Deep Features Based Multiview Gait Recognition

S. Balamurugan[a], Dr. V. Joseph Raj[b], Dr. S. John Peter[c]

[a]Research Scholar and Asst. Prof., Research Department of Computer Science, Kamaraj College, Thoothukudi – 628003, Tamilnadu, India (Affiliated to Manonmaniam Sundaranar University)
[b]Associate Professor and Head, Research Department of Computer Science, Kamaraj College, Thoothukudi – 628003, Tamilnadu, India (Affiliated to Manonmaniam Sundaranar University)
[c]Associate Professor, Research Department of Computer Science, St.Xavier's College (Autonomous), Palayamkottai – 627002, Tamilnadu, India (Affiliated to Manonmaniam Sundaranar University)
[a]psbalamathi2@gmail.com, [b]v.jose8@gmail.com, [c]jaypeeyes@rediffmail.com

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 28 April 2021

Abstract: Pattern recognition is one of the computing intensive application. As information technology has evolved to such an extend, it is easier to identify various patterns in images or videos. As an example consider fingerprint identification, which is image pattern identification application widely used in smartphones to authenticate a person. Similarly Gait recognition, which is a video pattern identification application, is used for surveillance purpose. A lot of research is going on gait recognition. In our purposed deep Convolutional Neural Network (CNN) architecture, Gait Energy image is used as the input and simple several layers are used in matrix analytics. The softmax classifier is used to measure the similarity between Gallery gaits and Probe gaits. This paper uses CASIA Dataset B for performance evaluation of Gait recognition using different angle view variation, normal walking, wearing clothing and carrying bags conditions. The experimental results show that the proposed deep CNN architecture's accuracy is better than Bag-of-Words (BoW), Histogram of Image Gradient (HOG) and 3D-Histogram of Merged Orientations (3D-HOMO)

Keywords: Convolutional Neural Network, Gait Energy Image, Gait Recognition, Multiview

1. Introduction

Human Identification is one of the most significant aspect in social security. Biometrics is used to identify a human by using the physical traits, like fingerprint, face, hand geometry, retina, iris, voice recognition and behaviour traits such as keyboard typing patterns and gait recognition. Human gait, is a behavioural feature, extracted from a raw video sequence which means human walking style. Gait recognition is used for surveillancing private or public space like street, shopping malls, banks, military, airports and so on. Artificial Intelligence (AI) is used in gait recognition, due to complexity of identification. If traditional neural network is used in gait recognition it requires multiple parameter, hence handling of the data training is difficultly. A CNN is one of the most accurate and gives better results than the traditional ones [1]. The CNN, part of Artificial Intelligence is also called as Deep Learning (DL).

2. Related Works

Gait recognition can be divided into two categories, Modal-based and Appearance-Based category. The Model-based approach gait recognition refers human on body structure to extract the gait with static or dynamic parameters such as speed and stride length [2]. G. Shakhnarovich, L. Lee [3] presented the Integrated face and gait recognition from multiple view. L. Wang, H. Ning proposes [4] Fusion of static and dynamic body biometrics for gait recognition. The disadvantages of Model-based approach gait recognition for using high resolution images and computationally expensive.

The Appearance-Based approach gait recognition was working well of low-resolution images and suitable of outdoor or public surveillance. The Appearance-Based model focus on gait sequence or gait of silhouettes on the motion of human body [5]. S.Yu, D. Tan [6] represents, A Framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. WorapanKusakunniran [7] proposes the Recognizing Gaits on Spatio-Temporal Feature Domain. Dr. V. Joseph Raj and S. Balamurugan [8] represent the Survey of current trends in Human Gait Recognition Approaches. S. Balamurugan and Dr. V. Joseph Raj [9] proposes An approach for Gait Recognition using Spatio-Temporal Interest Points based 3D-Histogram of Merged Orientations. Pa. Pa Min, Md. Shohel Sayeed, [10] proposes Gait Recognition Using Deep Convolutional Features. Muif Alotaibi [11] represents the Improved Gait Recognition based on Specialized Deep Convolutional Neural Networks. Mohd Shahrum [12] proposed Convolutional Neural Network (CNN) based Gait Recognition System using Microsoft Kinect Skelton Features. Wang. J., Peng. K [13] represents A Multi-view Gait Recognition Method using Deep Convolutional Neural Network and Channel Attention Mechanism. Suib., Yuz., Xin. [14] proposes Multiview Gait Recognition Based on a Spatial-Temporal Deep Neural Network.

3. Methodology

The Deep CNN architecture proposed in thing, Initially has Gait Energy Image as the input. Next, the deep CNN architecture with 16 layers is used. Finally, the softmax classifier is used to measure the similarity between Training data (gallery) and Testing data (probe).

Proposed Deep CNN Architecture

Gait Cycle Representation

Few GEI samples represented from the CASIA Dataset B [6] are shown below in Figure 1

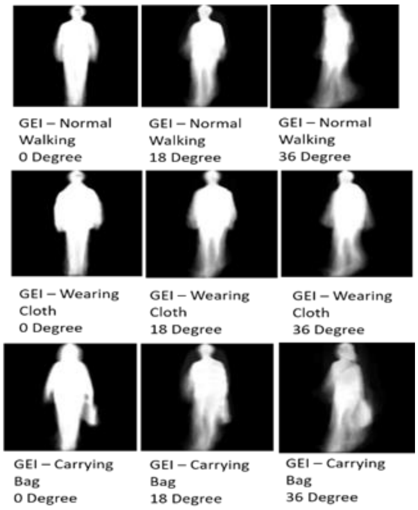

Figure 1: CASIA B – Gait Dataset (Sample Images)

Given the Gait Energy Image (GEI) [15] as the input of proposed CNN architecture. GEI is the gait feature descriptors. GEI is defined by [16]

$$GEI(x, y) = \frac{1}{2}\sum_{t=1}^{N} Gt(\mathrm{x}, \mathrm{y})$$

Where N represents the number of frames per gait cycle

Gt(x,y) represents the binary silhouette of subject at time t.

The silhouettes are normalized in fixed smaller size (160 * 160 * 1) and to consider training process speed up to the Convolutional Neural Network.

In Figure 2, the proposed Deep CNN Architecture, uses 16 layers: one Image Input Layer, four convolutional layers, four Relu (Rectifier Linear Units) layers, four maxpooling layers, two fully connected layers, one softmax classification layer.

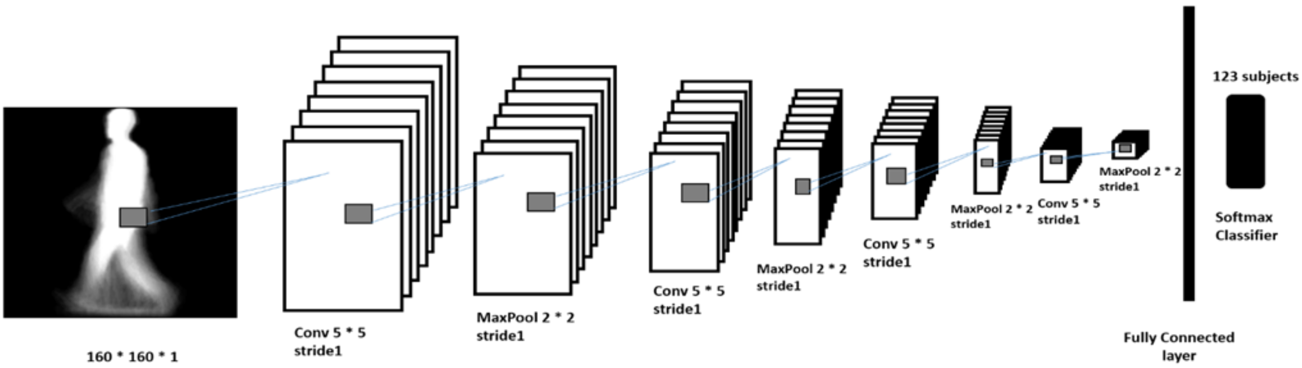

Figure 2: Proposed Deep CNN Architecture

The convolutional layer is used to convolve the images by using the horizontally and vertically filters. Convolutional layer uses the filter size matrix (5 * 5) and applied with stride=1, number of filters eight. The output of convolutional layer feature map is applied on relu Layer. In relu layer, obtained a threshold operation to the pixels of the input images and to transforms negative values to zero and also considering the positive values. It represents an activation function f(x)=max(0,x). The output of relu Layer feature map is applied on maxpooling layer. The maximum pooling layer obtained the maximum values of each region. In maxpooling layer, the pooling size matrix (2 * 2) is used to extract the maxpooling feature map. The output of maxpooling feature map is applied on the fully connected layer. The output of fully connected layer features map to extract the softmax features map. The softmax features map is applied on finally classification layer. The back propagation learning algorithm used these layers is trained. In our proposed Deep CNN architecture model, each feature map is connected to one feature map from the previous layer. The Stochastic gradient Decent with Momentum (SGDM) algorithms used as an optimizer with an initial learning rate of 0.001. In each training iteration the maximum number of epochs and size of mini-batch with 30 and 8 respectively. Finally, the softmax classifier is used to measure the similarity between training data (gallery) and testing data (prob) in the proposed Deep learning architecture.

4. Experimental Results

The Institute of Automation Chinese Academy of Science (CASIA) produced the CASIA Gait Dataset of gait recognition. There are three datasets in CASIA Gait Dataset. CASIA Gait Dataset A, CASIA Gait Dataset B and CASIA GAIT Dataset C. In this paper, CASIA Gait Dataset B is selected. The CASIA Dataset B [6] is a multiview gait dataset used to compare the proposed deep CNN architecture with other algorithms.

In the proposed deep CNN architecture, 123 subjects in CASIA Gait Dataset B are selected. Each subject contain, ten walking sequences and consider three types of walking styles such as six times repeatedly for normal walking, two times for wearing cloth and two times for carrying bag conditions. Each subject captured from three different view angles such as 0, 18, and 36.

In the proposed deep CNN architecture, three sets of experiment A, experiment B and experiment C are conducted. In each experiment the Gait Energy Image is divided into Training data (Gallery) and Testing data (Probe). For experiment A set, normal walking of first four sequence for training data and normal walking of two sequence for testing data was used. For experiment B set, normal walking of first four sequence for training data and wearing cloth walking of two sequence for testing data was used. For experiment C set, normal walking of first four sequence for training data and carrying bag walking of two sequence for testing data was used.

The Recognition Rate (%) obtained by the proposed deep CNN architecture of Experiment A set is given in below Table 1 and Figure 3

Table 1: Experiment A for Recognition Rate (%) Analysis

Experiment A (Training [nm] - Testing [nm])					
Gallery (Training)	Probe (Testing)	Recognition Rate (%)			
(normal walking)	(normal walking)	BoW	HOG	(3D-HOMO)	Proposed CNN Architecture (Accuracy)
0^o	0^o	0.912	0.9029	0.9903	0.9994
0^o	18^o	0.395	0.1068	0.3738	0.4981
18^o	0^o	0.239	0.1165	0.4757	0.9956
18^o	18^o	0.91	0.8689	0.9951	0.9992
18^o	36^o	0.397	0.199	0.6845	0.9871
36^o	18^o	0.388	0.1942	0.699	0.9911

36º	36º	0. 923	0.8 689	0.99 03	0.999

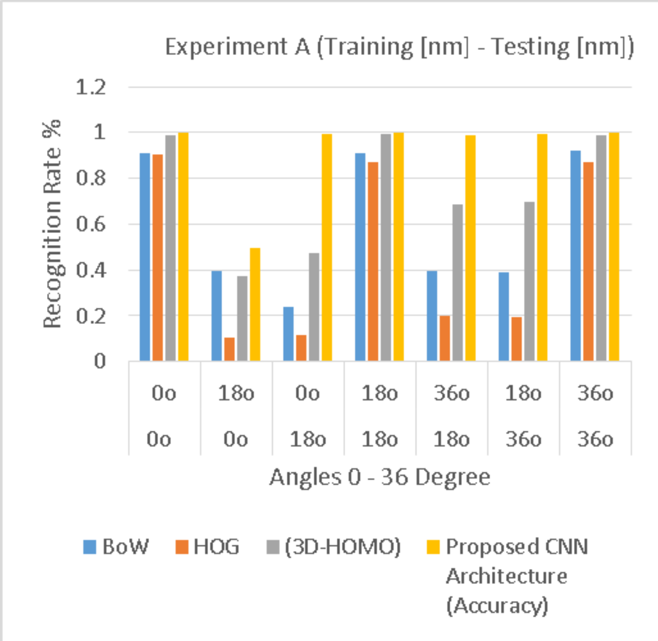

Figure 3: Experiment A for Recognition Rate (%) Analysis

From this analysis, it is found that the proposed deep CNN architecture features accuracy has improved the level of recognition rate of 18% (ie, 92%) compared to Bag-of-Words, Histogram of Image Gradient and 3D-Histogram of Merged Orientations with walking style sequence of Training data normal walking (nm) and Testing data normal walking (nm).

The Recognition Rate (%) obtained by the proposed deep CNN architecture of Experiment B set is given below in Table 2 and Figure 4

Table 2: Experiment B for Recognition Rate (%) Analysis

Experiment B (Training [nm] - Testing [cl])					
Gallery (Training)	Probe (Testing)	Recognition Rate (%)			
(normal walking)	(carrying bag)	BoW	HOG	(3D-HOMO)	Proposed CNN Architecture (Accuracy)
0º	0º	0.246	0.2233	0.6311	0.9868
0º	18º	0.056	0.0583	0.0874	0.9848
18º	0º	0.065	0.068	0.3155	0.9846
18º	18º	0.29	0.3058	0.7184	0.9862
18º	36º	0.157	0.0922	0.4903	0.9856
36º	18º	0.081	0.0874	0.4709	0.9849
36º	36º	0.343	0.2767	0.6553	0.9864

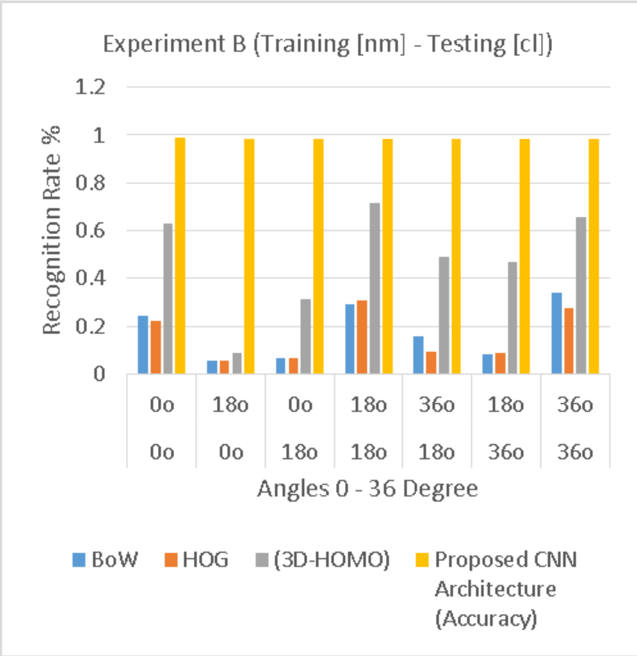

Figure 4: Experiment B for Recognition Rate (%) Analysis

From this analysis, it is found that the proposed deep CNN architecture features accuracy has improved the level of recognition rate of 50% (ie, 98%) compared to Bag-of-Words, Histogram of Image Gradient and 3D-Histogram of Merged Orientations with walking style sequence of Training data normal walking (nm) and Testing data wearing cloth walking (cl).

The Recognition Rate (%) obtained by the proposed deep CNN architecture of Experiment C set is given in below Table 3 and Figure 5

Table 3: Experiment C for Recognition Rate (%) Analysis

Experiment C (Training [nm] - Testing [bg])					
Gallery (Training)	Probe (Testing)	Recognition Rate (%)			
(normal walking)	(carrying bag)	BoW	HOG	(3D-HOMO)	Proposed CNN Architecture (Accuracy)
0^o	0^o	0.794	0.5146	0.8689	0.99
0^o	18^o	0.184	0.0437	0.2718	0.9854
18^o	0^o	0.151	0.0777	0.4272	0.9852
18^o	18^o	0.738	0.2961	0.5922	0.9893
18^o	36^o	0.399	0.1602	0.6117	0.9853
36^o	18^o	0.226	0.0825	0.3932	0.987

36º	36º	0.77	0.4 272	0.84 47	0.9887

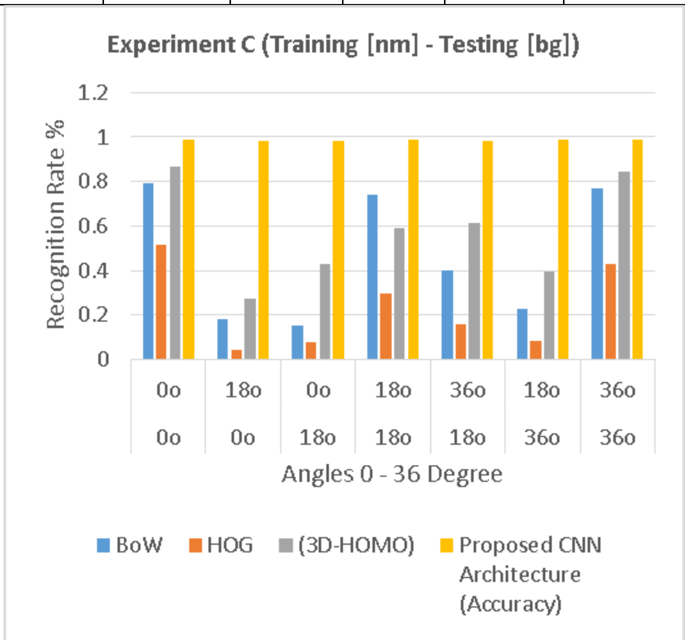

Figure 5: Experiment C for Recognition Rate (%) Analysis

From this analysis, it is found that the proposed deep CNN architecture features accuracy has improved the level of recognition rate of 41% (ie, 98%) compared to Bag-of-Words, Histogram of Image Gradient and 3D-Histogram of Merged Orientations with walking style sequence of Training data normal walking (nm) and Testing data carrying bag walking (bg).

5. Conclusion

This research proposes a deep CNN architecture based on the appearance-based gait recognition. In the architecture we used the simple convolutional neural network with different activation functions to find the classification deep features accuracy and better computational speed. From the experiment results, the deep CNN architecture features accuracy has significant improvement over other gait recognition techniques. In future, the architecture design in the deep CNN architecture can be improved.

References

1. Lawrence S., Giles, C.L., Tsoi, A.C., Back, A.D. (1997), Face recognition: A convolutional neural-network
2. approach, IEEE Transaction on Neural Networks, 8(1), 98 - 113
3. K. Bashir, T. Xiang and S. Gong, “Gait recognition using gait entropy image,” 3rd international Conference
4. on Imaging for Crime Detection and Prevention (ICDP 2009), 2009
5. G. Shakhnarovich, L. Lee, and T. Darrell Integrated face and gait recognition from multiple views. In Proc.
6. of the 2001 IEEE computer Society Conference on Computer Vision and Pattern Recognition, pages 1439-
7. 1446, Hawaii, USA, Dec. 2001.
8. L. Wang, H. Ning T. Tan and W. Hu. Fusion of static and dynamic body biometrics for gait recognition. IEEE
9. Transactions on Circuits and Systems for Video Technology, 14(2): 149-158, Feb 2004.
10. T. H. Lam, K. Cheung, and J. N. Liu, “Gait flow image: A silhouette-based gait representation for human
11. identification, “Pattern Recognition, vol. 44, no. 4, pp, 973-987, 2011.
12. S. Yu, D. Tan, and T. Tan, “A Framework for evaluating the effect of view angle, clothing and carrying
13. condition on gait recognition,” in Proce. IEEE Int. Conf. pattern Recognit. (ICPT), Aug. 2006, pp. 441-444
14. WorapanKusakunniran, “Recognizing Gaits on Spatio-Temporal Feature Doman,” IEEE Trans. Inf.
15. Forensics and Security, vol. 9, no. 9, pp. 1416-1423, Sep. 2014.

16. Dr. V. Joseph Raj and S. Balamurugan, “Survey of current trends in Human Gait Recognition Approaches”
17. Int. J. Adv. Res., 5(10), Oct. 2017, pp. 1851-1864
18. S. Balamurugan and Dr. V. Joseph Raj, “An approach for Gait Recognition using Spatio-Temporal Interest
19. Points based 3D-Histogram of Merged Orientations”. Test Engineering Management, ISBN: 0193-4120 pp.
20. 13787-13794 Jun.2020.
21. Pa. Pa Min, Md. Shohel Sayeed, Thian Song Ong, “Gait Recognition Using Deep Convolutional Features”,
22. 7th Int. Conf. IEEE 10.1109/ICoICT.2019.8835194, Sep 2019
23. Muif Alotaibi and Ausif Mahmood, “Improved Gait Recognition based on Specialized Deep Convolutional
24. Neural Networks”, Computer Vision and Image UnderStanding, Vol 164, pp. 103-110, Nov 2017
25. Mohd Shahrum Md Guntor, Rohilak Sahak, “Convolutional Neural Network (CNN) based Gait Recognition
26. System using Microsoft Kinect Skeleton Features” Int. J. Eng. & Tech. on vol 7, No 4.11, pp. 202-205, 2018
27. Wang. J., Peng. K. “A Multi-view Gait Recognition Method using Deep Convolutional Neural Network and
28. Channel Attention Mechanism”, Com. Mod. In Eng. & Sci., vol 123, No. 1, pp. 345-363, 2020
29. Suib., Yuz., Xin., “Multiview Gait Recognition Based on a Spatial-Temporal Deep Neural Network”, IEEE
30. Access, vol. 6, pp. 57583 – 57596, Oct. 2018.
31. Han, J; Bhanu, B., “Individual recognition using gait energy image”, in Pattern Analysis and Machine
32. Intelligence, IEEE Transactions on vol 28, no. 2, pp. 316-322, Feb 2006
33. L. Lee and W. E. L. Grimson, “Gait analysis for recognition and classification,” in Proc. 5th IEEE Int. Conf.
34. AFGR, pp. 148-155 2002.