


Article

FuseLNet: Fusion of Local and Global Information for Detection of Parkinson's Disease

Ming Chen ^{1,2} , Tao Ren ^{1,2}, Pihai Sun ^{1,3}, Jianfei Wu ^{1,3}, Jinfeng Zhang ^{1,3} and Aite Zhao ^{1,2,*}¹ College of Computer Science and Technology, Qingdao University, Qingdao 266071, China² Institute of Ubiquitous Networks and Urban Computing, Qingdao University, Qingdao 266071, China³ Turing Innovation Team, Qingdao University, Qingdao 266071, China

* Correspondence: zhaoaite@qdu.edu.cn

Abstract: In the past few years, the assessment of Parkinson's disease (PD) has mainly been based on the clinician's examination, the patient's medical history, and self-report. Parkinson's disease may be misdiagnosed due to a lack of clinical experience. Moreover, it is highly subjective and is not conducive to reflecting a true result. Due to the high incidence rate and increasing trend of PD, it is significant to use objective monitoring and diagnostic tools for accurate and timely diagnosis. In this paper, we designed a low-level feature extractor that uses convolutional layers to extract local information about an image and a high-level feature extractor that extracts global information about an image through the autofocus mechanism. PD is detected by fusing local and global information. The model is trained and evaluated on two publicly available datasets. Experiments have shown that our model has a strong advantage in diagnosing whether people have PD; gait-based analysis and recognition can also provide effective evidence for the early diagnosis of PD.

Keywords: Parkinson's disease; gait recognition; CNN; Transformer



Citation: Chen, M.; Ren, T.; Sun, P.; Wu, J.; Zhang, J.; Zhao, A. FuseLNet: Fusion of Local and Global Information for Detection of Parkinson's Disease. *Information* **2023**, *14*, 119. <https://doi.org/10.3390/info14020119>

Academic Editors: Jiquan Ma, Geng Chen, Hui Cui and Desi Shang

Received: 18 November 2022

Revised: 11 February 2023

Accepted: 12 February 2023

Published: 13 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Middle-aged and elderly people are more likely to suffer from Parkinson's disease (PD), a degenerative condition of the central nervous system (CNS). The main characteristics are slow movements, tremors in the hands and feet or other parts of the body, and loss of flexibility and stiffness in the body. In most cases, the condition changes over time, with different symptoms appearing at different times and the condition becoming more severe. When medical professionals diagnose PD, the Unified Parkinsons Disease Rating Scale (UPDRS) is primarily adopted, but UPDRS is more complex and contains more levels. In addition, it is difficult for UPDRS to assess the early mild signs and symptoms of Parkinson's patients. Therefore, it is very important to use computer-aided algorithms to diagnose PD. Gait recognition is a new type of biometric identification technology, which captures the characteristics displayed by various joints and parts of the body during walking. The goal of gait recognition is to extract these local and global features for correlation analysis and processing. Moreover, gait recognition can be obtained without the cooperation of the identified person. It is stable for a while. Gait analysis can not only provide a basis for the early diagnosis of PD but also furnish real-time monitoring and early warnings of changes.

In the field of gait recognition, most datasets are obtained by various sensors. For PD detection in the gait data obtained using the sensor, positive results were obtained in [1–6]. However, the use of motion sensors comes with a high cost because the sensors are not only expensive but also require professional guidance for their wearing and calibration, which is impractical for PD patients and inconvenient for large-scale promotion and application. With recent developments in deep learning technologies [7–10], vision-based methods have attracted increasing attention, such as stereo cameras, infrared thermal imaging, structured light, time-of-flight systems, or simply 2D cameras. These methods eliminate the need

to wear additional equipment that reduce the compliance of patients with Parkinson's disease [11]. Furthermore, using the structural features of the acquired images has many advantages, such as low cost, no contact, and easy access. The emergence of AlexNet [12] makes convolutional neural networks (CNN) begin to dominate in computer vision. It is well known that convolutional neural networks are highly effective at extracting local image features from the image. However, they still have some limitations in capturing global feature representations. Transformers [13] have been making a splash in the field of computer vision recently. The self-attention mechanism used by transformers solves the long-term dependence on input and output, which makes it achieve promising results in the field of image classification. However, the local information collection of transformers is not as strong as that of CNNs. Therefore, neural networks and Transformers complementing each other is the inevitable trend in the development of deep learning.

In this paper, we have designed two modules. The low-level feature extractor is responsible for extracting the local information of the gait image, and the high-level feature extractor is responsible for extracting the global information of the gait image. Through the fusion information, we can assess whether people have PD. The model has been trained on the GAIT-IST and GAIT-IT public datasets and has well captured the dynamic characteristics of PD patients. The implementation results show that the model is more effective and accurate than the existing methods.

2. Related Work

The use of sensor data to detect gait abnormalities in PD has gradually improved with the development of sensors. Zhang et al. [14] used an accelerometer attached to the lower back in 12 PD patients with freezing of gait (FOG) to collect acceleration signals on the patient and used AdaBoost to build two FOG prediction models, which can better predict gait abnormalities based on impaired gait features. The work of Borzì et al. [15] installed two inertial sensors on each tibia of 11 PD patients receiving and not receiving dopaminergic therapy, and they performed stepwise segmentation of angular velocity signals and extracted features from time and frequency domains. A wrapper approach was employed for feature selection, and different machine-learning classifiers were optimized to capture FOG and pre-FOG events.

To make full use of spatiotemporal information in gait, Zhao et al. [16] proposed a two-channel model that combined CNN and long short-term memory (LSTM) to better capture spatiotemporal information from the gait data of PD patients and used the captured information for PD detection. Tong et al. [17] proposed a novel spatial-temporal deep neural network (STDNN), which included a temporal feature network (TFN) and a spatial feature network (SFN). In TFN, a feature sub-network is used to extract low-level edge features from gait silhouettes. These features are input to the spatial-temporal gradient network (STGN) that adopts a spatial-temporal gradient (STG) unit and LSTM unit to extract the spatial-temporal gradient features. In SFN, spatial features of gait sequences are extracted by multilayer convolutional neural networks from a gait energy image (GEI). After training, TFN and SFN are employed to extract temporal and spatial features, respectively, which are applied to multi-view gait recognition. Experimental results show that STDNN has great potential in future practical applications. Currently, most gait analysis methods based on 2DRGB rely on compact gait representations, such as gait energy maps. To solve the problem that gait energy maps cannot fully capture the temporal information and dependencies between gait movements, a spatiotemporal deep learning method [18–21] was proposed by [22], which used the selection of keyframes to represent the gait cycle. This method combined CNN and recurrent neural networks to extract spatiotemporal features from gait sequences represented by selected binary contour keyframes. The method also used the GAIT-IT dataset for training and the GAIT-IST dataset for testing. The cross-dataset scenario showed that the proposed spatiotemporal deep learning gait classification framework has good generalization capabilities.

Nowadays, multimodal learning shines in the field of gait recognition. Because multimodal learning can aggregate information from multiple sources, the representation learned by the model is more complete. Zhao et al. [23] designed a spatial feature extractor to extract frequency and time domain features from heterogeneous data. Their work used two datasets; one was the vertical ground reaction force (VGRF) sequence collected by the force sensor, and the other was the acceleration data collected by the acceleration sensor. To capture temporal information from the two-modal data, they designed a novel correlation memory neural network (CorrMNN) architecture to extract temporal features and embed a multi-switch discriminator to model the fused features of each class. Cai et al. [11] proposed a new fusion strategy to integrate contour and skeleton information into a single compact contour skeleton representation of gait, which was demonstrated on GaitPart using the CASIA-B dataset to outperform any single ensemble-based or sequence-based performance. Because silhouette and pose heatmaps complement each other, the silhouettes of the same pedestrian under different clothing conditions are similar, and the pose heatmaps of different pedestrians are very similar. Li et al. [24] combined silhouette and pose heatmaps to construct a part-based multimodal architecture. They proposed a set transformer module network to fuse multimodal visual information and demonstrated the effectiveness of the multimodal approach on the CASIA-B and GREW datasets.

In addition, computer vision-based [25–28] systems have certain advantages in detecting PD gait [29] because it allows gait acquisition in an unconstrained environment. Maachi et al. [30] proposed a novel intelligent Parkinson’s detection system based on deep learning technology to analyze gait information, and they used a one-dimensional convolutional neural network (1D-Convnet) to build a deep neural network (DNN) classifier to process 18 1D signals from foot sensors measuring VGRF. A fully connected network was then used to concatenate the cascaded outputs of the 1D Convents to obtain the final classification. They used UPDRS to test the application of their algorithm in PD detection and disease severity prediction. Experimental results showed that the system was effective in detecting PD based on gait data. Recognizing the importance of telemedicine, Paulo et al. [31] designed a novel web application system for remote gait diagnosis. The system allowed users to upload videos of their walking and perform web services to classify their gait. Not only did it reduce memory requirements, but it also sped up execution time.

Inspired by the above methods, we designed a model that can fuse local and global information from RGB images and use this information to accurately predict whether PD exists. Fused information can expand the spatial and temporal information contained in the image, reduce uncertainty, increase reliability, and improve the robustness of the model.

3. Method

The FuseLGNNet architecture is shown in Figure 1. First, we output the RGB input image to the Patch Partition module for partitioning and set 4×4 adjacent pixels as a patch; then, the feature size of each patch was $4 \times 4 \times 3 = 48$. Then, we input this to the linear embedding layer to linearly transform the channel of each patch to any dimension (denoted C). Then, feature maps of different sizes were constructed through four stages, such as Swim Transformer, except that Stage 1 passed through a linear embedding layer first, and the remaining three stages were scaled down through a patch merging layer first. The repeatedly stacking FuseLG blocks, the windows multi-head self-attention (W-MSA) structure, and the shifted windows multi-head self-attention (SW-MSA) structure used herein were used in pairs. Therefore, the number of stacked FuseLG blocks was an even number. Here, we also introduced convolutional layers to obtain local features. By fusing local and global features, local regions help to characterize objects in more detail, so fusing local and global image features will achieve better classification performance. The fused feature is represented by Equation (1), where F is a fusion function.

$$\text{Output}(X) = F(\text{Conv}(X_i), (X_j)) \quad (1)$$

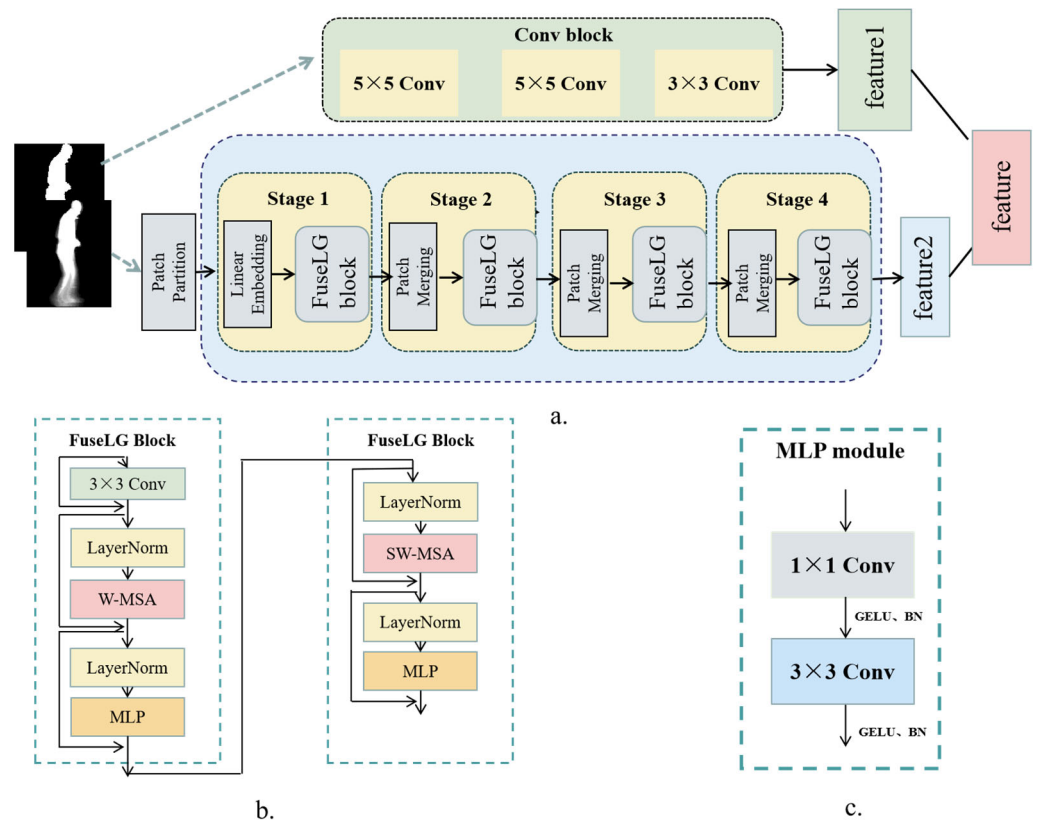


Figure 1. The structure of the proposed FuseLGNNet. (a) corresponds to the model overview; (b) represents the FuseLGN block; (c) depicts the MLP module. Note that W-MSA and SW-MSA are used in pairs.

Conv block: the CNN channel consisted of three convolutional layers. The convolution kernel size of the first two convolutional layers was 5×5 , and the convolution kernel size of the latter one was 3×3 . GELU and LayerNorm were set after each convolutional layer, respectively. The convolutional layer’s job was to extract features from the image. The input image was converted into a computer-readable matrix. The new matrix was a feature. Each convolution kernel generated a feature map after the sliding calculation. The height and width of the newly generated feature map are shown as follows:

$$H_{out} = \frac{W_{in} + 2P_h - K_h}{S_h} + 1 \tag{2}$$

$$W_{out} = \frac{W_{in} + 2P_w - K_w}{S_w} + 1 \tag{3}$$

where P denotes padding, K represents the size of the convolution kernel, and S depicts the size of the stripe.

Patch Merging: We input a 4×4 single channel feature map into patch merging. The patch merging module split every 2×2 adjacent pixel into a patch and then stitched together the pixels at the same position in each patch to obtain 4 feature maps. Before passing through the LayerNorm layer, the four feature maps were spliced in the channel direction, and finally, a fully connected layer was used to make a linear change in the channel direction of the feature map. At this time, the depth of the feature map changed from C to $C/2$. After passing through the patch merging layer, the height and width of the feature map was halved, and the number of channels was doubled.

W-MSA: The W-MSA module was introduced to reduce the amount of computation. As shown in Figure 2, the standard multi-head self-attention (MSA) module was used on the left. During the self-attention calculation process, each token in the feature map needs to be calculated with all pixels. However, on the right side of the figure, when using the

W-MSA module, the feature map was divided into windows according to the $M \times M$ size ($M = 2$ in the example), and then the internal processing of each window was performed using self-attention separately. The self-attention module is defined as follows:

$$Attention(Q, K, V) = SoftMax\left(\frac{QK^T}{\sqrt{d}}\right)V \tag{4}$$

$$SoftMax(y) = \frac{e^{y_i}}{\sum_{j=1}^k e^{y_j}} \tag{5}$$

Q: query, $Q \in X^{H \times W \times d_k}$, where H and W are the height and width of the input image, respectively.

K: key, $K \in X^{H \times W \times d_k}$, where d_k is the dimension of the key.

V: value, $V \in X^{H \times W \times d_v}$, where d_v is the dimension of the value.

k : the number of multiple outputs or categories of the neural network.

y : output vector; y_j is the value of the j th output or category in y , and i represents the category to be calculated at that time.

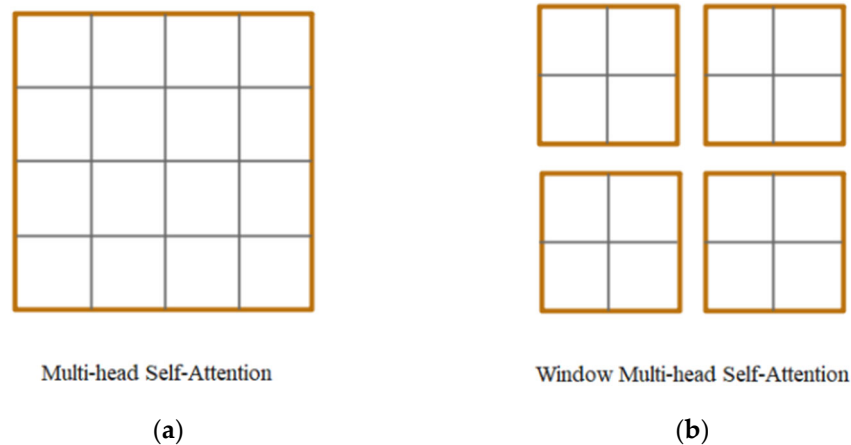


Figure 2. (a) Normal multi-head self-attention; (b) windows multi-head self-attention.

SW-MSA: The SW-MSA module was proposed to solve the problem of not being able to pass information between windows. As shown in Figure 3, W-MSA was used on the left (assuming layer C). Since W-MSA and SW-MSA were used in pairs, the C + 1 layer used SW-MSA (right). The window shift can be seen in Figure 3 (it can be understood that the window has shifted from the top left corner to the right and bottom, respectively, by $\lfloor \frac{M}{2} \rfloor$ pixels). For the offset window (right), for example, for a 2×4 window in the first row and the second column, it allows the two windows in the first row of layer C to communicate information and allows the same for the other windows. SW-MSA solves the problem that information cannot be exchanged between different windows.

FuseLG block: As shown in Figure 1, FuseLG added a 3×3 convolutional layer to extract local information before entering the LayerNorm layer and improved the cross-layer propagation capability of the gradient. In addition, the purpose of the residual structure was to stabilize network training. The information obtained was fed through the MSA module, which includes two LayerNorm layers and two MLP layers, each of which references redundant connections to stabilize the network training.

MLP block: Due to the translation invariance and local correlation features of CNN [32–34], locality focuses on adjacent points in the feature map, and translation invariance means using the same matching rules for different regions. CNN usually performs better on a small amount of data. Therefore, as shown in Figure 1, in this module, we added two convolutional layers for better performance.

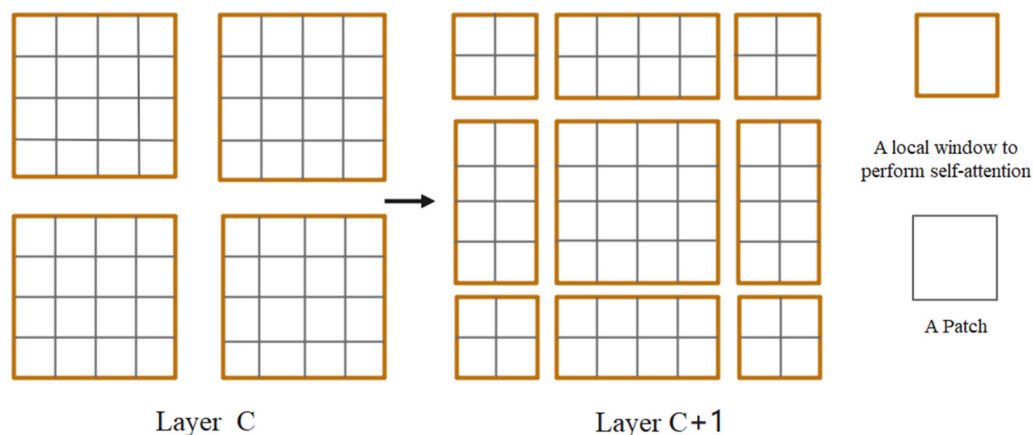


Figure 3. Illustration using a shift window method for self-attention calculation. W-MSA on the left; SW-MSA on the right.

4. Experiment

4.1. Dataset

Currently, in the field of Parkinson’s gait recognition, most gait datasets are obtained from sensors, often using gait signals from PhysioNet. Moreover, it is generally true that privacy and confidentiality become paramount when human data are involved. Therefore, there are not many public gait image databases dedicated to the study of PD gait. In this paper, we utilize two public datasets: the GAIT-IST dataset and the GAIT-IT dataset, which are presented separately below.

GAIT-IST dataset [35]: this dataset consists of image sequences obtained from 10 subjects simulating a pathological gait. The dataset includes sequences with 2 severity levels, 2 walking directions, and 2 repetitions per participant. For each sequence, the database includes multiple contour and skeletal images, as well as gait energy images and skeletal energy images.

GAIT-IT dataset [31]: this dataset is comprised of 21 volunteers aged 20 to 56; 2 participants simulated the type and severity of PD gait. The dataset captures two severity levels. Subjects provided 4 gait sequences for each severity level.

4.2. Training Details

In the experiment, the batch size was set to 8, the learning rate was set to 0.0001, and the epoch was set to 30. Because the experimental results of epochs less than 30 and higher than 30 are not as good as those equal to 30, our experimental epoch was set to 30. The ratio of training and test sets was 4:1. Due to the fewer gait data in PD datasets used, the experimental results were obtained by pre-training on the ImageNet-1k dataset to make the training cost smaller. This allowed us to achieve a faster convergence speed and effectively improved the performance of the model.

We conducted two sets of experiments, each of which included results for two walking directions (backward and forward). The first set of experiments used GEI from both datasets. GEI was obtained by averaging cropped, normalized size, and horizontally aligned binary silhouettes belonging to a gait cycle, as shown in Figure 4. The second set of experiments used silhouette maps from both datasets. In video processing, the background subtraction method is used to distinguish the background to extract the moving target’s gait profile, as shown in Figure 5. The purpose of the experiment is to detect PD with the gait of Parkinson’s patients and other control groups (CO). Table 1 describes the number of GEI and silhouettes of PD patients and other COs in the GAIT-IST and GAIT-IT datasets. The number of training iterations and batch size can be set according to the actual needs used when training the model.

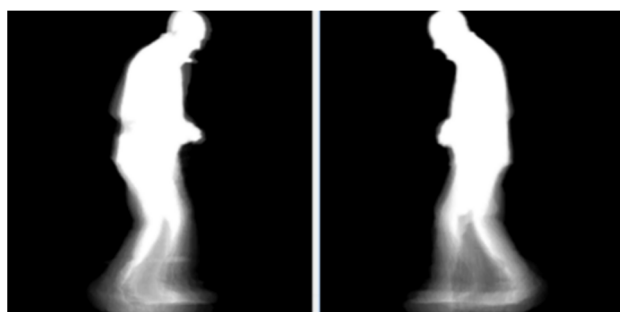


Figure 4. Examples of walking forward (**right**) and walking backward (**left**) of a gait energy map.



Figure 5. Examples of walking forward (**right**) and walking backward (**left**) of a gait contour map.

Table 1. The specific content of GAIT-IST and GAIT-IT datasets for PD patients and other controls.

DataSet	Direction	Data Type	PD Patients	CO
GAIT-IST	front	GEI Silhouettes	143 3996	312 12,216
	back	GEI Silhouettes	121 3471	304 11,767
GAIT-IT	front	GEI Silhouettes	330 23,858	998 85,882
	back	GEI Silhouettes	326 23,441	989 86,811

4.3. Experimental Results

We compared different deep learning methods for PD detection using a traditional classifier called k-nearest neighbor (KNN) and four deep learning methods, Efficient-Net, EfficientNetV2, Vision Transformer (ViT), and Swim Transformer (ST), all of which were compared with our model. Tables 2–5 show the results of the experiment.

Table 2. Results from gait energy maps using the GAIT-IST dataset.

	KNN	Efficient-Net	EfficientNetV2	ViT	ST	FuseLGNNet
front	91%	95.56%	96.67%	97.78%	98.81%	98.89%
back	91%	91.67%	92.86%	94.04%	96.43%	97.62%

Table 3. Results on gait silhouette maps using the GAIT-IST dataset.

	KNN	Efficient-Net	EfficientNetV2	ViT	ST	FuseLGNNet
front	56%	95.77%	95.80%	97.50%	98.94%	99.78%
back	87%	94.26%	95.37%	96.03%	98.66%	99.63%

Table 4. Results from gait energy maps using the GAIT-IT dataset.

	KNN	Efficient-Net	EfficientNetV2	ViT	ST	FuseLGNNet
front	87%	91.70%	93.96%	95.09%	96.24%	96.60%
back	82%	92.38%	93.51%	91.60%	95.42%	95.42%

Table 5. Results from gait silhouette maps using the GAIT-IT dataset.

	Efficient-Net	EfficientNetV2	ViT	ST	FuseLGNNet
front	91.51%	92.38%	93.81%	99.39%	99.78%
back	91.97%	92.52%	93.92%	98.58%	99.89%

The information given in Tables 2 and 3 is the result of using the GAIT-IST dataset. From the results, we created a confusion matrix, where 0 represents the control group and 1 represents Parkinson’s patients. Abnormal gaits in PD can be directly visualized by using a confusion matrix. As can be seen from Figure 6, which is divided into forward and backward walking, CO obtained 100% results when using the gait energy map. In the forward walking PD gait, one sample was incorrectly classified as CO, and two samples were incorrectly classified as CO in the backward walking PD gait group. In Figure 7, when using the contour plot, the forward CO has a better effect, and only one sample is poorly classified as having PD, while the backward CO has eight samples that are incorrectly classified as PD.

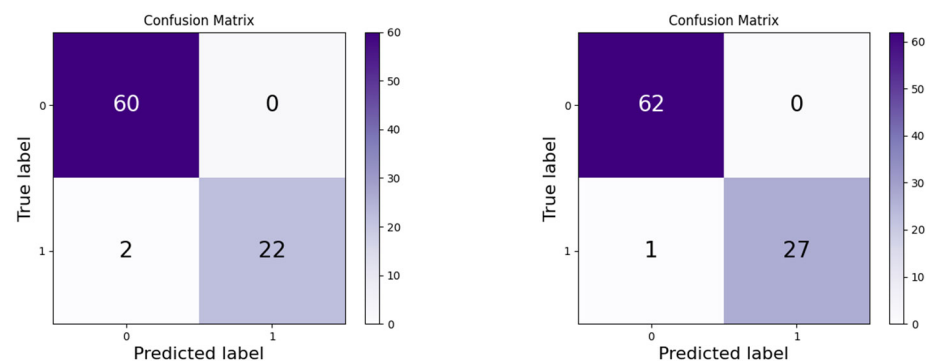


Figure 6. Confusion matrix for backward (left) and forward (right) classification results using IST dataset gait energy maps.

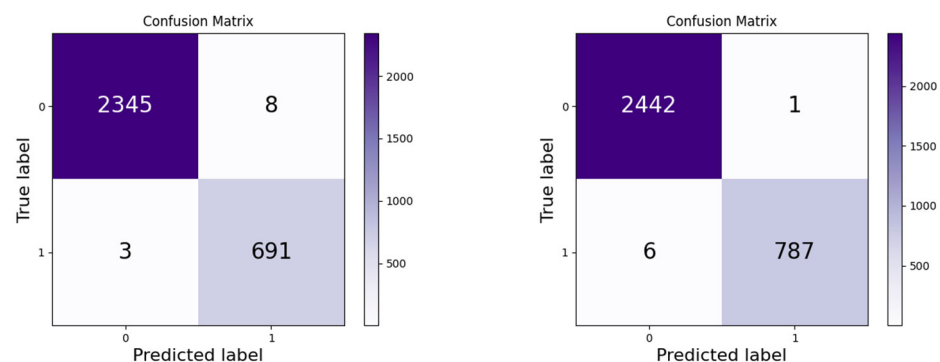


Figure 7. Confusion matrix for backward (left) and forward (right) classification results using IST dataset gait silhouette maps.

Tables 4 and 5 show the results of using the GAIT-IT dataset, and Figures 8 and 9 are the confusion matrix for the results. As can be seen from Figure 8, 11 CO samples that walked backward were incorrectly considered to have PD. In the forward walking PD

group, nine samples were incorrectly classified as CO. The reason why KNN is not used in Table 5 is that the number of images is too large.

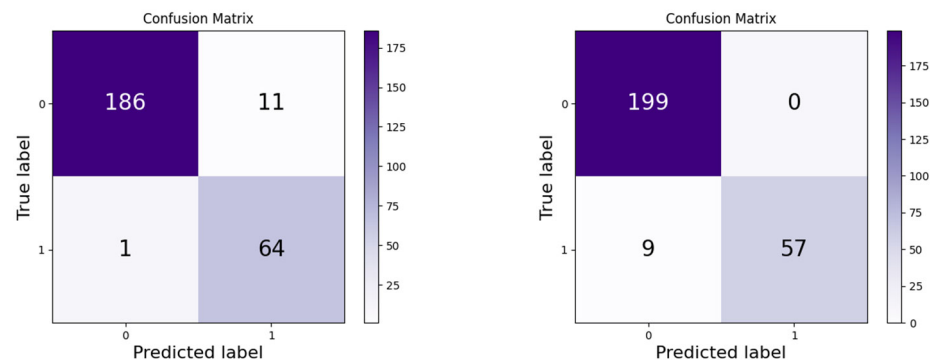


Figure 8. Confusion matrix for backward (left) and forward (right) classification results using IT dataset gait energy maps.

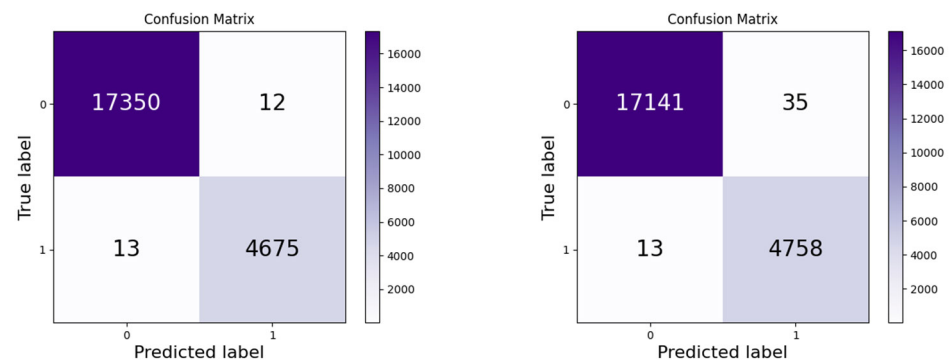


Figure 9. Confusion matrix for backward (left) and forward (right) classification results using IT dataset gait silhouette maps.

From the information given in the table, it can be seen that the classification results of KNN cannot be compared with the four deep learning methods. Among the deep learning methods in Tables 2–5, it can be seen that Swim Transformer achieved better performance, and its classification effect in PD detection has achieved a relatively ideal result. However, our model can not only obtain global features but also obtain local information by introducing a CNN. Utilizing local and global features enables our model to achieve the best classification effect. In addition, it can be seen from the experimental results that the results of different walking directions are different, and most of the data results of walking forward are better than those of walking backward. Because the GAIT-IST and GAIT-IT datasets correspond to a subject walking twice from left to right and from right to left from the side view, the number of consecutive key frames in the forward image is greater than that in the backward image.

5. Conclusions

In this paper, we use two gait image datasets, i.e., GAIT-IST and GAIT-IT, which simulate gait abnormalities in PD to detect whether a person has PD or not. In our work, two feature extractors are designed: one is a low-level feature extractor, and the other is a high-level feature extractor. The low-level feature extractor contains three convolutional layers for spatial feature extraction. In the high-level feature extractor, the self-attention mechanism in the FuseLGNNet architecture is used to quickly focus on important information with obvious gait changes. Furthermore, since convolution can use the characteristics of local information to improve the processing performance on the image, while the transformer's self-attention mechanism is superior in capturing global features, local and global features of various body parts in PD are obtained by combining the two features.

Our proposed method achieves excellent results on two publicly available image datasets compared with other state-of-the-art techniques, which solves the subjectivity brought by traditional assessment and improves the accuracy of diagnosis.

Due to privacy and ethical concerns, the current dataset only contained normal people simulating pathological gait, and there is not much research on PD severity classification. In future work, higher detection accuracy can be achieved by using multiple data types for fusion, such as skeleton contours and skeleton energy maps. Similarly, data on Parkinson's patients with different degrees of disease is also an issue to be addressed. The experiments in this paper show that the proposed model has advantages in diagnosing PD, so automatic and objective motion assessment through computer vision is of great significance for PD detection.

Author Contributions: Conceptualization, M.C. and T.R.; methodology, M.C.; software, M.C.; validation, A.Z., P.S., J.W. and J.Z.; formal analysis, M.C.; investigation, M.C.; resources, M.C.; data curation, M.C.; writing—original draft preparation, M.C.; writing—review and editing, M.C.; visualization, M.C.; supervision, A.Z.; project administration, A.Z.; funding acquisition, A.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported in part by the National Natural Science Foundation of China under Grant No. 62106117, the China Postdoctoral Science Foundation under Grant No. 2022M711741, and the Natural Science Foundation of Shandong Province under Grant No. ZR2021QF084.

Data Availability Statement: Data used in this article can be made available by the corresponding authors on reasonable request.

Acknowledgments: Ming Chen thanks Zhihao Xu from Qingdao University for her constructive comments about the article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Camps, J.; Samà, A.; Martín, M.; Rodríguez-Martín, D.; Pérez-López, C.; Arostegui, J.M.M.; Cabestany, J.; Català, A.; Alcaine, S.; Mestre, B.; et al. Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn inertial measurement unit. *Knowl. Based Syst.* **2018**, *139*, 119–131. [[CrossRef](#)]
2. Mileti, I.; Germanotta, M.; Alcaro, S.; Pacilli, A.; Imbimbo, I.; Petracca, M.; Erra, C.; Di Sipio, E.; Aprile, I.; Rossi, S.; et al. Gait partitioning methods in Parkinson's disease patients with motor fluctuations: A comparative analysis. In Proceedings of the 2017 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Rochester, MN, USA, 7–10 May 2017; pp. 402–407.
3. Chen, M.; Huang, B.; Xu, Y. Intelligent shoes for abnormal gait detection. In Proceedings of the 2008 IEEE International Conference on Robotics and Automation, Pasadena, CA, USA, 19–23 May 2008; pp. 2019–2024.
4. Nguyen, T.N.; Huynh, H.H.; Meunier, J. Skeleton-based abnormal gait detection. *Sensors* **2016**, *16*, 1792. [[CrossRef](#)] [[PubMed](#)]
5. Agostini, V.; Balestra, G.; Knaflitz, M. Segmentation and classification of gait cycles. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2013**, *22*, 946–952. [[CrossRef](#)] [[PubMed](#)]
6. Morris, T.R.; Cho, C.; Dilda, V.; Shine, J.M.; Naismith, S.L.; Lewis, S.J.G.; Moore, S.T. Clinical assessment of freezing of gait in Parkinson's disease from computer-generated animation. *Gait Posture* **2013**, *38*, 326–329. [[CrossRef](#)] [[PubMed](#)]
7. Lv, Z.; Li, J.; Dong, C.; Li, H.; Xu, Z. Deep learning in the COVID-19 epidemic: A deep model for urban traffic revitalization index. *Data Knowl. Eng.* **2021**, *135*, 101912. [[CrossRef](#)] [[PubMed](#)]
8. Xu, Z.; Lv, Z.; Li, J.; Sun, H.; Sheng, Z. A Novel Perspective on Travel Demand Prediction Considering Natural Environmental and Socioeconomic Factors. *IEEE Intell. Transp. Syst. Mag.* **2023**, *15*, 136–159. [[CrossRef](#)]
9. Cheng, Z.; Jian, S.; Rashidi, T.H.; Maghrebi, M.; Waller, S.T. Integrating household travel survey and social media data to improve the quality of od matrix: A comparative case study. *IEEE Trans. Intell. Transp. Syst.* **2020**, *21*, 2628–2636. [[CrossRef](#)]
10. Zhao, A.; Wu, H.; Chen, M.; Wang, N. DCACorrCapsNet: A deep channel-attention correlative capsule network for COVID-19 detection based on multi-source medical images. *IET Image Process.* **2022**, *53*, 126378. [[CrossRef](#)]
11. Cai, N.; Feng, S.; Gui, Q.; Zhao, L.; Pan, H.; Yin, J.; Lin, B. Hybrid silhouette-skeleton body representation for gait recognition. In Proceedings of the 2021 13th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), Hangzhou, China, 21–22 August 2021; pp. 216–220.
12. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. ImageNet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
13. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS'17), Long Beach, CA, USA, 4–9 December 2017; Curran Associates Inc.: Red Hook, NY, USA, 2017; pp. 6000–6010.

14. Zhang, Y.; Yan, W.; Yao, Y.; Ahmed, J.B.; Tan, Y.; Gu, D. Prediction of freezing of gait in patients with Parkinson's disease by identifying impaired gait patterns. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 591–600. [\[CrossRef\]](#)
15. Borzi, L.; Mazzetta, I.; Zampogna, A.; Suppa, A.; Olmo, G.; Irrera, F. Prediction of Freezing of Gait in Parkinson's Disease Using Wearables and Machine Learning. *Sensors* **2021**, *21*, 614. [\[CrossRef\]](#)
16. Zhao, A.; Qi, L.; Li, J.; Dong, J.; Yu, H. A hybrid spatio-temporal model for detection and severity rating of Parkinson's disease from gait data. *Neurocomputing* **2018**, *315*, 1–8. [\[CrossRef\]](#)
17. Tong, S.; Fu, Y.; Yue, X.; Ling, H. Multi-View Gait Recognition Based on a Spatial-Temporal Deep Neural Network. *IEEE Access* **2018**, *6*, 57583–57596. [\[CrossRef\]](#)
18. Li, H.; Lv, Z.; Li, J.; Xu, Z.; Yue, W.; Sun, H.; Sheng, Z. Traffic Flow Forecasting in the COVID-19, A Deep Spatial-Temporal Model Based on Discrete Wavelet Transformation. *ACM Trans. Knowl. Discov. Data* **2022**. [\[CrossRef\]](#)
19. Wang, Y.; Lv, Z.; Sheng, Z.; Sun, H.; Zhao, A. A deep spatio-temporal meta-learning model for urban traffic revitalization index prediction in the COVID-19 pandemic. *Adv. Eng. Inform.* **2022**, *53*, 101678. [\[CrossRef\]](#)
20. Lv, Z.; Li, J.; Dong, C.; Xu, Z. DeepSTF: A Deep Spatial–Temporal Forecast Model of Taxi Flow. *Comput. J.* **2021**, bxab178. [\[CrossRef\]](#)
21. Cheng, Z.; Rashidi, T.H.; Jian, S.; Maghrebi, M.; Waller, S.T.; Dixit, V. A Spatio-Temporal autocorrelation model for designing a carshare system using historical heterogeneous Data: Policy suggestion. *Transp. Res. Part C Emerg. Technol.* **2022**, *141*, 103758. [\[CrossRef\]](#)
22. Albuquerque, P.; Verlekar, T.T.; Correia, P.L.; Soares, L.D. A spatiotemporal deep learning approach for automatic pathological gait classification. *Sensors* **2021**, *21*, 6202. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Zhao, A.; Li, J.; Dong, J.; Qi, L.; Zhang, Q.; Li, N.; Wang, X.; Zhou, H. Multimodal Gait Recognition for Neurodegenerative Diseases. *IEEE Trans Cybern.* **2022**, *52*, 9439–9453. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Li, G.; Guo, L.; Zhang, R.; Qian, J.; Gao, S. TransGait: Multimodal-based gait recognition with set transformer. *Appl. Intell.* **2023**, *53*, 1535–1547. [\[CrossRef\]](#)
25. Liang, Y.; Li, Y.; Guo, J.; Li, Y. Resource Competition in Blockchain Networks Under Cloud and Device Enabled Participation. *IEEE Access* **2022**, *10*, 11979–11993. [\[CrossRef\]](#)
26. Xu, Z.; Lv, Z.; Li, J.; Shi, A. A Novel Approach for Predicting Water Demand with Complex Patterns Based on Ensemble Learning. *Water Resour. Manag.* **2022**, *36*, 4293–4312. [\[CrossRef\]](#)
27. Zhao, A.; Wang, Y.; Li, J. Transferable Self-Supervised Instance Learning for Sleep Recognition. *IEEE Trans. Multimed.* **2022**. [\[CrossRef\]](#)
28. Wang, Y.; Zhao, A.; Li, J.; Lv, Z.; Dong, C.; Li, H. Multi-attribute Graph Convolution Network for Regional Traffic Flow Prediction. *Neural Process. Lett.* **2022**. [\[CrossRef\]](#)
29. Zhao, A.; Li, J. Two-channel lstm for severity rating of parkinson's disease using 3d trajectory of hand motion. *Multimed. Tools Appl.* **2022**, *81*, 33851–33866. [\[CrossRef\]](#)
30. El Maachi, I.; Bilodeau, G.A.; Bouachir, W. Deep 1D-Convnet for accurate Parkinson disease detection and severity prediction from gait. *Expert Syst. Appl.* **2020**, *143*, 113075. [\[CrossRef\]](#)
31. Albuquerque, P.; Machado, J.P.; Verlekar, T.T.; Correia, P.L.; Soares, L.D. Remote Gait type classification system using markerless 2D video. *Diagnostics* **2021**, *11*, 1824. [\[CrossRef\]](#)
32. Lv, Z.; Li, J.; Li, H.; Xu, Z.; Wang, Y. Blind travel prediction based on obstacle avoidance in indoor scene. *Wirel. Commun. Mob. Comput.* **2021**, *2021*, 5536386. [\[CrossRef\]](#)
33. Lv, Z.; Li, J.; Dong, C.; Zhao, W. A deep spatial-temporal network for vehicle trajectory prediction. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Wireless Algorithms, Systems, and Applications, Qingdao, China, 13–15 September 2020*; Springer: Cham, Switzerland, 2020; pp. 359–369.
34. Lv, Z.; Li, J.; Xu, Z.; Wang, Y.; Li, H. Parallel computing of spatio-temporal model based on deep reinforcement learning. In *Lecture Notes in Computer Science, Proceedings of the International Conference on Wireless Algorithms, Systems, and Applications, Nanjing, China, 25–27 June 2021*; Springer: Cham, Switzerland, 2021; pp. 391–403.
35. Loureiro, J.; Correia, P.L. Using a skeleton gait energy image for pathological gait classification. In *Proceedings of the 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020), Buenos Aires, Argentina, 16–20 November 2020*; pp. 503–507.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.