts gait features through Dense Net-based transfer learning. Finally, the K nearest neighbor classifier (KNN) is used to classify and identify people. The method is first evaluated on the large public dataset CASIA-B in terms of same-view gait recognition. Experimental results show that it outperforms the other mainstream same-view gait recognition methods by a significant margin, and the average recognition rate can reach 98.86%. In addition, it has good robustness under different conditions. Compared with VGGNet network, the number of the network model parameters of the proposed method is reduced by 448M, which is about 84.85%. These results show that the proposed method effectively improves the speed and quality of gait recognition transfer generated images.
Keywords: Gait recognition; gait energy image; DenseNet; transfer learning.

1. Introduction

Gait recognition is a new biometric recognition technology, which has attracted more and more researchers in recent years. And it is also playing an increasingly important role in people's life. Gait is a biological feature that contains both physiological and behavioral characteristics. Everyone walks differently. Gait recognition is a method to identify people's identity by the way they walk. In short, you can tell who you are by not looking at your face. Compared with other kinds of biometric features such as it is and face, it is particularly suitable for long-distance human identification that the distance of gait recognition can reach 50 meters under the ultra hd camera. Whilst the iris recognition usually needs to recognize the target within 30 centimeters, and the face recognition needs to be within 5 meters. At the same time, gait recognition also has other characteristics, such as difficult to camouflage, fast recognition speed, uncontrolled [1] and so on. Besides, not only it plays an active role in areas such as security and anti-terrorism, smart home, community management, industrial inspection, smart transportation, smart medical care, medical and health care, human-computer interaction, new

*Received Dec 23, 2019 * Published Feb 2, 2020 * www.ijset.net*

retail, and intelligent video surveillance [2], but also it can be widely used in urban scenes such as stations, airports, museums, schools, scenic spots, shopping malls, docks, as well as important infrastructure such as nuclear power plants, power stations, petroleum and petrochemical bases. However, at present, the training of classic neural network models such as VGGNet takes a long time and the speed of generating gait features is slow, making gait recognition still challenging in practical applications. If the training process of neural network model can be accelerated and the occupation of computer resources can be reduced, it will help to put gait recognition into commercial application and bring more convenience to users. This paper proposes a method based on DenseNet transfer learning to extract gait characteristics and classify and identify human identity through nearest neighbor classifier. In addition, experiments were carried out on the large public dataset CASIA-B, and achieved an average recognition rate of 98.86% in the same perspective, which compared with VGGNet-based transfer learning improved by 3.16%. At the same time, the number of model parameters in this method is only 1/16 of VGGNet, which greatly helps to reduce the complexity of the network model and speed up the training and testing process. Although our experiments have not been tested on more data sets at present, the experimental results show that the proposed method effectively improves the speed and quality of gait recognition migration images, and it is better than VGGNet. What’s more, it has wide application prospects in gait-based age estimation [4], human motion recognition [5], complex traffic scene language segmentation [6], and so on.

The innovation of this paper is as follows: proposing a transfer learning method based on DenseNet, which reduce the need for training data and training time in the target domain, and improve the recognition rate; reduce the amount of parameters, speed up model training, save bandwidth, and reduce storage overhead; applying deep learning framework Keras in gait recognition, which is easy to learn and use.

We organize the rest of the paper as follows. In Section 2, the related work is reviewed. In Section 3, the network structure proposed in this paper is described. Finally, in Section 4, the experimental results are shown.

2. Related works

Transfer learning is the forefront of research in machine learning. Andrew Ng, a professor at Stanford University, believes that transfer learning will become the next driving force for the successful application of machine learning in the commercial field after supervised learning [7]. Although we have now entered the era of big data, there is a contradiction between big data and weak computing. Traditional machine learning can only rely on powerful computing capabilities, while transfer learning not only can reduce the use of computing resources, but also quickly transfer the knowledge that has been learned to a new field. In transfer learning, not only the training data and test data do not need to meet the conditions of independent and identical data distribution, but also the model in the target domain does not need to be trained from scratch, which can significantly reduce the need for training data and training time in the target domain [8].

Convolutional neural networks have become the most mainstream method in the field of computer vision. The classic networks of CNN are AlxNet, VGGNet, GoogLenet, ResNet, DenseNet. We can use these classic CNN networks as pre-trained models for transfer learning. In 2012, Alex Krizhevsky et al. proposed AlexNet [9], which applied the basic principles of CNN to a deep and wide network. The entire network has a total of 8 layers, the first 5 layers are convolution layers, and the last 3 layers are fully connected layers which using 11 * 11, 5 * 5, and 3 * 3 convolution kernels. AlexNet is the champion of the ILSVRC2012 competition, and the accuracy rate much better than the second place (Its top5 error rate is 15.3%, and the second place is 26.2%). In recent years, AlexNet-based transfer learning methods have been widely used in process industry image recognition [10], fiat currency recognition [11], and so on.

In 2014, Karen Simonyan et al. at Oxford University proposed VGGNet [12], which is runner-up at ILSVRC2014. It has two structures, VGG16 and VGG19.They are essentially the same, except that the network depth is different. Compared with AlexNet, the improvement of VGGNet is mainly to use several consecutive 3 * 3 small convolution kernels to replace the larger convolution kernel in AlexNet and use a deeper network, so that the network learns more complex models and improves parameter efficiency. In recent years, VGGNet-based transfer learning methods have been widely used in gait recognition [3], clothing picture recognition [13], and so on.

Also in 2014, Christian Szegedy et al. proposed GoogLeNet [14] which was the winner of the ILSVRC2014 competition. GoogLeNet cleverly uses the inception network structure, which not only maintains the sparseness of the network structure, but also utilizes the high computing performance of the dense matrix. Although GoogLeNet has 22 network layers, its number of parameters is about 5 million, only 1/12 of the number of AlexNet parameters and 1/36 of the number of VGGNet parameters. It uses computing resources more efficiently, and extracts more features under the same amount of calculation, which improves the training effect. In recent years, the GoogLeNet-based transfer learning method has been widely used in remote sensing image automatic classification [15] and human behavior classification [16].

In 2015, Kaiming He et al. proposed ResNet [17] which was the champion of ILSVRC2015. One of the biggest highlights of ResNet is that the author added residual blocks through the short circuit mechanism. The residual learning between two layers is studied for shallow ResNet, and the residual learning between three layers is performed for deep ResNet. The proposed residual network reduces a series of problems caused by deep networks, such as the disappearance of gradients. It not only enables deeper networks to be trained, but also improves the accuracy of recognition. In recent years, ResNet-based transfer learning methods have been widely used in human protein atlas image classification [18] and face recognition [19].

In 2016, Gao Huang et al. proposed DenseNet [20]. The basic idea of DenseNet is the same as that of ResNet. Its main contribution is to achieve feature reuse through the connection of features on the channel, and establish dense connections between all the front layers and the back layers. DenseNet has a narrower network width, fewer parameters, and less computational cost.It not only alleviates the vanishing-gradient problem, but also achieves better performance than ResNet. DenseNet also won the Best Paper Award of CVPR 2017.In recent years, DenseNet-based transfer learning methods have been widely used in human motion recognition [5] and complex traffic scene language segmentation [6].

According to the current research status at home and abroad, transfer learning based on CNN classic network has been widely applied in various fields that proving transfer learning is effective. In the field of gait recognition, there are few transfer learning methods based on CNN classic network. Therefore, according to the advantages and disadvantages of each classic CNN network, this paper proposes a gait recognition method based on DenseNet for transfer learning. The experimental details are explained in the third and fourth sections below.

3. Proposed method

DenseNet [20] was proposed by Gao Huang et al. in 2016, which is a convolutional neural network with dense connection characteristics. It introduces direct connections from any layer to all subsequent layers, and connects multiple inputs into a tensor through the $H_l(\cdot)$ function [20], then passes to the next layer of the network. The number of feature maps and the number of parameters are greatly reduced through feature reuse, dividing the network into multiple Bottleneck layers, and introducing Transiton layers etc. These improve the computational efficiency and achieve higher experimental performance. In this paper, DenseNet is used as a feature extractor, then fine-tune the model and obtain the recognition results.

3.1 Input Data

The gait energy image [21] is used as the input of the network. The gait energy image obtains each pixel value in the gait by calculating the average value of the pixels of the silhouette in a gait cycle, which effectively maintains the spatial information in the gait sequence. Not only does it save storage space and computation time, it is also less sensitive to the silhouette noise of a single frame. The calculation method is as follows：

$$\mathrm{GEI} = \frac{1}{N}\sum_{t=1}^{N} B_t \tag{1}$$

Where N is the number of frames in a single gait cycle, and Bt is the silhouette of the t^{th} frame in the gait cycle. The specific calculation process is shown in Fig. 1.

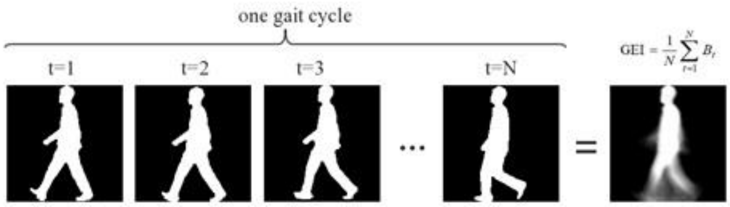

Fig. 1 The calculation process of gait energy image

3.2 Network Structure

The classic convolutional neural network which is density convolutional network (referred to as DenseNet) is used as the basis of our transfer learning. A pre-trained DenseNet model on the ImageNet dataset is loaded and used as a feature extractor. The gait energy image is used as the input of DenseNet to further obtain the gait depth information. In the experiment, there is no need to retrain the model, and the outputs of the fully connected layer and softmax layer can be changed from 1000 categories to 124 categories, so as to adapt to the situation of gait recognition classification problem in this paper, and realize feature transfer. We briefly introduce this network structure in Fig. 2. For more details, may refer to the paper [20]. Where n1,n2,n3,n4 correspond to different values. When [n1,n2,n3,n4] is [6,12,24,16], it is a DenseNet121 structure; when [n1,n2,n3,n4] is [6,12,32,32], it is a DenseNet169 structure; when [n1,n2,n3,n4] is [6,12,48,32], it is a DenseNet201 structure.

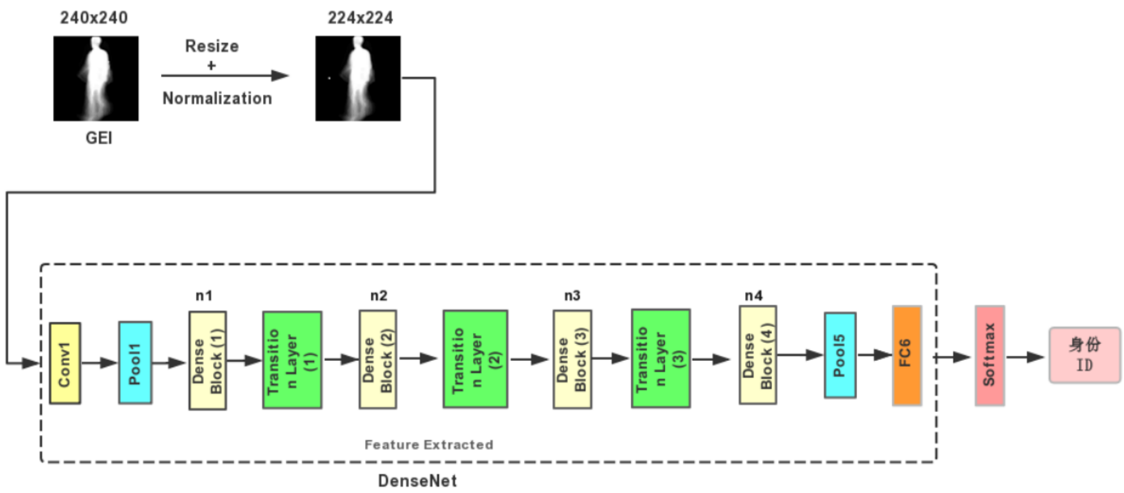

Fig. 2 An illustration of network architecture with 4 dense block

The Dense Block is composed of multiple conv_blocks through dense connections. The input of each conv_block is the connection of all the previous layers. Each conv_block consists of a 1*1 convolution layer and a 3*3 convolution layer. In order to reduce the number of feature maps and improve the calculation efficiency, a convolution layer of 1*1 is introduced. The structure diagram of Dense Block and the structure diagram of conv_block are shown in Fig. 3 and Fig. 4, respectively. In the Dense Block structure diagram, n is the number of conv_blocks, that is, each Dense Block is composed of n conv_block dense connections, where n corresponds to n1, n2, n3 and n4 in Fig.2.

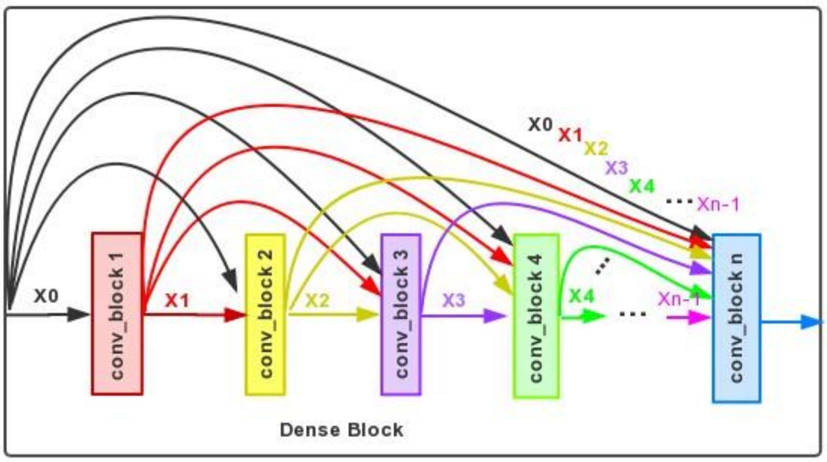

Fig. 3 Dense Block structure diagram, n is the number of conv_blocks, that is, each Dense Block is composed of n conv_block dense connections, where n corresponds to n1,n2, n3 and n4 in Fig.2.

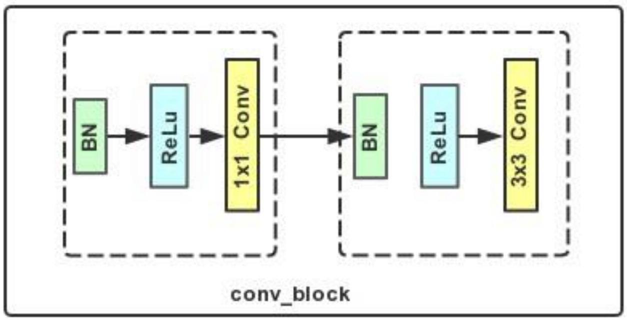

Fig. 4 Conv_block structure diagram

The transition layer consists of a batch normalization layer, a 1*1 convolution layer, and a 2*2 average pooling layer, which is used to reduce the dimensional size of the Dense Block output. The structure of the Transition layer is shown in Fig. 5.

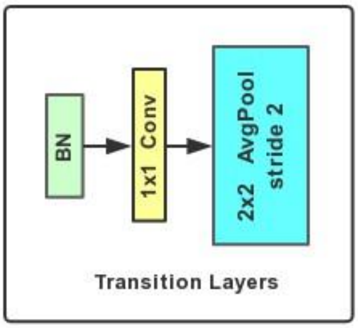

Fig. 5 Transition layer structure

In addition to the above main parts, in front of the first dense block is a 7*7 convolution layer with stride 2 and a maximum pooling layer with 3*3. In the last layer of the network, the softmax function is used to calculate the multi-class cross-entropy loss. In the test phase, the features obtained by DenseNet-based transfer learning are used to classify the identity of the person using the nearest neighbor classifier.

4. Experimental and results

This article uses the Keras open source deep learning framework to build a DenseNet network model for training. The training process includes two stages of training and testing. In this article, experiments I and II were performed and compared with mainstream gait recognition methods.

4.1 DataSet

CASIA Gait Database [22] is a database published by the Chinese Academy of Sciences Automation Research in 2006. It includes a small-scale database Dataset A created in 2001, a multi-view database Dataset B created in 2005, and an infrared database Dataset C. This article adopt the CASIA-B database. CASIA-B [23] is a large-scale, multi-view gait database with a total of 124 subjects, each of whom has 11 views from 0° to 180°, with an interval of 18 degrees. It also provides gait energy image [21] under three different conditions, which are normal walking conditions, carrying a bag condition, and wearing a coat condition. There are 10 sets of gait sequences in total, of which there are 6 groups of gait sequences under normal walking conditions, and both 2 groups under carrying a bag and wearing a coat conditions. Each person has 110 gait sequences. In Fig. 6, we list a sample gait energy image of a person in the CASIA-B database at 11 angles and 3 conditions.

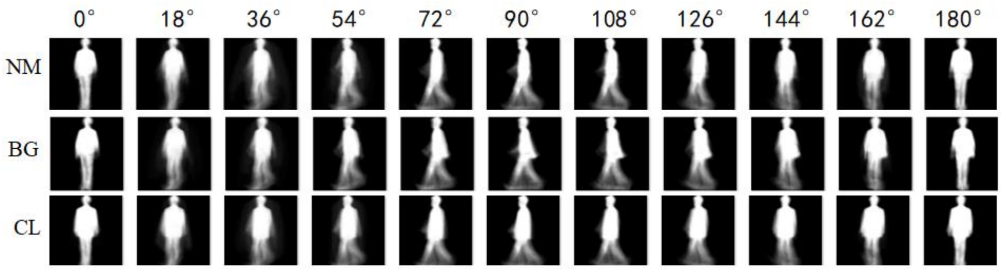

Fig. 6 Sample example of a person's gait energy image at 11 angles and 3 conditions

The image resolution of GEI [21] is 240 * 240 pixels. In this experiment, the data preprocessing uses the data normalization method mentioned in [24]. The pixel value is divided by 255 so that it is in the range [0,1] and resized to 224*224, making it conform to the input picture format of DenseNet network architecture.

4.2 Training

The gait recognition method based on DenseNet transfer learning proposed in this paper is designed and implemented using python script language and keras deep learning framework. The Python version used in this experiment is 3.6.2, and the Keras version is 2.0.5.The main parameters of the computer hardware used in this experiment are as follows: the GPU is GTX 1080Ti GAMINGX with 11G memory, the CPU is AMD Ryzen 5 2600X, the main frequency is 3.6GHz, and the memory size is 32GB.In the experiment, we use the gait sequence of the first 74 subjects as the training set and the last 50 as the test set. All networks are trained using stochastic gradient descent with a batch size of 16 and training of 100 epochs. The other parameters are set to DenseNet [20] trained on the ImageNet [25] dataset. For example, the initial learning rate is 0.1, the weight decay parameter is set to 10^{-4}, and the nesterov momentum [26] is set to 0.9, the dropout rate is set to 0.2.The loss function employs categorical crossentropy. It is worth noting that when using this objective function, the labels need to be converted into a two-valued sequence of the form as (n_samples, n_classes).Therefore, the label y should be adopted a one-hot encoding, converted to an array of the form as (n_samples, 124) with values 0 and 1,where n_samples is the number of gait sequences in the training or testing set.

4.3 Experimental results

In this paper, two experiments were performed to evaluate the rank-1 recognition rate of the CASIA-B dataset. One of the experiments is performed at different angles and different DenseNet network depths under normal walking conditions, the other is performed under three covariate conditions.

4.3.1 Experiment I

In this experiment, under normal walking conditions, the rank-1 recognition rate of CASIA-B data set at 0° -180 ° and three network structures such as DenseNet121, DenseNet169 and DenseNet201 were evaluated, where the gait sequence NM#1-4 serves as gallery and NM#5-6 as probe. Moreover, we compared it with VGGNet-based transfer learning [3] and other traditional methods. The experimental results are shown in Table 1. It can be known from Table 1 that the recognition rate of our proposed gait recognition method of DenseNet-based transfer learning not only exceeds the traditional method, but also exceeds the gait recognition method of VGGNet-based transfer learning. In three network structures, DenseNet201-based transfer learning method has the highest accuracy rate of 98.87%, which is 3.17% higher than VGGNet-based transfer learning [3], and also has fewer parameters than the VGGNet-based transfer learning method. As shown in Fig. 7, DenseNet121 has the least parameters, only 33M, which is about 1/16 of VGG16. In addition, DenseNet201 has the largest number of parameters, only about 1/7 of VGG16, which greatly improve the use efficiency of computer resources.

Table 1: Experimental comparison of Rank-1 recognition rate of CASIA-B data set under the NM condition

Method	0°	18°	36°	54°	72°	90°	108°	126°	144°	162°	180°	Mean
DM-GEI[3]	92.0	86.5	82.0	89.5	92.5	93.5	92.0	93.5	94.5	94.5	96.0	91.5
DM-MGEI	85.5	74.5	76.0	85.5	89.5	91.0	89.0	89.5	87.5	92.0	95.0	86.8
DM-GEnI	92.5	90.5	85.5	87.5	91.5	92.0	94.0	92.0	92.5	95.0	94.3	91.6
DM-FDF	93.0	87.5	82.0	88.5	91.5	93.0	91.5	92.5	94.5	94.0	95.5	91.2
DM-GFI	86.0	79.0	76.5	85.0	90.5	91.5	88.5	90.5	94.5	95.5	96.5	88.5
KSPP[27]	95.2	79.8	70.7	84.7	96.2	96.8	94.1	94.4	92.7	93.6	95.2	90.3
VGR-Net[28]	98.3	99.2	99.2	96.7	97.7	97.1	97.9	97.1	96.3	97.0	97.1	97.6
VGGNet[3]	95.4	91.3	92.1	94.2	97.1	97.7	97.7	96.7	95.8	97.5	97.1	95.7
DenseNet121	100	98.4	97.2	97.2	97.6	97.6	98.0	96.8	96.4	98.8	100	98.0
DenseNet169	99.2	99.2	96.8	97.6	98.0	96.8	98.4	97.2	97.6	98.8	100	98.15
DenseNet201	99.6	99.6	97.2	98.0	98.8	98.8	98.8	98.4	99.2	99.2	100	98.87

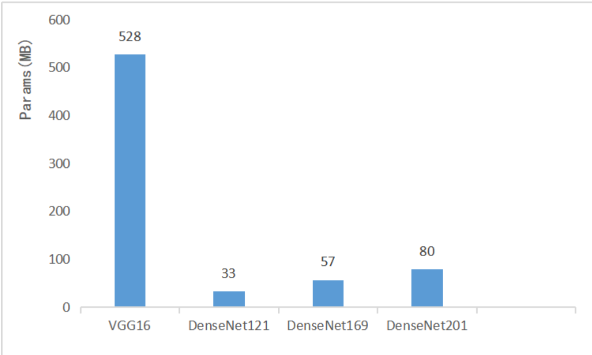

Fig. 7 Comparison of model parameters

4.3.2 Experiment II

In Experiment II, we tested our method under three covariate conditions, namely normal walking conditions, carrying a bag condition, and wearing a coat condition. The settings of the gallery and probe used in this experiment are as follows. Under normal walking conditions, the gait sequence NM#1-4 is used as gallery and NM#5-6 is used as probe, which is also the same as the setting of Experiment I; Under carrying a bag condition, the gait sequence BG#1 is used as a gallery, and BG#2 is used as a probe; Under wearing a coat condition, the gait sequence CL#1 is used as a gallery, and CL#2 is used as a probe. In the experiment, only the recognition rate at 180 ° is tested. The experimental results are shown in Table 2. It can be observed from the table that the recognition rate under the condition of carrying a bag and wearing a coat is not as high as that under normal walking. Both of these conditions have an impact on gait recognition that the effect is less when wearing a coat, and the effect is larger under carrying a bag. The reason is that carrying a bag will occlude part of the gait silhouette, and the features learned by the network are not highly discriminating, which affects the recognition rate.

Table 2: Rank-1 recognition rate (%) comparison of CASIA-B dataset under three covariate conditions

Method	nm-nm	bg-bg	cl-cl
DenseNet121	100.0	85.5	92.0
DenseNet169	100.0	87.1	96.0
DenseNet201	100.0	89.5	98.4

5. Conclusion

In this paper, a gait recognition method based on DenseNet transfer learning is proposed, which uses DenseNet as a feature extractor. The gait energy image is used as a feature to extract gait features, and implements feature transfer. Finally, the nearest-neighbor classifier is used to classify and identify people. This method was tested on the data set CASIA-B, and achieved an average recognition rate of 98.86% in the same perspective. Besides, it was robust to different conditions. The recognition rate of the proposed method is improved by 3.17% compared with the VGGNet-based transfer learning method [3] network. The method in this paper not only reduces the amount of parameters and reduces the storage overhead, but also speeds up the model training speed and improves the recognition rate. In the future, it can be further studied in the following aspects. Further verification and research of the proposed method in other data sets; attempts to fuse DensNet with ResNet to perform feature transfer.

Acknowledgements

This work was supported by the Science and Technology Project of Guangdong (No.2017A0 10101036), the Fundamental Research Funds for the Central Universities (No.21619413).

References

[1] J. Wang, M. She, S. Nahavandi and A. Kouzani: A Review of Vision-based Gait Recognition Methods for Human Identification. *2010 International Conference on Digital Image Computing: Techniques and Applications*(Sydney, NSW, Australia, December 1-3,2010). p. 320-327.

[2] K. Jin, S.C. Chang: Status and Development of Gait Recognition, Computer Science, Vol. 46(2019) No.6A, p.30-34. (In Chinese)

[3] X. Zhang, S. Sun, C. Li, X. Zhao, Y. Hu: DeepGait: A Learning Deep Convolutional Representation for Gait Recognition. *Chinese Conference on Biometric Recognition*(Shen Zhen, China, October 28-29,2017). p.447-456.

[4] A. Sakata, Y. Makihara, N. Takemura, D. Muramatsu and Y. Yagi: Gait-Based Age Estimation Using a Dense Net. *Computer Vision – ACCV 2018 Workshops* (Perth, Australia, December 2–6, 2018). Vol 11367, p.55-63.

[5] J.Q. Zhong: *Human Action Recognition based on Spatial-temporal DenseNet* (MS., South China University of Technology, China 2018), p.1.

[6] B. Jiang, W.X. Tu, C. Yang, H.Y. Liu, Z.L. Zhao: Semantic Segmentation Method for Complex Traffic Scene Based on DenseNet, Pattern Recognition and Artificial Intelligence, Vol.32 (2019) No.5,p. 472-480.(In Chinese)

[7] X.P. Liu, X.D. Luan, Y.X. Xie, M.Z. Huang: Transfer Learning Research and Algorithm Review,

Journal of Changsha University, Vol.32 (2018) No.5,p.28-31+36.(In Chinese)

[8] S.J. Pan and Q. Yang: A Survey on Transfer Learning, IEEE Transactions on Knowledge and Data Engineering, Vol 22(2010) No.10,p.1345-1359.

[9] K. Alex, I. Sutskever, GE. Hinton: Imagenet Classification with Deep Convolutional Neural Networks, Communications of the ACM, Vol. 60 (2017) No. 6, P.84-90.

[10] M. Zhou, Z.Y. Liu, M. Ch. Chen, Y.Y. Zhao, L.J. Yang: Transfer Learning of AlexNet on Image Recognition Application in Process Industry, Industrial Control Computer, Vol. 31 (2018) No. 11, p.80-82.(In Chinese)

[11] S.Y. Liu, J. Huang, G.X. Liu, W.J. Luo: Technology of Multi-Category Legal Currency Identification under Multi-light Conditions Based on AlexNet, China Test, Vol. 45 (2019) No.9, p.118-122.(In Chinese)

[12] K. Simonyan, A. Zisserman: Very Deep Convolutional Networks for Large-Scale image Recognition, *International Conference on Learning Representations* (The Hilton San Diego Resort & Spa, San Diego, May 7 - 9, 2015). arXiv preprint (2014). arXiv:1409.1556.

[13] H.Y. Tang, Z.W. Sun, J. Wang, M. Qian: Clothing image recognition based on VGG-19 hybrid migration learning model, Journal of Xi'an University of Posts and Telecommunications, Vol.23 (2018) No.6.p.87-93.(In Chinese)

[14] C. Szegedy et al.: Going Deeper with Convolutions, *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*(Boston, USA, June 7-12, 2015). p. 1.

[15] B. Chen, H.Z. Wang, X.L. Xu, S.T. Wang, Y.Q. Zhang: The Classification by Medium Resolution Remote Sensing Images Based on Deep Learning Algorithm of GoogleNet Modle, Bulletin of Surveying and Mapping, (2019) No.6,p.29-33+40.(In Chinese)

[16] Y. Bai, H.L. Wang, C.J. Bai: Study on Human Behavior Classification in Still Images Based on GoogLeNet, Computer Knowledge and Technology, Vol.13 (2017) No.18. p.186-188.(In Chinese)

[17] K. He, Z.X.Y. Zhang, S.Q. Ren, J. Sun: Deep Residual Learning for Image Recognition, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Las Vegas, 26 June-1st July,2016), p.770.

[18] C. Chang: Research on Human Protein Atlas Image Classification Method Based on ResNet Deep Network, Journal of Medical Informatcs, Vol. 40 (2019) No.7, p.45-49. (In Chinese)

[19] Y. Bai, H.L. Wang, C.J. Bai: Study on Human Behavior Classification in Still Images Based on GoogLeNet, Computer Knowledge and Technology, Vol.13 (2017) No.18. P.186-188.(In Chinese)

[20] G. Huang, Z. Liu et.al.: Densely Connected Convolutional Networks, *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (Hawaii Convention Center, USA, July 21-26, 2017), p. 4700-4708.

[21] J. Han, B. Bhanu: Individual Recognition Using Gait Energy Image, Pattern Analysis and Machine Intelligence, Vol. 28 (2006) No.2, p.316-322.

[22] Center for Biometrics and Security Research, CASIA on http://www.cbsr.ia.ac.cn.

[23] S. Yu, D. Tan, T. Tan: A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition,*18th International Conference on Pattern Recognition (ICPR'06)*（Hong Kong, China, August 20-24,2006),p. 441-444.

[24] S. Zagoruyko, N. Komodakis: Wide Residual Networks, arXiv preprint arXiv: 1605.07146.

[25] J. Deng, W. Dong, R. Soche, et al.: ImageNet: A Large-Scale Hierarchical Image Database, *2009 IEEE Conference on Computer Vision and Pattern Recognition* (Miami, FL, USA, June 20-25, 2009), p.248-255.

[26] I. Sutskever, J. Martens, G. Dahl, G. Hinton: On the Importance of Initialization and Momentum in Deep Learning, *the 30th International Conference on Machine Learning* (Atlanta, USA, June 16-21, 2013), Vol.28, No.3,p.1139-1147.

[27] K.J. Wang, T. Yan, Z.W. Lv, et al.: Kernel Sparsity Preserving Projections and Application to Gait Recognition, Journal of Image and Graphics, Vol.18(2013) No.3, p.257-263. (In Chinese)

[28] D. Thapar, D. Aggarwal, P. Agarwal, et al.: VGR-Net: A View Invariant Gait Recognition Network, 2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)(Singapore, January 11-12,2018),p.1-8.