

An Advanced Attention-Enhanced Hybrid Deep Learning Model for WBSNs' Gait Pattern Recognition

Jianning Wu (✉ jianningwu@fjnu.edu.cn)

Fujian Normal University

Qiaoling Tan

Fujian Normal University

Xiaoyan Wu

University of Bristol

Research Article

Keywords: Gait pattern recognition, Wireless body sensor network, Hybrid deep learning model, Attention mechanism, Gait variability

Posted Date: January 13th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1251973/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

An Advanced Attention-enhanced Hybrid Deep Learning Model for WBSNs' Gait Pattern Recognition

Jianning Wu^{1*}, Qiaoling Tan¹, Xiaoyan Wu²

*Correspondence:

jianningwu@fjnu.edu.cn
college of Computer and
Cyber Security, Fujian
Normal University, People's
Republic of China
Full list of author
information is available at
the end of the article

Abstract

Background: The deep learning techniques have been attracted increasing attention on wireless body sensor networks (WBSNs) gait pattern recognition that has a great contribution to monitoring gait change in clinical application. However, in existing studies, there are some challenging issues such as low generalization performance and no potential interpretation for gait variability. It is necessary to search for the advanced deep learning models to resolve these issues.

Method: A public WARD database including acceleration and gyroscope data acquired from each subject wearing five sensors was selected, and the gait with different combination of on-body multi-sensors is considered as a WBSNs' gait pattern. An advanced attention-enhanced hybrid deep learning model of DCNN and LSTM for WBSNs' gait pattern recognition was proposed. In our proposed technique, the combination model of DCNN with LSTM is firstly to discover the spatial-temporary gait correlation features. And then the attention mechanism is introduced to exploit the more valuable intrinsic nonlinear dynamic correlation gait characteristics associated with gait variability hidden in spatial-temporary gait space obtained. This significantly contributes to enhancing the generalization performance and taking insight on gait variability in a certain anatomical region.

Results: The ten gait patterns are randomly selected from WARD database to evaluate the feasibility of our proposed method. Our experiments demonstrated the superior generalization ability of our method to some models such as CNN-LSTM, DCNN-LSTM. Our proposed model could classify ten gait patterns with the highest accuracy and F1-score of 91.48% and 91.46%, respectively. Moreover, we also found that the classification performance of a certain gait pattern was almost same best when the combinations of three or five on-body sensors were employed respectively, suggesting that our method possibly take insight on gait variability in a certain anatomical region.

Conclusion: Our proposed technique could feasibly discover the more intrinsic nonlinear dynamic correlation gait characteristics associated with gait variability from on-body multi-sensors gait data, which greatly contributed to best generalization performance and potential clinical interpretation. Our proposed technique would hopefully become a powerful tool of monitoring gait change in clinical application.

Keywords: Gait pattern recognition, Wireless body sensor network, Hybrid deep learning model, Attention mechanism, Gait variability

Background

In recent 20 year, wireless body sensor networks(WBSNs), as an emerging wearable sensing technology for easily acquiring gait data during freely walk in-home or outdoor setting [1-2], has been popularly applied in gait analysis [3-5]. Usually, WBSNs-based gait pattern can be constructed by different sensors node that are located on the different anatomical regions of body such as head, shoulder, elbow, arm, wrist, waist, hip, thigh, knee, ankle, heel, foot and so on. Many studies have found that automatic recognition of WBSNs gait pattern has greatly contributes to accurately evaluate the human gait function change in daily life [6-9], which benefits clinical diagnosis, rehabilitation assessment and early prediction of fall risk of elderly, and so on. In recent studies, automatic recognition of WBSNs gait pattern has been considered as a gait classification task, and the challenging issue is how to achieve the best generalization performance and potential interpretation for gait variability[8-9].

As we know, a key step in developing gait classification model with high quality is to capture more valuable intrinsic gait features information associated with gait variability hidden in gait data. The same is true with how WBSNs gait classification model with best quality is constructed [10-11]. Recent studies have shown that some deep learning algorithms with excellent ability to learn representative features have been applied in WBSNs' gait pattern recognition. These algorithms mainly include convolutional neural network (CNN), dilated convolutional neural network (DCNN), recurrent neural network (RNN) and long short-term memory(LSTM) [12-15]. Their basic idea is to take advantage of superior learning ability to exploit the spatial or temporary correlation gait features hidden in multi-sensors data for improving generalization performance. Especially, DCNN and LSTM algorithms have more and more been attracted attention on the related studies because they have strong capability to capture the intrinsic spatial or temporal dependency gait features. Despite these algorithms can obtain better classification performance, they hardly capture the more valuable gait characteristics information from multi-sensors data. This is because these algorithms possibly impede the efforts on the valuable intrinsic spatial-temporal correlation gait features in WBSNs gait classification task.

Recently, the different efforts on the quantitative analysis of multi-sensors gait data have been devoted to discovering the intrinsic spatial-temporal correlation gait feature information embedded in WBSNs' gait pattern, in order to produce the best generalization ability. Some recent studies have concentrated on the application of data fusion-based feature extraction technique to exploit the valuable implicit spatial-temporal correlation gait features information from both inter-sensors and intra-sensors data in WBSNs [16-18]. A typical example is that the distributed compressed sensing (DCS) technique is utilized to simultaneously compress multi-sensors gait data for capturing spatio-temporal correlation information hidden in WBSNs' gait pattern [16]. However, the DCS technique difficultly discovers the valuable intrinsic gait feature from non-sparsity gait data of multi-sensors dynamic sequence because it must satisfy data joint

sparsity. Meanwhile, the others recent studies have demonstrated that many endeavors have been made to construct the deep learning-based WBSNs' gait classification model with high quality [19-22]. Their basic idea is to take advantage of the superior ability of learning representation feature of deep learning algorithms to discover the intrinsic spatial-temporal interaction information embedded in WBSNs' gait pattern in an end-to-end manner. Some existing studies have showed that the feasibility of application of hybrid deep learning algorithms for capturing the more intrinsic spatial-temporal correlation information hidden in on-body multi-sensors gait data, in order to enhance the generalization ability of dynamic sequence gait data [19-21]. For example, F. J. Ordoez et al. proposed a hybrid deep learning model of CNN and LSTM to automatically learn the implicit spatial-temporal gait features from multi-sensor dynamic data for best classification performance. Unfortunately, there exist some limitations in recent studies. Firstly, these existing studies have ignored the developed WBSNs gait classification model for potential interpretation for gait variability, that is, they can not accurately find the best location of on-body sensors associated with gait change for potential interpretation. Secondly, it is well-known that the human gait is generated by the interaction between the central nervous system, peripheral nervous system, and musculoskeletal effectors system. In other words, human gait depends on the high interaction between central nervous system and various muscles which allow the individual keep a continual interchange between mobility and stability. There exists the intrinsic characteristics of the complex dynamic in gait that is derived from the locomotors sensory feedback from muscles, joints, and other receptors related to locomotion change. The intrinsic "interesting" gait characteristics information on local nonlinear dynamic is possibly hidden in the interaction among multi-sensors in WBSNs' gait pattern. However, these existing studies do not take into account the intrinsic valuable characteristic highly associated with gait variability such as the local nonlinear dynamic characteristics embedded in spatial-temporal correlation of human gait. How to discover these intrinsic valuable "interesting" gait characteristics for the best generalization ability and potential interpretation for gait variability based on on-body multi-sensors data has been a challenge endeavor in the recent existing studies [20-22]. Theoretically, it is possible that valuable characteristics with gait variability are locally and saliently distributed in the intrinsic non-linear dynamics correlation of human movement. It is critical to exploit the local intrinsic non-linear dynamics correlation from on-body multi-sensors data, which greatly contributes the best generalization performance of gait classification. This motivates us to search for the advanced hybrid deep learning models that can discover the valuable local intrinsic "interesting" feature from multi-sensors gait data for accurately detecting WBSNs' gait pattern change.

To our best knowledge, attention mechanism has a superior ability to focus on the salient parts of the object rather than the whole object. Recently, the attention mechanism has been popular in deep learning algorithms for exploiting valuable implicit features hidden in the dynamic data. The representative examples include attention-based RNN model and attention-based LSTM model, which improve intrinsic time correlation in dynamic time-series data [23-24]. Some recent studies have showed the feasibility of application of attention-enhanced LSTM network to identify Parkinson' disease(PD) characteristics using multi-sensors GRF gait data on foot[25]. However, no

studies have been reported that attention-based hybrid deep learning model is developed to discover local intrinsic non-linear dynamics correlation embedded in spatial-temporal correlation of gait features using on-body multi-sensors data. Therefore, inspired by deep fusion of excellent learning performance of attention mechanism, DCNN and LSTM, we proposed a novel attention-based hybrid deep learning model of DCNN and LSTM to capture the local intrinsic non-linear dynamics correlation information associated with gait variability hidden in WBSNs' gait pattern defined by the combination of different on-body multi-sensors. That is, with on-body multi-sensors gait data, we take advantage of the attention mechanism to exploit the local implicit nonlinear dynamic correlation gait features containing more valuable gait variability information embedded in spatial-temporary gait correlation space obtained by DCNN and LSTM, which greatly contribute to the best generalization performance and potential interpretation for gait variability based on different anatomical regions of body. In this study, an open wearable action recognition database (WARD) [26] including acceleration and gyroscope data from all five on-body sensors of each subject was employed to evaluate the feasibility of our proposed model. The experimental results demonstrated that our proposed model can achieve the state-of-art generalization performance and find different anatomical regions for potential interpretation on gait variability. The main contribution of this study are summarized as follow:

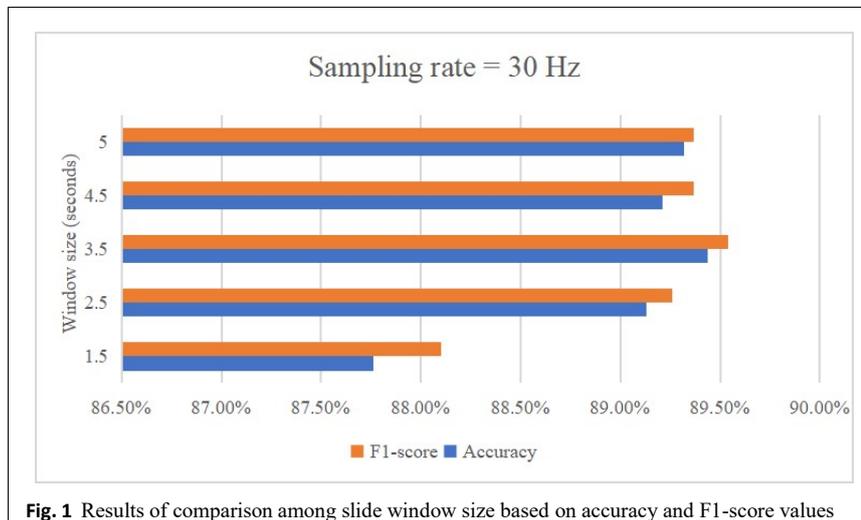
1. We proposed a novel attention-based hybrid deep learning model of DCNN and LSTM for WBSNs' gait pattern recognition with high quality.
2. Our proposed model can discover the valuable local intrinsic “interesting” characteristics associated with gait variability such as nonlinear dynamic correlation features, which would contribute to searching for specific details of pathological gait deviations.
3. Our proposed technique can find the best combination location of different on-body sensors associated with anatomical regions, which could provide potential gait variability interpretation for gaining deeper insight into gait compensatory mechanism.

Results

The selection of proper fixed-length of sliding window

In this study, with the selected accelerometer and gyroscope data from multi-sensors, we objectively evaluate the effects of the selection of best parameters on our proposed model. These selected parameters mainly include of sliding window size, dilated rate, batch size and epoch size on the classification performance. Firstly, we evaluated the impact of the selection of the proper the fixed-length of sliding window on the classification performance. In view of the normal gait speed (1.1-1.5 m/s) and gait cycle (1-1.3s), we concentrated on the selection of the proper fixed-length sliding window from window length range from 1.5s to 5s. In this experiment, our proposed model consist of seven dilated convolution layers, two long short-term memory layer, attention strategy and softmax layer. To decrease the training time, the time of epoch was set to 60. Accuracy and F1-score were used as evaluation criteria of the classification performance. The results of effect of sliding window size on classification

performance were presented in Fig. 1. From Fig. 1, we obviously observe that both F1-score and accuracy increase while the window length increases from 1.5s to 3.5s, whereas both F1-score and accuracy decrease when the windows length increases from 1.5s to 3.5s, whereas both F1-score and accuracy decrease when the window length increases from 3.5s to 5s. In comparison, both F1-score and accuracy can, respectively, achieve the maximum values of 89.54% and 89.44% when the window length is fixed to 3.5s. Both F1-score and accuracy can, respectively, obtained the minimum values of 88.1% and 87.8% when the window length is fixed to 1.5s. These results demonstrated that data segmentations obtained by sliding window with length of 1.5s possibly contain the less valuable information about gait change, which could deteriorate the generalization performance of our proposed classification model. Similarly, data segmentations yielded by sliding window with length of 4.5s or 5s can retain the more redundancy information hidden in gait data, which could worsen the classification performance. Fortunately, data segmentations produced by sliding window with length of 3.5s can contain the more valuable discriminative information in gait, which significantly improve the generalization ability.



The selection of best dilated rates

In this study, it is very important to select the best dilated rate for exploiting the valuable “interesting” gait features. As we know, dilated rate is set to 8 and 16 corresponding to receptive field of 31 and 63, respectively. Theoretically, the receptive field size with dilated rate size of 16 is closed to the input data size of 105 in our proposed model, which could obtain the more useful information about gait change. In this experiment, for comparison, we designed the two different schemes of selection of dilated rate. The first scheme is that the dilated rates are set to 1, 2, 4, 8, 16, 2, 4, 8, 16, respectively; and the second scheme is that the

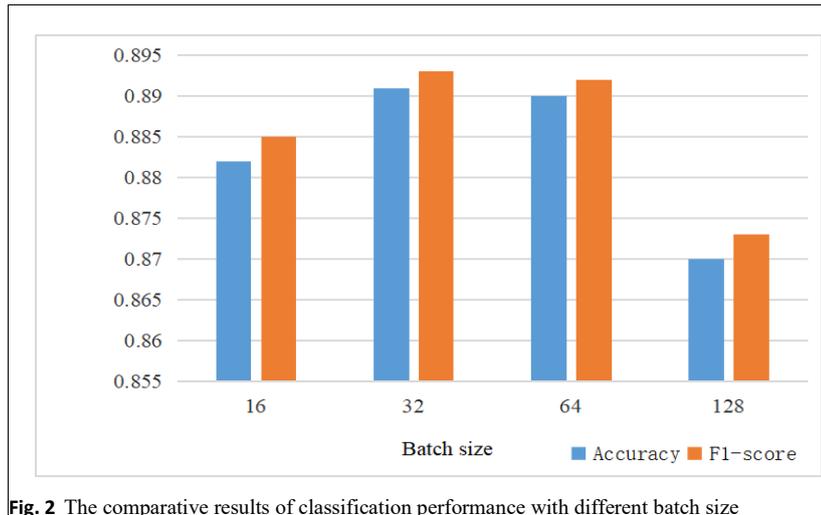
dilated rates are set to 1, 2, 4, 8, 2, 4, 8, respectively. The best parameters corresponding classification model are same as those in the former. The comparative results are showed in Table 1. From Table 1, we can see that the F1-score and accuracy can achieve the higher values of 91.46% and 91.48%, respectively, when the second selection scheme of dilated rates was employed. In comparison, it is obvious that the second selection scheme of dilated rates is superior to the first selection scheme. These results suggested that, with our selected gait data, when the dilated rates are selected as 1, 2, 4, 8, 2, 4, 8 in the different DC layers (i.e. seven DC layers), we can gain the more valuable intrinsic gait spatial features for best classification performance.

Table 1 The comparative results of classification performance with different dilated rates

Dilated rate setting	F1-score	Accuracy
1,2,4,8,16,2,4,8,16	90.42%	90.29%
1,2,4,8,2,4,8	91.64%	91.48%

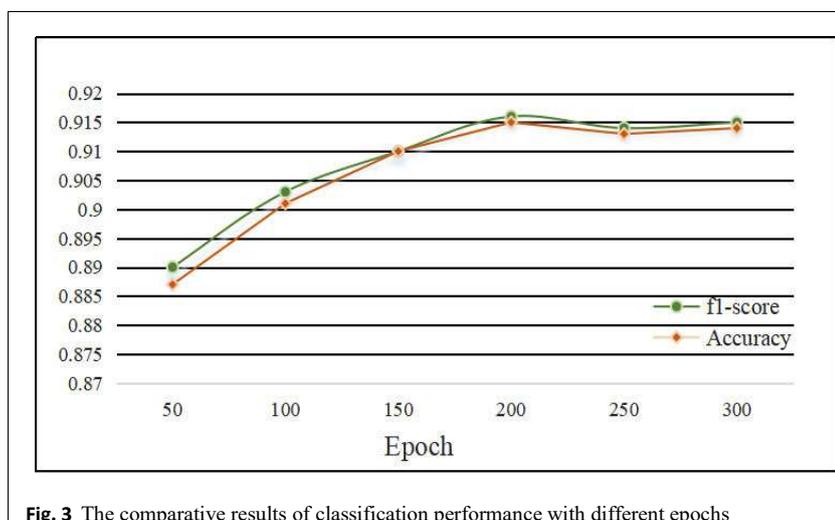
The proper batch size of our proposed model

Besides, we also evaluated the effects of the choice of the different batch size on the classification performance. In this experiment, in view of our selected sliding window with 3.5s length, we selected the batch size of 16, 32, 64 and 128, respectively, to implement our gait classification task. The epoch of classification model was set to 60, and accuracy and F1-score were also employed as evaluation criteria of assessing the generalization performance. The comparative results were given in Fig. 2. As shown in Fig. 2, it is obviously that both accuracy and F1-score can increase with the increase of batch size from 16 to 64, especially, they can obtain the almost same higher values respectively when the batch size increases from 32 to 64. In comparison, both accuracy and F1-score can reach the best values of 89.1% and 89.3% respectively when batch size was set to 32, suggesting that the selected batch size of 32 possibly contain the more intrinsic valuable spatial-temporal correlation information hidden in gait data for best gait classification performance. However, both accuracy and F1-score only obtain the minimum value when the batch size was fixed to 128, demonstrating that the batch size fixed to 128 can retain the more redundant information for the poor classification performance. These results illustrated that the properly selected batch size of gait data such as acceleration and gyroscope data can greatly benefit to the improvement of the generalization ability of our proposed model.



The selection of best epochs size of training model

Also, we evaluated the impact of the epochs in training process on the generalization ability of our proposed model. In this experiment, the fixed-length of sliding window and the batch size were set to 3.5s and 32, respectively. The other best parameters corresponding classification model are presented in Table 2. Accuracy and F1-score were used as evaluation criteria of classification model. The different epochs such as 50, 100, 150, 200, 250 and 300 were employed to train our proposed model, respectively. The comparative results were presented in Fig. 3. From Fig. 3, we can evidently see that both accuracy and F1-score increase with the increase of epochs when the epochs size increase from 50 to 200, and they slightly decrease and almost keep no change when the epochs size increase from 200 to 300. In comparison, both accuracy and F1-score can reach the almost same highest value of 91.5% while the epoch size is set to 200. These results illustrated that when the selected epochs size were more than 200 in training process, our proposed model could avoid the overfitting and underfitting limitation for gaining the best generalization performance.



The results of evaluation of our proposed model with high quality

In this section, we described the results of the evaluation of our proposed model with high quality. For comparison, some recent relevant deep learning models such as CNN, LSTM, CNN-LSTM, DCNN-LSTM, attention-based DCNN-GRU, attention-based DCNN-RNN and attention-based CNN-LSTM were employed. In this experiment, to objectively and accurately evaluate the classification performance, we adopted the same gait data, architecture of model, and the training and testing scheme. The fixed-length of sliding window, batch size and epoch size were set to 3.5s, 32, and 200, respectively. Both accuracy and F1-score were used as the evaluation criteria of the classification performance. With our selected multi-sensors data, the best architecture of our proposed model was presented in Fig. 4. As illustrated in Fig. 4, this architecture mainly includes seven dilated convolution layers, two long short-term memory layer, attention strategy and softmax layer. For all seven “DC” layers, we successively set dilation rate to 1, 2, 4, 8, 2, 4, 8, respectively. The numbers before and after “@” refer to dimension of a feature map and the number of feature maps in each DC layer. In first and fourth “DC” layers, there are the same 32 feature maps. In the reminder “DC” layers, there are the same 64 feature maps. In addition, in the first LSTM layer, the best 128 neurons can be obtained to effectively capture the more valuable temporary dependency change corresponding to the extracted local gait spatial features. In the second LSTM layer, 10 neurons are equal to ten selected gait patterns to be classified. The neurons of models are dropped out with a probability of 0.25. The comparative results were given in Table 2. As shown in Table 2, it is obviously that our proposed model can achieve the best classification performance (F1-score of 91.6%, accuracy of 91.5% and computational time of 428s) while the optimal architecture of our model consist of seven dilated convolution layers, two long short-term memory layer, attention strategy and softmax layer. In comparison, although the selected attention-based DCNN-GRU model can produce the almost same F1-score (91.4%) and accuracy (91.2%) as our proposed model, it needs to spend more computational time cost (612s) than our proposed model (computational time of 428s) [46]. In addition, we can see that despite the selected CNN, LSTM and attention-based DCNN-RNN models could spend the lowest computational time cost, their classification performances were poor. Especially, although the selected CNN model could cost the lowest computational time of 17s, its generalization performance is poorest (F1-score of 49.7%, accuracy of 60.7%). Meanwhile, from Table 2, we can also observe that some selected hybrid deep learning, such as attention-based CNN-LSTM, CNN-LSTM and DCNN-LSTM, can yield the better classification performance and low computational time cost. These results illustrated that our proposed technique with proper computational time cost is capable of exploiting the more valuable intrinsic nonlinear dynamic correlation features hidden in spatial-temporary gait change space when compared with some traditional or hybrid deep learning models, which significantly enhances the generalization ability of WBSNs’ gait classification.

Next, with our proposed model, we implemented the classification task of all selected ten gait patterns (i.e stand (St), sit (Si), lie down (LD), walk forward (WF), walk left-circle (WLC), walk right-circle (WRC), turn left (TL), turn right (TR), upstairs (UP) and downstairs(Down)), in order to

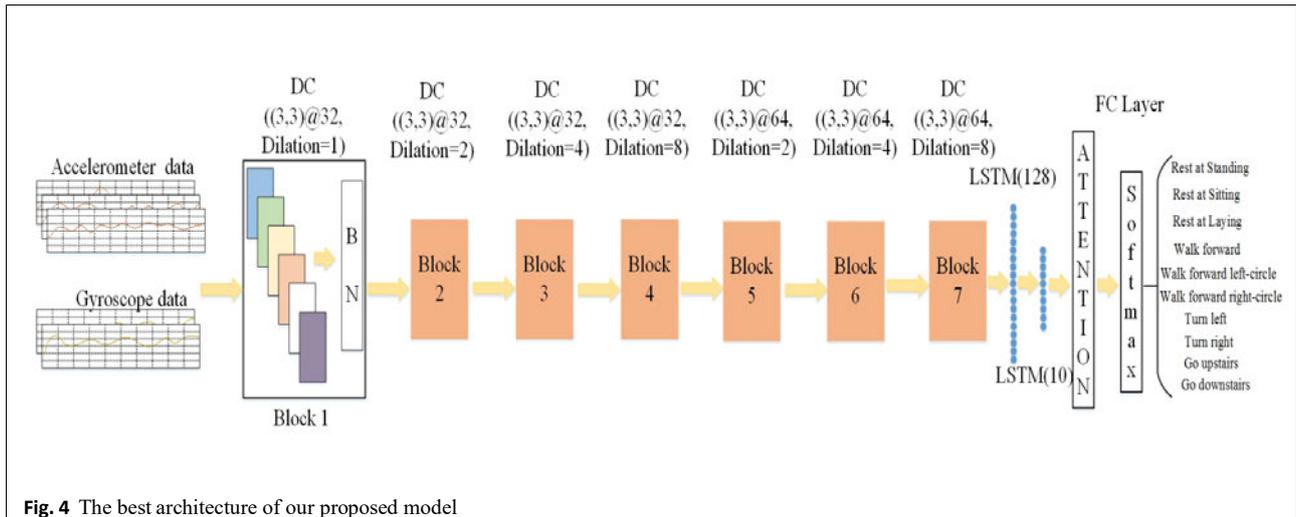


Table 2 The comparative results of best classification performance with the different deep learning models

Model	Model parameters				Performance		Time(s/Epoch)	
	The number of feature maps	DC rates	The number of neurons	Attention	The number of layers	F1-score/%		Accuracy/%
CNN [41]	64,64,64			No	3	49.7	60.7	17
LSTM [42]			32,32	No	2	73.2	72.9	56
CNN-LSTM [43]	512		100	No	2	80.7	80.2	267
Self-attention [44]				YES	5	75.1	75.2	876
DCNN-LSTM		1,2,4,8,2,4,8	128,10	No	9	89.5	89.4	411
Attention-DCNN-GRU		1,2,4,8,2,4,8	128,10	YES	10	91.4	91.2	612
Attention-DCNN-RNN		1,2,4,8,2,4,8	128,10	YES	10	61.7	64.2	263
Attention-CNN-LSTM	32,32,32,32,64,64,64		128,10	YES	10	90.8	90.7	441
Our proposed model		1,2,4,8,2,4,8	128,10	YES	10	91.6	91.5	428

further validate the excellent generalization ability of classification WBSNs' gait patterns. Here, in order to effectively assess the superior ability of accurate detection of each selected gait pattern, both recall and precision were used as the evaluation metrics of classification performance. The evaluation results with confusion matrix were presented in Table 3. In such confusion matrix, the i th row and j th column denote the total number of gait pattern with class i that is classified as class j . According to the results of confusion matrix in Table 3, the whole averaged precision and recall across all selected ten gait patterns can reach the maximum of 91.3% and 91.2%, respectively, suggesting that our proposed model can recognize all selected WBSNs-based gait patterns with higher accuracy. In comparison, from Table 3, we can find that the precision and recall from three "static" gait patterns such as stand, sit and lie down are higher while the precision and recall from two "active" gait patterns are lower. Especially, stand pattern can gain the best recall value of 95.6%, and both sit and lie down patterns can obtain the maximum precision values of 94.5%.

In contrast, downstairs pattern can obtain poor precision and recall values of 84.7% and 87.1%, respectively, which illustrated that some downstairs patterns could be misclassified as upstairs and

walk forward patterns. In addition, for walk left-circle pattern, its precision can reach the highest value of 93.1% whereas its recall value is only 89%. On the contrary, for walk forward pattern, its best recall reaches 90.3% while its precision is only 87.6%. The above results demonstrated that some walk left-circle and walk forward patterns can be misclassified as turn left and upstairs patterns. According to the results of confusion matrix in Table 3, the whole averaged precision and recall across all selected ten gait patterns can reach the maximum of 91.3% and 91.2%, respectively, suggesting that our proposed model can recognize all selected WBSNs' gait patterns with higher accuracy. In comparison, from Table 3, we can find that the precision and recall from three "static" gait patterns such as stand, sit and lie down are higher while the precision and recall from two "active" gait patterns are lower. Especially, stand pattern can gain the best recall value of 95.6%, and both sit and lie down patterns can obtain the maximum precision values of 94.5%. In contrast, downstairs pattern can obtain poor precision and recall values of 84.7% and 87.1%, respectively, illustrating that some downstairs patterns could be misclassified as upstairs and walk forward patterns. In addition, for walk left-circle pattern, its precision can reach the highest value of 93.1% whereas its recall value is only 89%. On the contrary, for walk forward pattern, its best recall reaches 90.3% while its precision is only 87.6%. The above results demonstrated that some walk left-circle and walk forward patterns can be misclassified as turn left and upstairs patterns.

Table 3 The evaluation results with confusion matrix for ten gait patterns classification

	St	Si	LD	WF	WLC	WRC	TL	TR	Up	Down	Total	Recall
St	2111	50	34	0	0	0	4	1	0	4	2204	95.6%
Si	99	1979	68	1	1	0	2	4	2	2	2158	91.7%
LD	28	54	2056	0	4	0	1	2	3	13	2161	95.1%
WF	1	0	9	1411	25	12	5	8	68	72	1611	87.6%
WLC	0	0	2	17	1924	19	50	24	13	17	2066	93.1%
WRC	4	0	2	7	48	1873	21	42	15	19	2031	92.2%
TL	1	1	1	9	96	14	1760	18	14	25	1939	90.8%
TR	2	3	0	17	26	52	49	1821	32	24	2026	89.9%
Up	0	4	2	61	15	10	4	16	1697	93	1902	89.2%
Down	7	3	1	39	23	9	7	7	125	1490	1711	87.1%
Total	2253	2094	2175	1562	2162	1989	1903	1943	1969	1959		91.2%
Precision	93.7%	94.5%	94.5%	90.3%	89%	94.2%	92.5%	93.7%	86.2%	84.7%	91.3%	

The evaluation results of the effect of intrinsic nonlinear gait correlation captured on gait change

As we know, a certain gait pattern is characterized by the implicit nonlinear correlation in the different anatomical regions of body during human movement, and discovering such implicit degree of a certain gait pattern significantly contributes to finding the valuable on-body anatomical regions associated with a certain gait pattern change, which could provide the insight into the variability of gait pattern. In this experiment, the different combination of sensors at different positions on body is considered as on-body different anatomical regions associated with a certain gait pattern, in order to accurately evaluate the effect of intrinsic nonlinear gait correlation exploited from different anatomical regions on gait change. The detailed evaluation scheme was described as follows. That is, according to

Table 4 The comparative results of best F1-score from the different combination of different sensors based on a certain gait

Class	S1	S2	S3	S4	S5	All5	S123	S124	S125	S134	S135	S145	S234	S235	S245	S345
St	0.92	0.93	0.91	0.87	0.94	0.95	0.96	0.94	0.94	0.93	0.96	0.95	0.95	0.95	0.94	0.94
Si	0.90	0.91	0.87	0.88	0.91	0.93	0.93	0.91	0.93	0.92	0.94	0.93	0.93	0.92	0.92	0.92
LD	0.93	0.89	0.89	0.91	0.92	0.95	0.93	0.93	0.94	0.92	0.95	0.95	0.92	0.93	0.92	0.93
WF	0.81	0.74	0.77	0.77	0.82	0.89	0.88	0.86	0.88	0.87	0.85	0.85	0.85	0.87	0.87	0.87
WLC	0.84	0.85	0.84	0.80	0.86	0.91	0.91	0.88	0.91	0.89	0.90	0.89	0.90	0.91	0.90	0.90
WRC	0.85	0.84	0.80	0.80	0.82	0.93	0.91	0.90	0.91	0.88	0.91	0.89	0.90	0.90	0.91	0.91
TL	0.83	0.81	0.82	0.79	0.85	0.92	0.89	0.87	0.89	0.88	0.90	0.89	0.88	0.90	0.89	0.91
TR	0.85	0.82	0.80	0.82	0.84	0.92	0.89	0.90	0.89	0.88	0.90	0.89	0.90	0.90	0.89	0.91
Up	0.79	0.76	0.79	0.73	0.83	0.88	0.88	0.82	0.85	0.85	0.86	0.84	0.86	0.87	0.85	0.87
Down	0.77	0.77	0.77	0.72	0.78	0.86	0.86	0.81	0.83	0.82	0.84	0.81	0.84	0.84	0.82	0.83

WARD database, we randomly selected one sensor, three sensors and all five sensors as different combination of sensors that is considered as the different anatomical regions. As shown in Fig. 7 in experiment section, on-body sensor1, sensor2, sensor3, sensor4 and sensor5 were denoted as S1, S2, S3, S4 and S5, respectively. They are located on left and right forearm, the waist, the left and right ankles of body, respectively. The combination of three sensors on body was represented as a subset including three sensors selected arbitrarily. For example, S123 was denoted as a combination of on-body sensor1, sensor2 and sensor3. Our proposed model with the best parameters was adopted for classification of WBSNs gait pattern with the combination of different on-body sensors. Here, F1-score was selected as the evaluation criteria of the generalization performance. The comparative results of best F1-score were given in Table 4. It was obviously observed from Table 4 that, for each gait classification, F1-scores values are best when the combination of all five sensors or three sensors were used. This suggested that the combination of all five sensors or three sensors could contain the more valuable implicit nonlinear gait correlation characteristics information than one sensor, which significantly enhanced the generalization ability of gait classification. In terms of a certain gait pattern, most of F1-scores from gait patterns of stand, sit and lie down can almost gain more than 90% while those from the remainder seven gait patterns such as walk forward, walk left-circle, walk right-circle, turn left, turn right, go upstairs, go downstairs were less than 90%. In comparison, F1-score from gait patterns of stand can achieve the on a whole best values while F1-score from gait pattern of walk forward obtain all poorest values. These results demonstrated that, with the combination of different sensors, our proposed technique was able to capture the more valuable implicit nonlinear gait correlation feature information from “static” gait patterns such as stand, sit and lie down when compared to the “active” gait patterns such as walk forward, walk left-circle, walk right-circle, turn left, turn right, go upstairs, go downstairs. Especially, from Table 4, we can find that F1-score can reach higher value using a certain combination of three sensors than all five sensors. For example, in stand gait detection task, F1-score values can achieve the maximum of 96% using S123 or S135 whereas it can gain the highest value of 95% using all five sensors. This also suggested that a certain anatomical regions, such as including waist, left and right wrists, or left wrist, left and right ankle, possibly contain the more valuable implicit nonlinear correlation information associated with gait pattern of stand

compared to the anatomical regions consisting of waist, left and right wrists, left and right ankles. In conclusion, these results demonstrated that our proposed method could exploit the intrinsic valuable nonlinear gait correlation features associated with gait variability in a certain anatomical regions, which could help to not only produce the best generalization performance but also take insight on gait change from a certain anatomical regions for potential gait variability interpretation.

Discussion

The experiment results of the present study showed that our proposed attention-based hybrid deep learning model of DCNN and LSTM could feasibly exploit the more intrinsic valuable nonlinear dynamic gait correlation characteristics from the on-body multi-sensors gait data such as acceleration and gyroscope data during walking, which greatly enhanced the generalization ability of WBSNs' gait classification. Currently, finding the best deep learning model for WBSNs' gait recognition with high quality has been a challenging endeavor in gait analysis. A key issue to be addressed in recent related work is how to exactly exploit the intrinsic most representative characteristics highly associated with gait change from on-body multi-sensors gait data. Theoretically, gait variability is highly depended on gait compensatory mechanism that is characterized by the whole coupling of anatomical structure of human locomotion. Such mechanism is to a great extent reflected by the correlation between joints and muscle in a certain anatomical regions during walking. The aim of this study is to find an advanced deep learning technique that can capture the intrinsic valuable nonlinear gait correlation features from a certain anatomical regions, which significantly contribute to improving WBSNs' gait classification performance and taking insight on gait variability.

In this study, we aimed to take advantage of attention mechanism to capture the implicit valuable nonlinear dynamic correlation information from the spatial-temporal correlation gait features space obtained by hybrid model of DCNN and LSTM. The evaluation of the effectiveness and feasibility of our proposed model was performed based on an open WARD database. The best parameters of our proposed model were given in Fig. 1, Table 1, Fig. 2 and Fig. 3, respectively. In order to gain data segmentation containing the more useful gait characteristics information, we selected the proper fixed-length of sliding window for the highest generalization performance. As shown in Fig. 1, the optimal fixed-length of 3.5s was selected to reach the best classification performance, when compared to the fixed-length of approximately 5s selected usually in previous study [45]. This is because data segmentation with fixed-length of 3.5s can contain the more valuable gait features information associated with gait change, whereas data segmentation with fixed-length of 5s can contain the more redundant information that reduce classification performance. In addition, in this study, it is very important to select the optimal number of DC and LSTM layers for best classification performance. As illustrated in Table 1 and Table 2, seven layers of DC and two layers of LSTM were chose to produce the highest generalization performance. Especially, the dilated rates in DCNN are experimentally selected as 1, 2, 4, 8, 2, 4, 8 to exploit the more valuable representative spatial gait features when compared to theoretically selected dilated rates of 1, 2, 4, 8, 16, 2, 4, 8, 16. The possible reason is that the dynamic nonlinearity and non-stationary of multi-sensors acceleration and gyroscope data selected are possibly sensitive to receptive field size

expanded by the proper selection of dilated rates in DCNN, which could benefit to dynamically capturing the intrinsic spatial gait features at different scales. In this experiment, we also found that all optimal parameters must be carefully selected by trial and error method for best generalization performance because the best value of each parameter could change with the different values of other parameters. Currently, there have been no general agreements on standards for the selection of the optimal parameters in deep learning algorithms. The similar reports have been found in [46].

In the present study, we focused on the investigation of our proposed model for WBSNs' gait classification with high quality. As shown in Table 2, our proposed model was best when compared with some deep learning models used in recent studies. In the experiment, we can find that most of hybrid deep learning models can produce superior generalization ability than individual deep learning model such as CNN and LSTM. The main reason is that hybrid deep learning model can capture the valuable spatial-temporary correlation information between both intra-sensors and inter-sensors on body, whereas CNN or LSTM only learn the spatial or temporary features in gait data. Similar results have been reported in [21,47]. Especially, our proposed technique can significantly exploit the more intrinsic nonlinear dynamic correlation features from spatial-temporary space associated with gait change while the lower computational time cost is spent. This is because attention-based DCNN-LSTM can concentrate on the most valuable spatial-temporal correlation features between both intra-sensors and inter-sensors, and thus it could suppress the redundant and possibly confusing information of multi-sensor gait data to reduce the overall training time [21]. It is hopeful that our proposed technique has a great potential application in real-time monitoring gait change. No similar results have been reported in recent studies.

In addition, we further evaluated the discriminative ability of our proposed model with high quality by classifying all selected ten gait patterns. Unlike some previous studies, this study randomly selected some similar gait patterns including "static" gait patterns such as stand, sit, and lie down, and "active" gait patterns such as walk forward, walk left-circle, walk right-circle, turn left, turn right, upstairs and downstairs, in order to accurately detect the smaller difference among WBSNs' gait patterns. As illustrated in Table 3, our proposed model can feasibly detect each gait pattern with a higher recall and precision. This further suggests that our proposed technique is capable of discovering the valuable intrinsic gait variability information between selected "static" or "active" gait patterns. The main reason is that our proposed method can take advantage of attention mechanism to exploit the more local distinctive nonlinear information associated with the dynamic spatial-temporal correlations of gait [28, 22-24], which significantly contribute to recognizing gait change. Similar studies have been reported in [16].

Besides, in this study, we also investigate the effect of combination of different on-body sensors on detecting a certain gait by our proposed technique. We aimed to find the impact of the intrinsic interaction between joint and muscle in a certain anatomical regions on gait change based on gait compensatory mechanism. As shown in Table 4, our proposed technique can accurately detected

each gait pattern when the combinations of three or all five on-body sensors were adopted. Especially, for a certain gait pattern such as stand, the superior identification ability of some combination of three sensors such as S123 is much the same as that of all five sensors. This further illustrated that our proposed model had excellent learning ability to capture intrinsic local interesting information, such as specific details of gait deviations highly associated with gait compensatory mechanism, which resides in a certain anatomical regions. Recent similar studies have been reported in [17]. In comparison, our proposed model is superior. The main reason is that the local energy-based shape histogram method used in recent study hardly discover the more valuable intrinsic gait features from multi-sensors data, however, our proposed technique has the superior learning capability to exploit the more local implicit nonlinear characteristics information about gait variability from a certain anatomical regions. This greatly benefits to gaining deeper insight into the pathologies and physiological change of gait compensatory mechanism based on a certain anatomical regions.

Conclusion

This study proposed a novel attention-based hybrid deep learning model of DCNN and LSTM for WBSNs' gait pattern recognition with high quality. Our proposed model with superior learning ability can feasibly exploit the valuable intrinsic nonlinear gait correlation characteristics associated with gait variability from on-body multi-sensors gait data such as acceleration and gyroscope data, which significantly improves the generalization performance of WBSNs' gait classification. With our proposed technique, a certain anatomical regions highly associated with gait pattern change can be obtained by the combination of different sensors, which greatly contributes to take insight on the intrinsic correlation of gait deviations residing in gait compensatory mechanism for potential clinical interpretation. Further work will evaluate the effectiveness and feasibility of our proposed model using the different kinds of multi-sensors gait data from patients subjects such as PD patients. It is hopeful that our proposed method would become a effective tool for monitoring gait change in clinical environment in the future.

Method

In this study, in view of multi-sensors dynamic time-series gait data with nonlinearity, nonstationary and stochasticity, we proposed an advanced attention-based hybrid deep learning model of DCNN and LSTM for WBSNs' gait classification. The objective of our study is to take advantage of deep fusion of the excellent learning ability of attention mechanism, DCNN and LSTM for WBSNs' gait pattern recognition with high quality in clinical application. The basic idea is that, with WBSNs' gait pattern defined by the combination of different on-body multi-sensors, we firstly took advantage of DCNN technique to exploit intrinsic spatial gait feature hidden in on-body multi-sensors gait data. Secondly, LSTM was utilized to capture the intrinsic temporal dependency corresponding to the extracted intrinsic spatial gait features. Namely, we constructed a spatial-temporal correlation space embedded in multi-sensors gait data by hybrid deep learning model of DCNN and LSTM. Finally, the attention mechanism was introduced to discover the valuable local nonlinear dynamic correlation information highly associated with gait variability

from the constructed spatial-temporal gait correlation space. This significantly improves the generalization performance can find a certain anatomical regions for potential clinical interpretation.

Based on the above basic idea, the architecture of our developed model mainly consist of input layer, H blocks based on dilated convolution (DC) layers with different dilation rates, K layers of long short-term memory with different number of neurons $K_1 \dots K_m$, attention strategy and softmax layer, as shown in Fig. 5 The input layer includes multi-sensors data such as acceleration data and gyroscope data. For DC networks capacity, each block includes a DC layer and a batch normalization(BN) layer. In the process of gait feature learning, the dilated convolution layers are firstly utilized to extract the intrinsic local short-term spatial gait features embedded in multi-sensors gait data. And then, long short-term memory layers are implemented to model temporal dependency corresponding to the local short-term spatial gait features extracted, in order to capture the implicit spatial-temporal correlation gait characteristics hidden in gait data. Next, attention strategy is employed to discover the local intrinsic valuable nonlinear dynamic gait feature embedded in spatial-temporal correlation feature space. Therefore, we can gain the more valuable discriminative information among WBSNs gait patterns, which greatly contributes to softmax layer that accurately recognize gait pattern with maximum probability. The detailed description of our proposed model is described as follows.

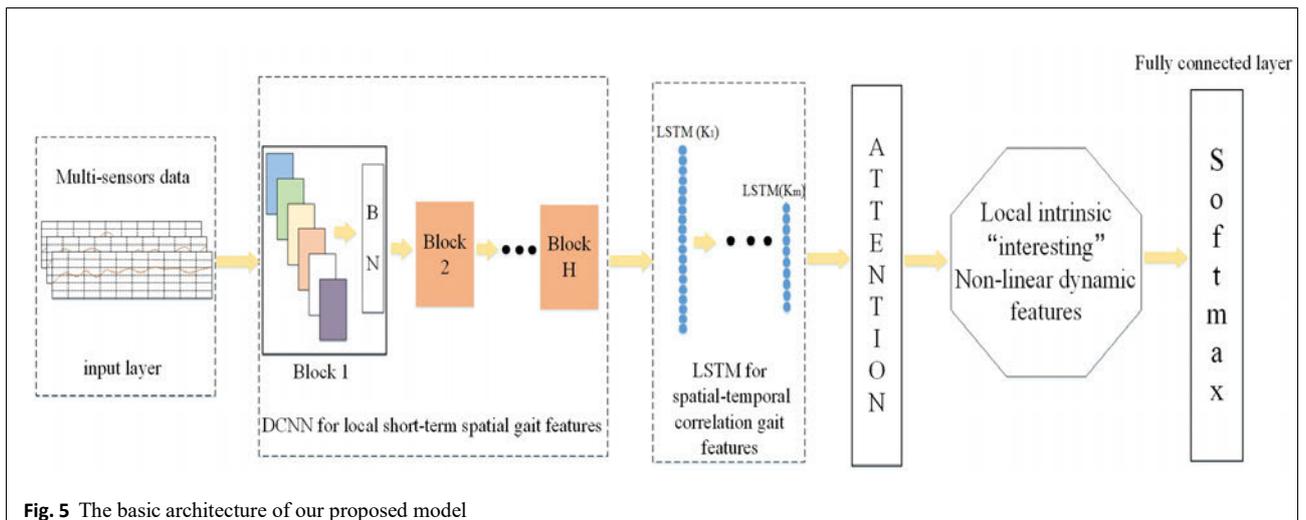


Fig. 5 The basic architecture of our proposed model

Definition of WBSNs' gait patterns

Firstly, the WBSNs' gait pattern is defined by the original multi-sensors time-series gait data in input layer. Given that there are a total of L sensors on body, the time-series gait data of each sensor j at time t can be defined as a vector $a^j \in R^g$ where g denotes the total of the data features from all sensors such as accelerometer with triaxial and gyroscope with biaxial.

According to the defined $a^i \in \mathbb{R}^g$, the gait data of sensor j during the length of time h can be defined as a vector s^j :

$$s^j = [a^j(1), a^j(2), \dots, a^j(h)] \in \mathbb{R}^{gh} \quad (1)$$

Next, according to equation (1), the gait data containing all L sensors during the length of time h can be defined as a vector v

$$v = [s^1, s^2, \dots, s^L]^T \in \mathbb{R}^{gLh} \quad (2)$$

And then, according to equation (2), the time-series gait data including all L sensors with gait pattern class i during the length of time h can be defined as a vector V_i :

$$V_i = [v_{i,1}, v_{i,2}, \dots, v_{i,n_i}] \in \mathbb{R}^{m \times n_i} \quad (3)$$

where n_i denotes all samples of the gait pattern class i , and $m = gLh$.

So, according to equation (3), we can define the WBSNs gait pattern using the combination of the different on-body sensors. In addition, the time-series gait data consisting of all L sensors with all gait pattern classes K during the length of time h can be defined as a vector D :

$$D = [V_1, V_2, \dots, V_k] = [v_{1,1}, v_{1,2}, \dots, v_{k,n_k}] \in \mathbb{R}^{m \times n} \quad (4)$$

where a total of WBSNs' gait samples $n = n_1 + n_2 + \dots + n_k$. According to equation (4), we can obtain the sample training set of WBSNs gait patterns to be classified.

DCNN for local short-term spatial gait features

Next, based on the defined WBSNs' gait samples set D , DCNN technique is employed to exploit the implicit local short-term spatial gait features hidden in on-body multi-sensors time-series gait data. As we know, the advantages of DCNN is to adopt dilation convolution operator to change the receptive field size while the size of field map (i.e. convolutional kernel) keeps no change. That is, in each block, assuming that λ denotes a dilation factor, the λ -dilated convolution operation \otimes_λ for local intrinsic gait feature extraction can be defined as

$$(F \otimes_\lambda k)(x, y) = \sum_{m=-r}^r \sum_{n=-r}^r F(x - \lambda m)(y - \lambda n)k(m, n) \quad (5)$$

where the discrete function F is generalized as $F_0, F_1, \dots, F_{n-1} : \mathbb{Z}^2 \rightarrow \mathbb{R}$, and the discrete filter k responding to F is also generalized as $k_0, k_1, \dots, k_{n-1} : \Omega_1 \rightarrow \mathbb{R}$. Here, because the dilated filter varies with exponentially increasing dilation factors such as $2^0, 2^1, \dots, 2^j$, the receptive field size in each F_j is equal to $(2^{j+1} - 1) \times (2^{j+1} - 1)$. Therefore, it is important to choose the proper dilation factors to effectively expand receptive field size that benefits to dynamically capturing the intrinsic spatial gait features at different scales. With our proposed DCNN model, we can gain the more intrinsic local short-term spatial gait features from on-body multi-sensors gait data

LSTM for spatial-temporal correlation gait features

And then, we take advantage of LSTM technique to discover the intrinsic temporal dependencies corresponding to the local spatial gait features extracted by DCNN, in order to exploit the spatial-temporal correlation gait characteristics hidden in WBSNs gait. As we know, the basic unit of recurrent architecture of LSTM consists of one state memory cell and three gate functions including input gate, forget gate and output gate. Such architecture cell can perform the recurrence computation that significantly relies on the corresponding computation in previous time step. That is, when the local spatial gait features F^t at time t are fed into the LSTM cell with the $t-1$ timestamp memory C_{t-1} and hidden state h_{t-1} , the input gate i_t , forget gate f_t and output gate o_t will be calculated by the functions as follow.

$$f_t = \sigma(W_f \cdot [F^t, h_{t-1}] + b_f) = \sigma(W_f F^t + W_f h_{t-1} + b_f) \quad (6)$$

Where σ, W_f, b_f are the non-linear function, the weighted matrix and bias vector to be learned during training, t represents the current time state, h_{t-1} denotes output value of the previous time $t-1$. Next, the updated information on temporal correlation at time t is determined by following input gate function i_t and candidate memory cell \tilde{C}_t :

$$i_t = \sigma(W_i \cdot [F^t, h_{t-1}] + b_i) \quad (7)$$

$$\tilde{C}_t = \tan(W_c \cdot [F^t, h_{t-1}] + b_c) \quad (8)$$

where W_i and W_c denote the weighted matrix. b_i and b_c denote bias vectors to be learned during training. And then, according to equation (7) and (8), the updated memory state cell for temporal correlation corresponding to local spatial feature of gait at time t is defined as C_t :

$$C_t = i_t \cdot \tilde{C}_t + f_t \cdot C_{t-1} \quad (9)$$

Also, the output gate function o_t at time t is defined as

$$O_t = \sigma(W_o \cdot [x_t, h_{t-1}] + b_o) \quad (10)$$

where W_o and b_o denote the weighted matrix and bias vector to be learned during training, respectively.

So, according to equation (10), the output of current memory state containing temporal dependencies corresponding to local spatial gait feature at time t is defined as h_t ,

$$h_t = o_t * \tan(C_t) \quad (11)$$

Therefore, according to equation (11), we can define the spatial-temporal correlation gait characteristics as a vector

$$s = [h^1(1), \dots, h^n(T)] \quad (12)$$

where T is total number of time steps and n denotes the number of dimensions of each time step. The detailed procedures for solution of recurrence computation in LSTM are presented in literature [38]. In addition, it is key to select the proper number of LSTM layers to achieve the best generalization performance.

Attention-based technique for local intrinsic nonlinear dynamic correlation gait features

Next, attention-based technique is employed to exploit the intrinsic specific nonlinear dynamic gait characteristics embedded from the spatial-temporal correlation space obtained by hybrid deep learning model of DCNN and LSTM. Here, assuming that $h^i \in R^n$ denotes a intrinsic dynamic gait feature at the t -th time step in the spatial-temporal correlation vector data $s = [h^1(1) \dots, h^n(T)] \in R^n$ with length n , we define the attention score a_t at time step t as a probability value that illustrates the significant contribution of the intrinsic dynamic gait feature h at time step t to the whole spatial-temporal correlation vector data $s = [h^1(1) \dots, h^n(T)] \in R^n$. A scheme of solution score α_t is presented [39], as shown in Fig. 6. Firstly, a specific nonlinear dynamic gait feature $s'(t)$ can be obtained by the learnable nonlinear function $g(\cdot)$ and the hidden state sequence $h(t)$ at time step t ,

$$s'(t) = g(h(t)) \tag{13}$$

And then, the attention score α_t at time t can be is defined as a probability vector by

$$\alpha(t) = \frac{\exp(s'(t))}{\sum_{k=1}^T \exp(s'(k))} \tag{14}$$

According to equation (14), the attention score α_t at time t obtained can be considered as a weighted value that reflects the significant contribution to intrinsic specific nonlinear dynamic gait features in the whole spatial-temporal correlation. So, we can obtain the intrinsic valuable specific nonlinear dynamic correlation gait feature s^* hidden in spatial-temporal correlation space by computing the following equation:

$$s^* = \sum_{t=1}^T \alpha(t)h(t) \tag{15}$$

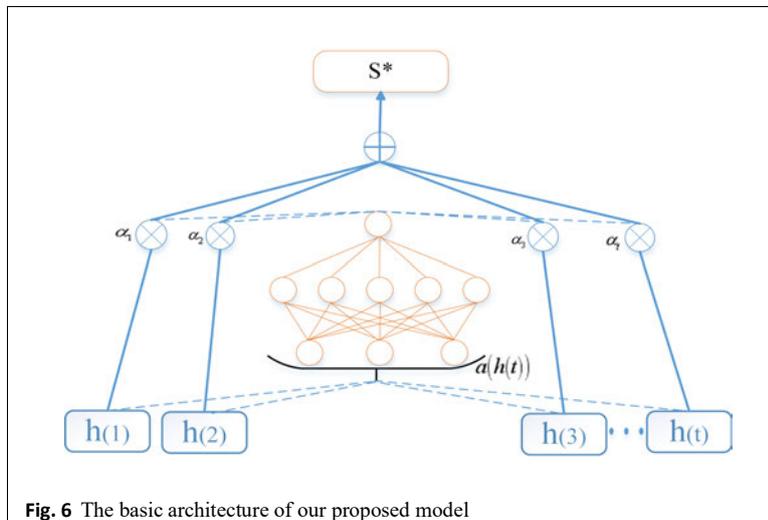


Fig. 6 The basic architecture of our proposed model

Accurate recognition of WBSNs' gait patterns

In softmax layer, according to the intrinsic valuable specific nonlinear dynamic correlation gait feature s^* obtained, we can define a softmax function as the probability distribution

$$\hat{p}(y | D) = \text{soft max}(W^{(D)}s^* + b^{(D)}) \quad (16)$$

Thus, for the all WBSNs gait pattern sample dataset D including all K classes, each pattern \hat{y} to be predicted can be accurately classified by the maximum probability distribution

$$\hat{y} = \arg \max \hat{p}(y | D) \quad (17)$$

where $W^{(D)}$ and $b^{(D)}$ are the weighted matrix and bias vector, respectively.

In order to achieve the best learning performance, we utilized the common cross-entropy error function as cost function.

$$L(\theta) = -\frac{1}{k} \sum_{i=1}^k y_i \log(p_i) \quad (18)$$

where k represents the total number of target classes to be predicted, and $y \in R^k$ denotes a one-hot vector that represents the label of the d -th data of the m -th sample. The $p_i \in R^k$ denotes the probability distribution for accurately predicting gait class i , and θ is the parameters of the learning network.

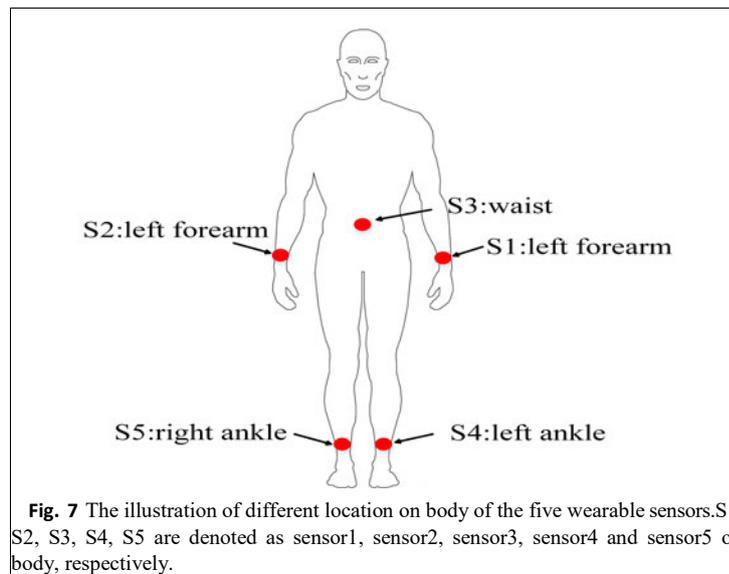
Experiment

The selection of on-body multi-sensors gait data

In this study, an open wearable action recognition database (WARD) including on-body multi-sensors gait data from University of California was selected, in order to evaluate the effectiveness of our proposed technique. In the selected WARD database, there are a total of 20 participants including 7 females and 13 male subjects, and the age of all subjects range from 19 to 75 years. Each subject had no known injuries or abnormalities associated with human walking pattern. The on-body multi-sensors gait data were acquired by the five wireless motion sensors that are, respectively, located on the right and left wrists, the waist, the right and left ankles using elastic belt, as shown in Fig. 7. Each wireless motion sensor was equipped with a triaxial accelerometer within the range of $\pm 2g$ and a biaxial gyroscope within the range of $\pm 500dps$ while each axis is reported as a 12bit value to the on-body sensor. The sampling frequency is set to 30 Hz for data recording while each subject wearing all five sensors was asked to walk at a natural self-selected walking speed. Each subject performs the data recording of a total of 13 different walking patterns, such as stand, sit, lie down, walk forward, walk left-circle, walk right-circle, turn left, turn right, go upstairs, go downstairs, jog, jump and push wheelchair, were asked to perform. The detailed scheme of data collection can be found in reference [26].

Data denoise and normalization

In order to obtain the gait data with high quality, the collected data denoise and normalization was firstly performed. Here, a combination denoise scheme of a fifth order median filter with a finite impulse response (FIR) filter was proposed. Its basic idea is that a fifth order median was firstly



used to reduce noise for smoothing the data, and then, a FIR filter was employed to remove the redundancy information for obtaining the data containing the more valuable information about gait. In the FIR filter, the size of filter window, cut-off frequency and sampling frequency were set to 64, 0.3Hz and 30Hz, respectively. And then, the denoise gait data was normalized to interval [-1,1] by using the linear Min-Max Normalization algorithm. This significantly contributes to avoiding the difference of the dimension and order of magnitude between data while retaining the original data structure for further data processing.

Data segmentation

In this study, data segmentation was implemented by the common sliding window method [36], in order to greatly contribute to performing further quantitative analysis of dynamic multi-sensors time-series gait data. Here, the fixed-length sliding window with two thirds overlap next to window was designed to perform data segmentation, in order to maintain the intrinsic best correlation in dynamic time-series gait data. In view of the significant impact of the length of sliding window (i.e. sliding window size) on the superior ability of both the valuable gait feature exploitation and gait classification, the selection of proper sliding window size was performed based on the best generalization performance of our proposed model.

Training and testing scheme of our proposed model

In order to effectively evaluate the best prediction ability of our proposed model, we build a training and testing sample data set including all multi-sensors gait data of 20 subjects in WARD database, each subject randomly selecting 10 gait patterns from all 13 gait patterns, such as stand, sit, lie down, walk forward, walk left-circle, walk right-circle, turn left, turn right, go upstairs, go downstairs. In this study, 105 data segmentations of each axis in accelerometer and gyroscope were arbitrarily extracted from the gait data of each participant. In view of all selected data segmentations, all samples data size can be obtained by

$$Samples = \frac{N_{DL} - N_{OV}}{L_{SL} - N_{OV}} \quad (19)$$

where N_{DL} represents the number of data length (i.e. 20 subjects \times each activity repeats 5 times \times the total number of a activity sequence number of more than 10 seconds = 2311048), N_{OV} denotes the number of the overlap size and L_{SL} is the segmentation length [40]. According to equation (19), the total of samples data size were 66027.

And then, the detailed scheme of training and testing was given as follows. The 70 percent of all samples data were randomly selected as the training data set while the rest 30 percent of all sample data were used as the testing dataset. It is noted that the data for testing can not be included in the training set, in order to objectively and accurately evaluate the generalization ability. In the training process, the backpropagation learning algorithm with Adam-based iteration rule for gradient descent was selected to search for the optimal parameters of our classification model while achieving the minimization of loss function. In this experiment, Adam optimizer with initial learning rate was set to 0.001, and the default parameters β_1 and β_2 were set to 0.9 and 0.999, respectively. All parameters of classification models were randomly initialized by glorot uniform. In the testing process, the final average results were obtained for all subjects while all optimal parameters of classification model were determined by the best generalization performance.

In addition, all deep learning algorithms programs were developed in the Tensorflow using python language, and they were implemented on computer with Intel(R) Core(TM) i5-6500 3.2GHz CPU, 8.00 GB RAM and Windows 7 operating system.

Evaluation criteria for our proposed model

In order to objectively evaluate the generalization performance of our proposed classifier, we adopted some common assessment criteria for gait classification task such as accuracy, recall, precision and F1-score. The detailed description of these criteria adopted was given as follows.

Accuracy is used to measure the generalization ability of our proposed model to recognize gait patterns in testing process, and is defined as [34].

$$Accuracy = \frac{TP}{NP} \times 100\% \quad (20)$$

Where TP and NP denote the total number of gait patterns truly detected and the whole number of gait patterns to be identified, respectively.

Recall is considered as a measure of the classification ability to specify gait with class j in the confusion gait between class i and class j , and is defined as [34].

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (21)$$

Precision is employed to objectively evaluate the ability of classification model to exactly discriminate gait patterns with class i in the confusion gait between class i and class j , and is defined as [34].

$$\text{Precision} = \frac{TP}{TP + FP} \times 100\% \quad (22)$$

where TP and FP denote the number of gait patterns accurately identified as class i and class j , respectively.

F1-score is utilized to objectively assess the ability to exactly predict the class of each gait pattern in testing process, and is defined as [34].

$$F1\text{-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \times 100\% \quad (23)$$

Abbreviations

WBSNs: wireless body sensor networks; CNN: Convolutional neural networks; LSTM: Long short-term memory; RNN: recurrent neural network; DCNN: dilated convolutional neural network;

Acknowledgements

We are thankful for the National Natural Science Foundation of China (Grant Number 82072043) and Fujian Provincial Natural Science Foundation of China (Grant Number 2020J01163). In addition, the authors wish to acknowledge the open Wearable Action Recognition Database (WARD) database from University of California, Berkeley.

Authors' contributions

WJN conceived the original idea of the study, and a corresponding author. He participated in designing model and drafting the manuscript. TQL participated in designing the algorithm and developing programme for analyzing the data, and WXY participated in writing the part of manuscript. All authors read and approved the final manuscript.

Funding

This study was funded by National Natural Science Foundation of China (Grant Number 82072043) and Fujian Provincial Natural Science Foundation of China (Grant Number 2020J01163).

Availability of data and materials

The WARD database is available for download at: <http://www.eecs.berkeley.edu/~yang/software/WAR/>.

Declarations

Ethics approval and consent to participate

This study was in accordance with the ethical standards of the institutional and/or national research committee.

Consent for publication

Not applicable.

Informed consent

For this type of study, formal consent was not required.

Competing interests

The authors declare that they have no conflicts of interest.

Author detail

¹ College of Computer and Cyber Security, Fujian Normal University, Fuzhou 350117, China.

² School of Management, University of Bristol, Bristol, BS8 1QU, UK

References

1. E.M. Tapia, S.S. Intille, and K. Larson, "Activity Recognition in the Home Setting Using Simple and Ubiquitous Sensors," Proc. Int'l Conf. Pervasive, pp. 158-175, 2004.
2. M. Philipose, K.P. Fishkin, M. Perkowitz, D.J. Patterson, D. Fox, H. Kautz, and D. Hahnel, "Inferring Activities from Interactions with Objects," IEEE Pervasive Computing, vol. 3, no. 4, pp. 50-57, Oct.-Dec. 2004.
3. Movassaghi S, Albolhasan M, Lipman J, Smith D, Jamalipour A. Wireless body area networks: A Survey[J]. IEEE Communication and Tutorials, 2014, 16(3): 1658-1686.
4. Rawat P, Singh KD, Chaouchi H, Bonnini JM. Wireless sensor networks: A survey on recent developments and potential synergies[J]. Journal of Supercomputing, 2014, 68: 1-48.
5. Nigussie E, Xu T, Potkonjak M. Securing wireless body sensor networks using bijective function-based hardware primitive [C]. 2015 International Conference on Intelligent Sensors, Sensor Networks and Information Processing Singapore, pp. 1-6, Apr, 2015.
6. Gu, F.; Khoshelham, K.; Valaee, S.; Shang, J.; Zhang, R. Locomotion Activity Recognition Using Stacked Denoising Autoencoders. IEEE Internet Things J. 2018, 5, 2085–2093.
7. Hsu, Y.L.; Yang, S.C.; Chang, H.C.; Lai, H.C. Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network. IEEE Access 2018, 6, 31715–31728.
8. Tahafchi P, Judy J W. Freezing-of-Gait Detection Using Wearable Sensor Technology and Possibilistic K-Nearest-Neighbor Algorithm[C]. international conference of the IEEE engineering in medicine and biology society, 2019: 4246-4249.

9. Paraschiakos, S.; Cachucho, R.; Moed, M.; van Heemst, D.; Mooijaart, S.; Slagboom, E.P.; Knobbe, A.; Beekman, M. Activity recognition using wearable sensors for tracking the elderly. *User Model. User Adapt. Interact.* 2020, 30, 567–605.
10. Anguita, D.; Ghio, A.; Oneto, L.; Parra, X.; Reyes-Ortiz, J.L. A public domain dataset for human activity recognition using smartphones. In *Proceedings of the 21th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning, ESANN 2013, Bruges, Belgium 24–26 April 2013*; pp. 437–442.
11. Zhang, M.; Sawchuk, A.A. USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing, Pittsburgh, PA, USA, 5–8 September 2012*; pp. 1036–1043.
12. Avilés-Cruz, C.; Ferreyra-Ramírez, A.; Zúñiga-López, A. Coarse-Fine Convolutional Deep-Learning Strategy for Human Activity Recognition. *Sensors* 2019, 19, 1556.
13. R. Xi, M. Li, M. Hou, M. Fu, H. Qu, D. Liu, and C. R. Haruna, "Deep dilation on multimodality time series for human activity recognition," *IEEE Access*, vol. 6, pp. 53381–53396, 2018.
14. Murad, A.; Pyun, J.Y. Deep Recurrent Neural Networks for Human Activity Recognition. *Sensors* 2017, 17, 2556.
15. Y. Guan and T. Ploetz, "Ensembles of deep LSTM learners for activity recognition using wearables," 2017, arXiv:1703.09370. [Online]. Available: <http://arxiv.org/abs/1703.09370>.
16. Wu J, Wang J, Ling Y, et al. An advanced Hybrid Technique of DCS and JSRC for Telemonitoring of Multi-sensor Gait Pattern[J]. *Sensors*, 2017, 17(12):2764-.
17. Muhammad Usman Shahid Khan, Assad Abbas, Mazhar Ali, Muhammad Jawad, Samee U Khan, Keqin Li, and Albert Y Zomaya. On the correlation of sensor location and human activity recognition in body area networks (bans). *IEEE Systems Journal*, 12(1):82–91, 2018.
18. Alexander, Turner, Stephen, et al. The Classification of Minor Gait Alterations Using Wearable Sensors and Deep Learning.[J]. *IEEE Transactions on Bio Medical Engineering*, 2019.
19. F.J.Ordonez, D. Roggen, Deep convolutional and lstm recurrent neural networks for multimodal wearable activity recognition, *Sensors*, 16(1): 115, 2016.
20. Omid D , Mojtaba T , Raghvendar C V . IMU-Based Gait Recognition Using Convolutional Neural Networks and Multi-Sensor Fusion[J]. *Sensors*, 2017, 17(12):2735.
21. Kun Wang, Jun He, and Lei Zhang. Attention-based Convolutional Neural Network for Weakly Labeled Human Activities Recognition with Wearable Sensors. *IEEE Sensors Journal*, 19(17):7598–7604, September 2019. arXiv: 1903.10909.
22. R. Chavarriaga, H. Sagha, A. Calatroni, The Opportunity challenge: A benchmark database for on-body sensor-based activity recognition, *Pattern Recognition Letters*, 34(15): 2033–2042, 2013.
23. B. Sun, M. Liu, R. Zheng and S. Zhang, "Attention-based LSTM Network for Wearable Human Activity Recognition," 2019 Chinese Control Conference (CCC), Guangzhou, China, 2019, pp. 8677-8682, doi: 10.23919/ChiCC.2019.8865360.
24. Yang D , Liu Y , Zhou Z , et al. Decoding visual motions from EEG using attention-based RNN[J]. *Applied Sciences*, 2020, 10(16):5662.
25. Xia Y , Yao Z M , Ye Q , et al. A Dual-Modal Attention-Enhanced Deep Learning Network for Quantification of Parkinson's Disease Characteristics[J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2020, 28(1):42-51.
26. A. Y. Yang, R. Jafari, S. S. Sastry, and R. Bajcsy, "Distributed recognition of human actions using wearable motion sensor networks," *J. Ambient Intell. Smart Environments*, vol. 1, no. 2, pp. 103–115, 2009.
27. T. H. Nguyen, T. P. Pham, C. Q. Ngo, et al. A SVM Algorithm for Investigation of Tri-Accelerometer Based Falling Data. *American Journal of Signal Processing*, 2016; 6(2): 56-65.
28. R. Lemoyne and T. Mastroianni, "Implementation of a Smartphone as a Wearable and Wireless Gyroscope Platform for Machine Learning Classification of Hemiplegic Gait Through a Multilayer Perceptron Neural Network," 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA), Orlando, FL, 2018, pp. 946-950, doi: 10.1109/ICMLA.2018.00153.
29. Rodriguezmartin D, Sama A, Perezlopez C, et al. Home detection of freezing of gait using support vector machines through a single waist-worn triaxial accelerometer[J]. *PLOS ONE*, 2017, 12(2): 1-22.
30. Mahoney J M, Rhudy M B. Methodology and validation for identifying gait type using machine learning on IMU data[J]. *Journal of Medical Engineering & Technology*, 2019, 43(1): 25-32.
31. Mehrang S, J Pietilä, Korhonen I. An Activity Recognition Framework Deploying the Random Forest Classifier and A Single Optical Heart Rate Monitoring and Triaxial Accelerometer Wrist-Band[J]. *Sensors*, 2018, 18(3).
32. A. Ignatov, "Real-time human activity recognition from accelerometer data using convolutional neural networks," *Appl. Soft Comput.*, vol. 62, pp. 915–922, Jan. 2018.
33. Ward M, Min C H, Salamy H, et al. Dilated Temporal Convolutional Neural Network Architecture with Independent Component Layer for Human Activity Recognition[C]// 2019 26th IEEE International Conference on Electronics, Circuits and Systems (ICECS). IEEE, 2019.
34. Thapa, K.; Abdullah Al, Z.M.; Lamichhane, B.; Yang, S.-H. A Deep Machine Learning Method for Concurrent and Interleaved Human Activity Recognition. *Sensors* 2020, 20, 5770.
35. T. Lv, X. Wang, L. Jin, Y. Xiao and M. Song, "A Hybrid Network Based on Dense Connection and Weighted Feature Aggregation for Human Activity Recognition," in *IEEE Access*, vol. 8, pp. 68320-68332, 2020, doi: 10.1109/ACCESS.2020.2986246.
36. Xia K, Huang J, Wang H. LSTM-CNN Architecture for Human Activity Recognition[J]. *IEEE Access*, 2020, 8:56855-56866.
37. Ma H J , Li W , Zhang X , et al. AttnSense: Multi-level Attention Mechanism For Multimodal Human Activity Recognition[C]// Twenty-Eighth International Joint Conference on Artificial Intelligence IJCAI-19. 2019.
38. Gers F A, Schraudolph N N, Schmidhuber J, et al. Learning precise timing with lstm recurrent networks[J]. *Journal of Machine Learning Research*, 2003, 3(1): 115-143.
39. C. Raffel and D. P. Ellis, "Feed-forward networks with attention can solve some long-term memory problems," arXiv preprint arXiv:1512.08756, 2015.
40. Wang Y , Shuang C , Yu H . A Data Fusion-Based Hybrid Sensory System for Older People's Daily Activity and Daily Routine Recognition[J]. *IEEE Sensors Journal*, 2018, 18:6874-6888.
41. Wan, S.; Qi, L.; Xu, X.; Tong, C.; Gu, Z. Deep learning models for real time human activity recognition with smartphones. *Mob. Netw. Appl.* 2020, 25, 743–755.

42. Wang L K, Liu R Y. Human activity recognition based on wearable sensor using hierarchical deep LSTM networks[J]. *Circuits, Systems, and Signal Processing*, 2020, 39(2): 837-856.2
43. Abbaspour S, Fotouhi F, Sedaghatbaf A, et al. A comparative analysis of hybrid deep learning models for human activity recognition[J]. *Sensors*, 2020, 20(19): 5707.
44. Mahmud S, Tonmoy M, Bhaumik K K, et al. Human activity recognition from wearable sensor data using self-attention[J]. *arXiv preprint arXiv:2003.09018*, 2020.
45. Reiss, A.; Stricker, D. Introducing a New Benchmarked Dataset for Activity Monitoring. In *Proceedings of the 2012 16th International Symposium on Wearable Computers*, Newcastle, UK, 18-22 June 2012; pp. 108-109.
46. Zhang X, Chen X, Yao L, et al. Deep Neural Network Hyperparameter Optimization with Orthogonal Array Tunning [J]. 2019.
47. Liu Y, Jiang X, Sun T, et al. 3D Gait Recognition Based on a CNN-LSTM Network with the Fusion of SkeGEI and DA Features [C]// 2019 16th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS). IEEE, 2019.