

Article

WM-STGCN: A Novel Spatiotemporal Modeling Method for Parkinsonian Gait Recognition

Jieming Zhang ¹, Jongmin Lim ¹, Moon-Hyun Kim ², Sungwook Hur ¹ and Tai-Myoung Chung ^{1,*}

¹ Department of Computer Science and Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea; jieming2021@skku.edu (J.Z.); jm.lim@skku.edu (J.L.); swh500@gmail.com (S.H.)

² Hippo T&C Inc., Suwon 16419, Republic of Korea; mhkim@skku.edu

* Correspondence: tmchung@skku.edu

Abstract: Parkinson's disease (PD) is a neurodegenerative disorder that causes gait abnormalities. Early and accurate recognition of PD gait is crucial for effective treatment. Recently, deep learning techniques have shown promising results in PD gait analysis. However, most existing methods focus on severity estimation and frozen gait detection, while the recognition of Parkinsonian gait and normal gait from the forward video has not been reported. In this paper, we propose a novel spatiotemporal modeling method for PD gait recognition, named WM-STGCN, which utilizes a Weighted adjacency matrix with virtual connection and Multi-scale temporal convolution in a Spatiotemporal Graph Convolution Network. The weighted matrix enables different intensities to be assigned to different spatial features, including virtual connections, while the multi-scale temporal convolution helps to effectively capture the temporal features at different scales. Moreover, we employ various approaches to augment skeleton data. Experimental results show that our proposed method achieved the best accuracy of 87.1% and an F1 score of 92.85%, outperforming Long short-term memory (LSTM), K-nearest neighbors (KNN), Decision tree, AdaBoost, and ST-GCN models. Our proposed WM-STGCN provides an effective spatiotemporal modeling method for PD gait recognition that outperforms existing methods. It has the potential for clinical application in PD diagnosis and treatment.



Citation: Zhang, J.; Lim, J.; Kim, M.-H.; Hur, S.; Chung, T.-M.

WM-STGCN: A Novel Spatiotemporal Modeling Method for Parkinsonian Gait Recognition. *Sensors* **2023**, *23*, 4980. <https://doi.org/10.3390/s23104980>

Academic Editor: Arnold Baca

Received: 4 April 2023

Revised: 12 May 2023

Accepted: 19 May 2023

Published: 22 May 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/>).

Keywords: gait recognition; graph convolution network; Parkinson's disease

1. Introduction

With the increase in the aging population, age-related cognitive disorders have become more prevalent in recent years. Parkinson's disease (PD), a common progressive degenerative disease of the central nervous system, is characterized by movement disorders such as muscle stiffness, hand tremor, and slow movement. Early detection of PD is crucial for timely treatment and proper medication.

Gait is an important indicator of health status, and the detection of gait abnormalities can serve as an indication to obtain further medical assessment and treatment. Reference [1] observes that analyzing a patient's gait could be utilized as a clinical diagnostic tool to help doctors recognize two dementia subtypes, Alzheimer's disease (AD) and Lewy body disease (LBD). This study distinguished LBD and AD using four key gait features: step time variability, step length variability, step time asymmetry, and swing time asymmetry. Beauchet et al. [2] found that a high mean and coefficient of variation of stride length were characteristic of moderate dementia, while an increased coefficient of variation of stride duration was associated with mild cognitive impairment status. Mirelman A. et al. [3] studied the effect of Parkinson's disease on gait. They highlighted the gait features unique to Parkinson's disease. In the early stages of Parkinson's disease, patients have a slower gait and shorter stride length compared to healthy individuals. These gait changes are

common in patients with Parkinson's disease but are not unique, as many diseases reduce gait speed. However, decreased arm swing and smoothness of movement and increased interlimb asymmetry are more specific to Parkinson's disease and are usually the first motor symptoms. Gait stiffness and staggering may also occur in later stages.

Clinical gait assessment is a commonly used method for performing gait analysis, which is an assessment performed by a clinician. Specifically, the physician needs to observe the patient's walking performance and then give a score based on criteria of the Unified Parkinson's Disease Rating Scale (UPDRS) [4] and Simpson–Angus Scale (SAS) [5]. Moreover, utilizing different types of sensors is a popular method. For example, sensors are embedded in the shoe insoles to measure the pressure of the foot against the ground while walking [6]; inertial measurement units and goniometers are fixed to joints, such as the waist and elbow, to measure the walking speed and acceleration [7]. Moreover, some studies have proposed video-based methods [8–10]. For example, reflective markers are attached to diverse locations on the human body. The location and trajectory of the markers are analyzed to provide kinematic information by recording with a digital camera. The Vicon Vantage system [10] requires about 8–14 high-precision cameras to provide accurate 3D motion data for gait analysis.

These existing gait analysis methods either require specialist assessment or particular sensors and equipment. It is too costly to deploy such systems. Furthermore, constructing a specific testing environment and training a team to calibrate the system and manage complex data necessitate substantial investment.

To solve this issue, a convenient, low-cost, and clinically practical method is needed to recognize Parkinsonian gait. In clinical practice, Parkinson's disease screening, follow-up, regular examination, and evaluation of treatment efficacy can be performed in a way that is easily implemented in a clinical setting and is both feasible and effective for patients. With advancements in computer vision, advanced techniques, such as human pose estimation algorithms, have made remarkable progress. Pose estimation is a process that involves localizing a person's joints in an image or video, and it has been applied to vision-based gait analysis. Previous work on vision-based gait assessment explored the use of the Microsoft Kinect sensor, thus using the 3D joint position provided by the system to analyze Parkinson's disease gait [11,12]. However, due to the technical limitations of the Kinect depth sensor, 3D joint positions can only be accurately extracted when the participant is located between (0.5 and 4.5) meters from the sensor, which limits the scenarios that can be widely used [13,14].

Recently, there has been an upsurge of interest amongst researchers in conducting gait analysis on conventional color video, which eliminates the requirement for depth sensors and enables the analysis of whole walking durations using a solitary camera. The emergence of novel computer vision techniques and machine learning algorithms has enabled more robust and automated analysis of video data captured by consumer-grade devices. In particular, advanced human pose estimation libraries, such as OpenPose, Detectron, and AlphaPose, have demonstrated their proficiency in extracting precise 2D joint pixel coordinates from video recordings [15–17]. Prior research has investigated the utilization of 2D joint trajectories to compute domain-specific features for the identification of Parkinsonian gait and dyskinesia rating from color videos, as highlighted in Refs. [18–21]. Moreover, the study conducted by Lu et al. [22] delved into the utilization of 3D joint trajectories extracted from video for predicting gait scores related to Parkinson's syndrome.

Model training in deep learning requires an extensive amount of data. However, there are various restrictions on medical sample acquisition: video collection is restricted by laws and patient privacy, while clinicians are not sufficiently motivated to record patients walking data. The lack of data hinders the application of deep learning. An alternative approach to obtaining real data is to generate synthetic data [23,24]. For example, random noise can be added to existing data, thus extending the available real data and training deep learning models [25]. Hence, data augmentation may be a valuable tool to overcome the inaccessibility of real data in the medical field [26].

Moreover, the input data in the spatial domain is skeletal data, which can be represented in graphical form, while convolution functions on the time axis can be used to capture temporal features such as joint dynamics (frequency, velocity). Naturally, the spatiotemporal graph convolutional network (ST-GCN) [27] is a well-suited model, as it leverages the inherent graph structure of human skeletons, providing an efficient mechanism for learning directly from joint trajectories. The advantage is that it is no longer necessary to develop and compute engineered gait features from joint trajectories, as ST-GCN can learn to utilize the most significant aspects of gait patterns directly from joint trajectories. ST-GCNs have been effectively combined with human pose estimation libraries to score Parkinsonian leg agility [28]. However, the use of these models to recognize Parkinsonian gait directly on a forward video remains unexplored.

In this paper, we hypothesize that Parkinson's patients have unique gait features that reflect disease-specific cognitive features and underlying pathology. We focus on developing a novel video-based Parkinsonian gait recognition method, using the skeleton and joint location from pose estimation to extract gait features and detect PD gait. The correct identification of brain damage diseases is very useful for clinicians to design appropriate treatment methods.

The present work offers major contributions in three aspects: (1) We propose to use a novel spatiotemporal modeling method based on skeleton data to recognize Parkinsonian gait; in addition, we construct a graph neural network to capture the topological properties of the human skeleton; (2) We design the weighted matrix with virtual connections to meet the specific demands in gait skeleton modeling and propose a multi-scale temporal convolution network to improve the temporal aggregation capability; and (3) An experiment on the dataset shows that compared to other machine learning methods, the proposed model achieves superior performance.

2. Related Work

This section provides a review of related works from two perspectives: gait patterns analysis and Parkinson's gait analysis using machine learning.

2.1. Gait Patterns Analysis

In the gait analysis domain, two main data modalities are commonly employed: sensor-based and vision-based approaches. The promising performance of sensors has drawn interest in their application to gait analysis. Lou et al. [29] developed an in-shoe wireless plantar pressure measurement system with a flexible pressure sensor embedded to capture plantar pressure distribution for quantitative gait analysis. Camps et al. [30] proposed to detect the freezing of gait in Parkinson's disease patients by using a waist-worn inertial measurement unit (IMU). Seifert et al. [31] used radar micro-Doppler signatures to classify different walking styles. Although the sensor-based approach has demonstrated the ability to reflect human kinematics, the need for specific sensors or devices and their requirement to be worn on the human body have limited their convenience in some applications. The vision-based approaches are more convenient and only require cameras for data collection. Prakash et al. [32] utilized an RGB camera to capture joint coordinates from five reflective markers attached to the body during walking, while Seifollahi et al. [33] employed a markerless system using Kinect cameras to capture RGB-D data to detect Alzheimer's disease from gait.

Recently, skeleton data have become a popular choice in gait analysis. Some studies have utilized the Microsoft Kinect camera and its camera SDK to generate 3D skeleton data. For example, Nguyen et al. [34] proposed an approach to predict the gait abnormality index by using the joint coordinates of the 3D skeleton as inputs for auto-encoders and then distinguishing abnormal gaits based on reconstruction errors. Elsewhere, Jun et al. [35] proposed a two-recurrent neural network-based autoencoder to extract features from 3D skeleton data for abnormal gait recognition and assessed the performance of discriminative

models with these features. In our study, we propose to extract gait features using the skeleton and joint locations obtained from pose estimation.

2.2. Parkinson's Gait Analysis Using Machine Learning

Researchers have experimented with data collected by various sensors for Parkinson's disease gait analysis. Shalin et al. [36] utilized LSTM to detect freezing of gait (FOG) in PD from plantar pressure data. The experiment required participants with PD to wear pressure-sensitive insole sensors while walking a predefined, provoking path. Labeling was then performed, and 16 features were manually extracted. The best FOG detection model had an average sensitivity of 82.1% and an average specificity of 89.5%. However, these particular sensors and devices are too costly to deploy. In addition, they need to be operated on in a specific place under the guidance of a professional doctor.

Due to the advances in action recognition [27,37–41], a growing number of researchers have applied it to gait recognition [42–44], and several studies have used video-based methods to automatically analyze dyskinesia symptoms in PD patients. Mandy Lu et al. [21] proposed a novel temporal convolutional neural network model to assess PD severity from gait videos, which extracts the 3D body skeleton of the participant and estimates the MDS–UPDRS score. Li et al. [20] extracted human joint sequences from videos recorded by PD patients and calculated motion features using a pose estimation method. Then, they applied random forest for multiclass classification and assessed clinical scores based on the UPDRS and Unified Dyskinesia Rating Scale (UDysRS) [45]. Sabo et al. [19] proposed the utilization of a spatiotemporal graph convolutional network (ST–GCN) architecture and training procedure to predict clinical scores of Parkinson's disease gait from videos of dementia patients. K. Hu et al. [46] proposed a graph convolutional neural network architecture that represents each video as a directed graph to detect PD frozen gait. The experimental results based on the analysis of over 100 videos collected from 45 patients during clinical evaluation have indicated that the proposed method performs well, achieving an AUC of 0.887.

Based on our literature survey, although several studies have evaluated gait videos of Parkinsonian patients, their focus has primarily been on estimating Parkinson's severity and detecting frozen gait, while recognizing PD gait versus normal gait from the forward video has yet to be reported. Additionally, traditional engineering solutions have proven insufficient to accurately assess motor function based on videos. To address this limitation, we have developed a novel deep-learning based framework to extract skeletal sequence features from forward videos of PD patients, with the ultimate goal of recognizing Parkinson's gait.

3. Materials and Methods

This part explains our dataset and how the data was preprocessed, and then the model is explained clearly. Figure 1 shows our methodology framework. Our method consists of two phases: feature extraction and gait recognition. Firstly, we augmented the video and then used OpenPose to extract skeleton data. In addition, we augmented the joint coordination space. Secondly, the skeleton data was constructed into a spatiotemporal graph and input to WM–STGCN, and the information in both temporal and spatial dimensions was aggregated by the spatiotemporal graph convolution operation to perform Parkinsonian gait recognition.

3.1. Dataset

We collected the data in an enclosed room for the normal walking video. The wall color was white, with no other colors. The space was 8 m long and 3 m wide, so the cameras could be located. Figure 2 shows the data collection environment. We used two Samsung mobile phones as our recording devices. The video parameters were 1080 × 1920 pixels at 30 Hz. As depicted in Figure 3, the cameras should be placed in forward of the patient's walking direction.

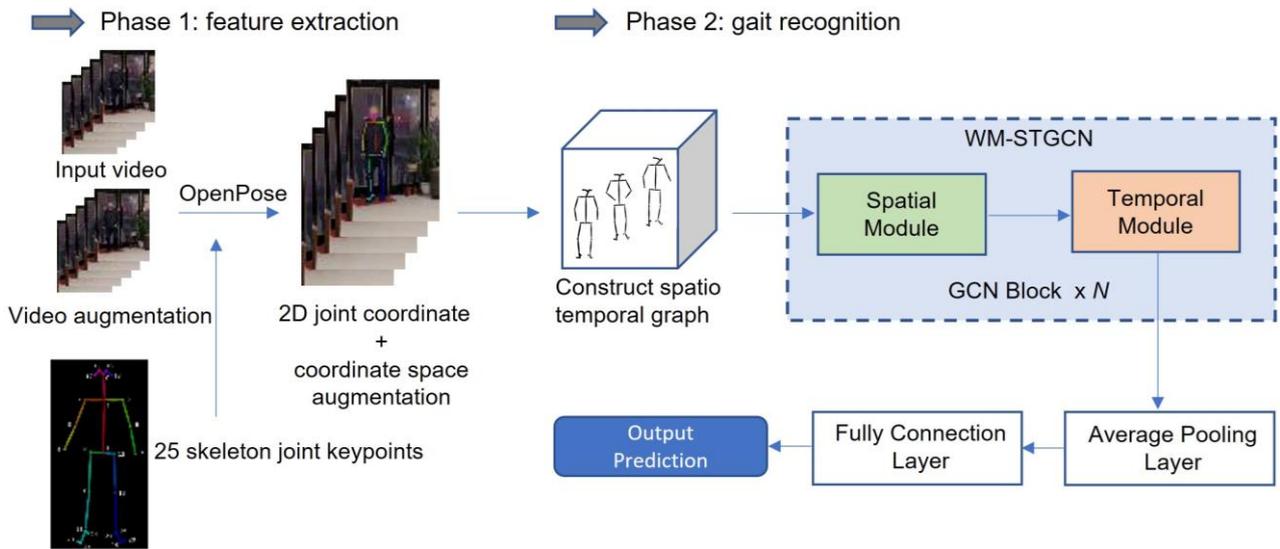


Figure 1. The overall framework of the proposed method.



Figure 2. Experiment environment.

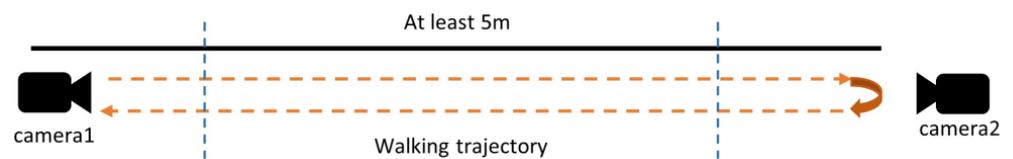


Figure 3. Walking trajectory and camera locations.

Participants wear their comfortable clothes (recommended wear: pants and sweatshirt or T-shirt) and walk straight from beginning to end, then turn around and walk back.

During the walk, participants should walk at a normal speed, and for each sequence, the time length is kept to approximately (10 to 20) seconds.

After that, we processed the data to make sure the content was only the frontal view walking. Table 1 lists the collected data details.

Table 1. Details of the collected data.

Type	Normal
Number of Participants	50
Mean height	174.6 cm
Resolution	1080 × 1920 pxl
Frame rate	30 fps
Length of sample video	10–20 s
Steps of sample video	6–8 steps

We obtained six videos from YouTube for Parkinsonian walking data [47–52]. To ensure clarity, their resolution was at least 652×894 pixels, and the frame rate was 30 fps. The video clips of a Parkinson's patient walking toward the camera without the assistance of others were selected as the data used in our study.

3.2. Data Augmentation

The difficulty in obtaining videos of PD patients walking resulted in a low amount of data. To reduce the class imbalance, we needed to perform data augmentation. Additionally, augmentation can increase the generalization capability of the system. There are two approaches: video augmentation and joint coordinate space augmentation; Figure 4 shows the augmentation pipeline.

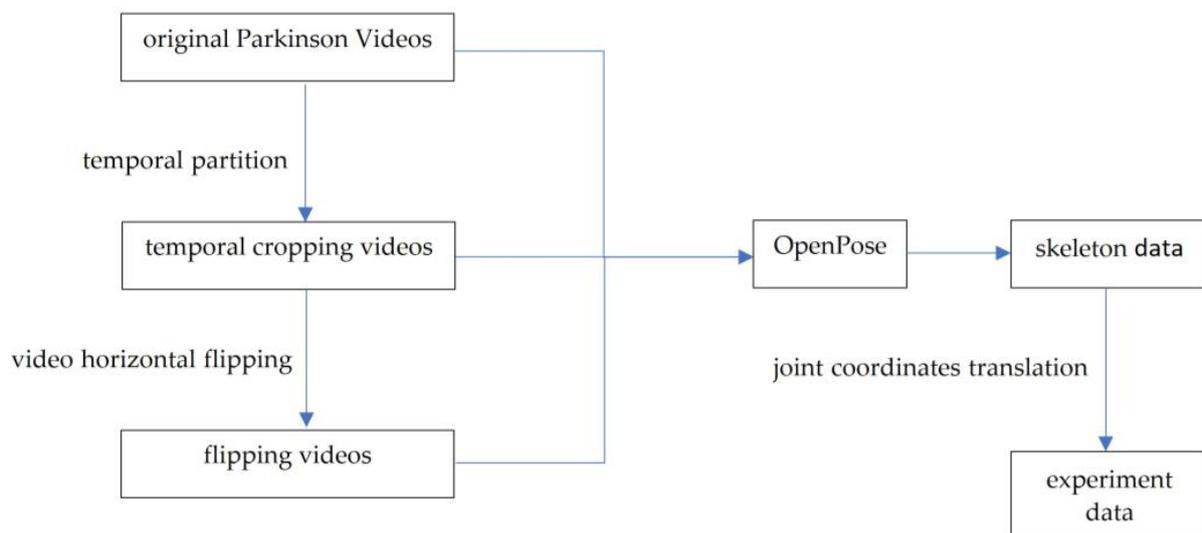


Figure 4. Data augmentation pipeline.

We first used temporal partition to crop the original videos, then flipped the video horizontally. After extracting skeleton data, we made joint coordinates space augmentation by translating and adding Gaussian noise.

3.2.1. Video Augmentation

In the video augmentation field, temporal partition and horizontal flipping are two effective tools to augment data on videos.

We used temporal cropping to implement partition: each video sequence of length l was temporally cropped to a fixed new sequence length k , where $k = 90$ frames, as shown in Figure 5. This allowed a video sequence to be partitioned with an interval of 20 frames. For horizontal flipping, we flipped the entire video to obtain a new video sequence.

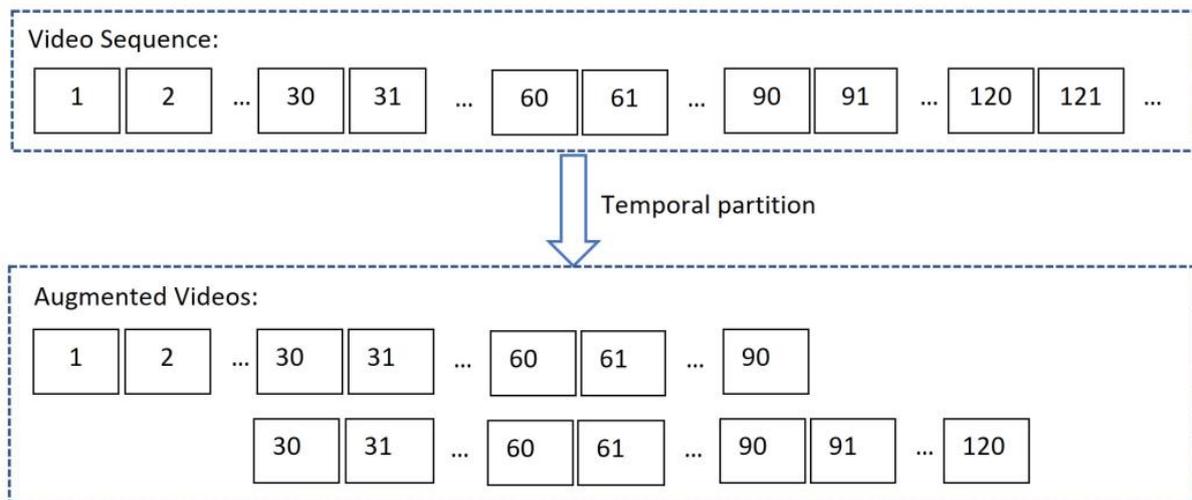


Figure 5. Temporal partition.

3.2.2. Joint Coordinate Space Augmentation

After we extracted skeleton data from videos, a natural idea to augment data is to directly focus on the joint coordinates space. The skeleton data are stored as a dictionary data structure (JSON format files) to allow key and value search to modify the joint value.

We performed the coordinate space augmentation processing in the following two ways:

1. Joint coordinates were translated in the horizontal direction to a new position to allow change in the viewing angle. As shown in Figure 6a, we set the offset $\Delta = (-0.1, 0.15, 0.2)$, which means we translated the coordinates of the skeleton data with Δ .
2. Gaussian noise was added to the joint coordinate. Figure 6b shows that the addition of appropriate noise perturbs the skeletal data within a certain range, which allows errors in joint coordinate calculation—for example, interference with the environment, such as background color or cloth texture. We set three Gaussian parameter groups for the experiment for $\varphi(\mu, \sigma)$, $\mu = 0$, $\sigma = (0.01, 0.05, 0.1)$.

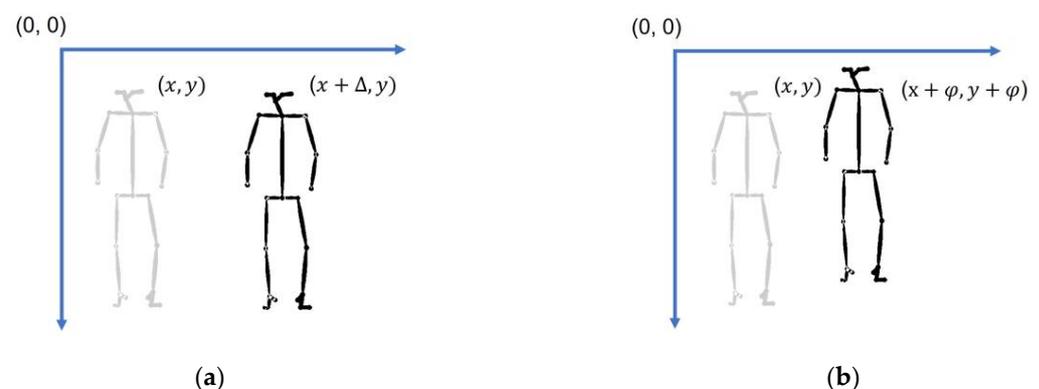


Figure 6. Joint coordinate space augmentation. (a) Joint coordinate translation; (b) Addition of Gaussian noise to the skeleton data.

3.3. Data Preprocessing

3.3.1. Skeleton Data Extraction

The video sequences are processed to extract 2D skeleton features, where each frame is analyzed using OpenPose, owing to its proficient and robust detection capabilities for 2D joint landmarks in upright individuals. We extract 25 landmarks in the OpenPose-skeleton format, which encompass 2D coordinate values (x , y) and an associated confidence score c that indicates the level of estimation reliability.

The key points roughly correspond to body parts: 0: Nose, 1: Neck, 2: RShoulder, 3: RElbow, 4: RWrist, 5: LShoulder, 6: LElbow, 7: LWrist, 8: MidHip, 9: RHip, 10: RKnee, 11: RAnkle, 12: LHip, 13: LKnee, 14: LAnkle, 15: REye, 16: LEye, 17: REar, 18: LEar, 19: LBigToe, 20: LSmallToe, 21: LHeel, 22: RBigToe, 23: RSmallToe, 24: RHeel (L, left; R, right; Mid, middle).

To obtain sequential key-point coordinate data for each gait sequence, we performed 2D real-time 25-key point body estimation on every image using OpenPose. Figure 7 illustrates the resulting skeleton sequence for a typical normal participant.

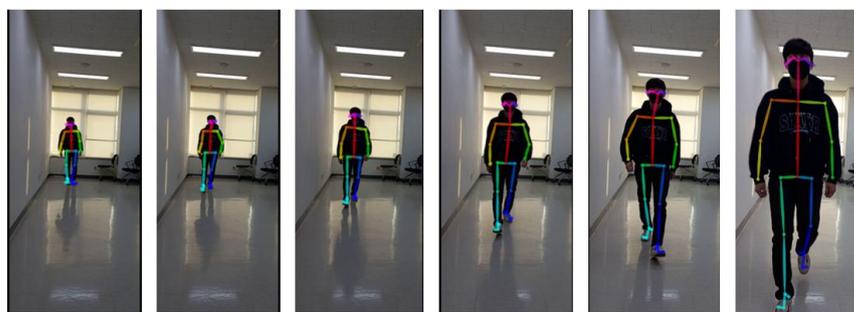


Figure 7. One normal skeleton sequence example.

3.3.2. Graph Structure Construction

To construct a spatiotemporal graph structure from a sequence comprising N nodes and T frames [27], we employed a pose graph $G = (V, E)$. The node set $V = \{v_t^i \mid t = 1, \dots, T, i = 1, \dots, N\}$ denotes the joint positions, where v_t^i represents the i -th joint at the t -th frame. The feature vector of v_t^i consists of the two-dimensional coordinate of this joint and the confidence score.

The edge set E includes: (a) the intra-skeleton connections, which connect the nodes of each frame according to the connections of human joints, where these edges form spatial edge; Figure 8a shows that we notate it as $E_s = \{v_t^i v_t^j \mid (i, j) \in H\}$, where H is a set of naturally connected human joints. (b) The inter-frame connections that connect the same joints (nodes) in two consecutive frames, where these edges form temporal edges. Figure 8b shows that we notate it as $\{v_t^i v_{t+1}^i\}$.

3.4. WM-STGCN

3.4.1. WM-STGCN Structure

Figure 9 shows the proposed WM-STGCN model architecture, which takes a sequence of human joint coordinates extracted from gait videos as input and predicts the gait category. Figure 9a provides an overall depiction of the proposed structure, whereas Figure 9b depicts the spatial module, and Figure 9c shows the temporal module.

The whole network comprises N GCN blocks ($N = 9$), with output channels of 64, 64, 64, 128, 128, 128, 256, 256, and 256, respectively. A global average pooling layer is added to the back end of the network, and the final output is sent to a SoftMax classifier to obtain the ultimate prediction result. To ensure training stability, residual connections are included in each basic block.

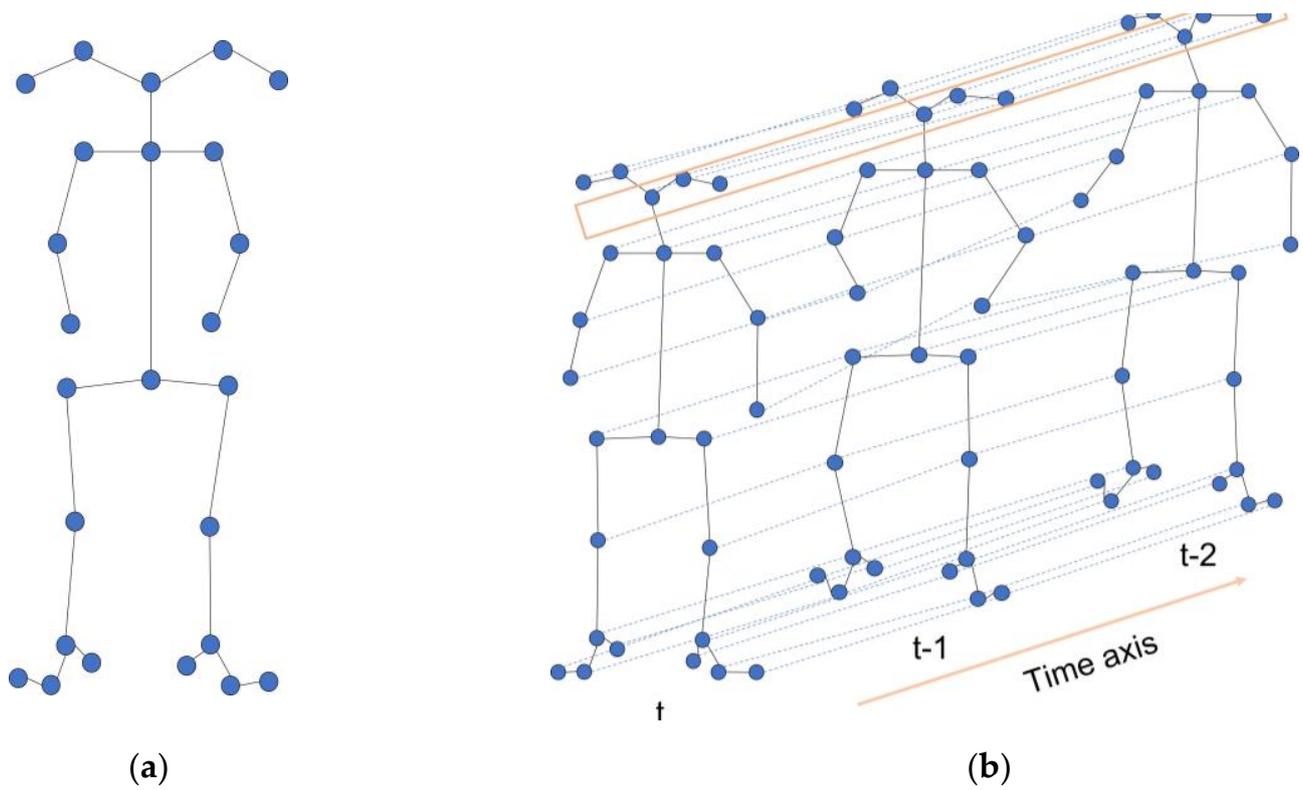


Figure 8. Spatiotemporal graph construction. (a) Spatial edges; (b) Temporal edges.

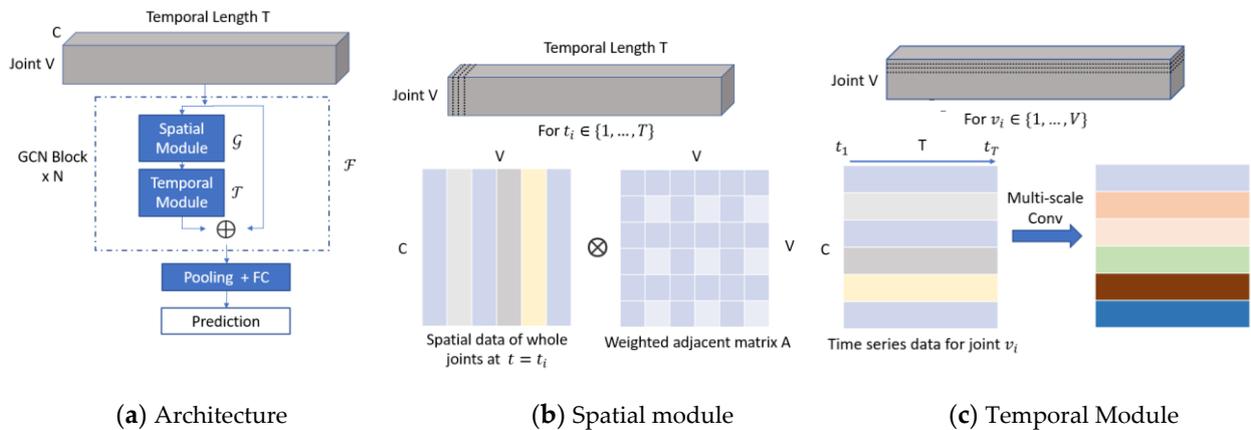


Figure 9. WM-STGCN framework. (a) The overall architecture of the proposed network; (b) The spatial module leverages the adjacency matrix to fuse features across joints; (c) The temporal module employs multi-scale temporal convolutions to capture temporal features.

Each GCN block \mathcal{F} comprises a spatial module \mathcal{G} and a temporal module \mathcal{T} . The spatial module \mathcal{G} combines the features of different joints using sparse matrices derived from the adjacency matrix A , as illustrated in Figure 10a. The output of \mathcal{G} is subsequently processed by \mathcal{T} to extract temporal features. The computations of \mathcal{F} can be summarized as follows:

$$\mathcal{F}(X) = \mathcal{T}(\mathcal{G}(X, A)) + X \tag{1}$$

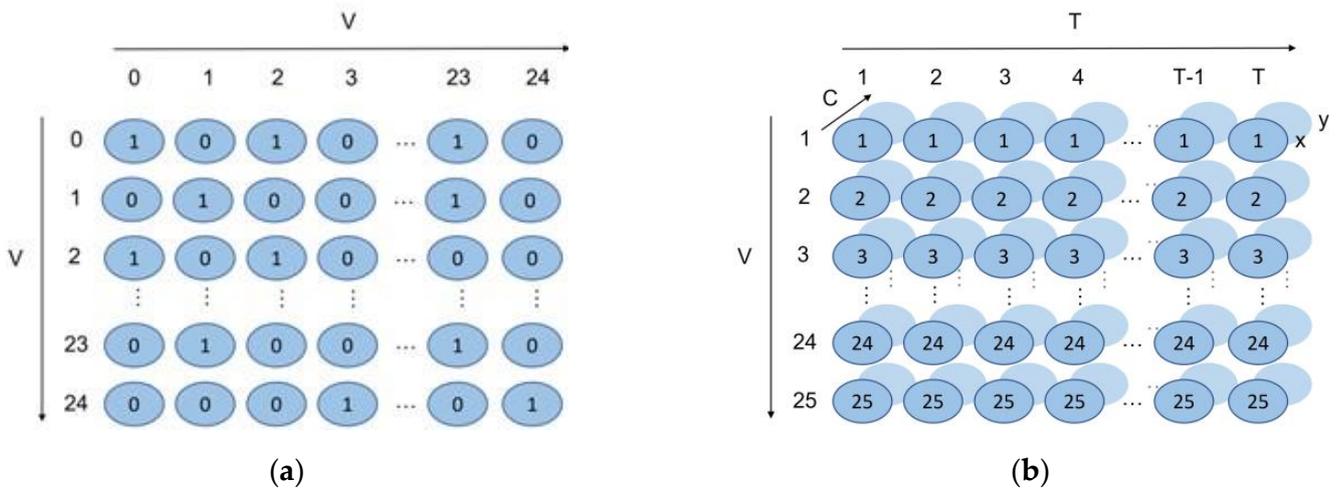


Figure 10. Input data. (a) Adjacency matrix A; (b) Input feature map of the first GCN block.

Figure 10b illustrates the input feature map of the first GCN block, wherein a skeleton feature $X \in R^{T \times V \times C}$ is given as input, where T denotes temporal length, V represents the number of skeleton joints, and C signifies the number of channels. Notably, the C input to the first GCN block equals 3.

3.4.2. Spatial Module \mathcal{G} : Graph Convolution in the Spatial Domain

In the spatial domain, the convolution of the graph on a certain node v_i is defined as follows:

$$f_{out}(v_i) = \sum_{v_j \in B_i} \frac{1}{Z_{ij}} f_{in}(v_j) \cdot \omega(l_i(v_j)) \tag{2}$$

where f_{in} and f_{out} represent the input and output feature maps, respectively; v_i represents a particular node in the spatial dimension; B_i represents the sampling area for the convolution of that node (in this work, B_i is the 1-neighbor set of v_i); Z is the normalizing term, which equals the cardinality of the corresponding subset; and w represents the weight function that provides the weight matrix.

We divided the neighborhood B into three subsets of self-connection, physical connection, and virtual connection, and different labels can be assigned to each subset. We discuss the virtual connection in Section 3.4.3. Here, l_i is a mapping function: $l_i(v_j) \rightarrow \{0, \dots, K, (K = 3)\}$, which maps a node in the neighborhood to its subset label.

Figure 11a shows a graph of the input skeleton sequence, and x_1 represents the root node itself (orange), x_2 represents the physically connected node (blue), and x_3 represents the virtually connected node (green). We use node 1 as the root node of this convolutional computation to explain the mapping strategy. Nodes 2, 4, 9 are its sampled neighboring nodes, which form the neighborhood B , where node 9 provides a virtual connection. Accordingly, as shown in Figure 11b, the adjacency matrix of node 3 is divided into three submatrices A_k , but ensure that where $A = \sum_k A_k, k = 1, 2, 3$.

Simplifying and transforming Equation (2), the spatial graph convolution can be implemented using the following:

$$F_{out} = \sum_k \omega_k(F_{in} A_k), k = 3 \tag{3}$$

$$A_k = \Lambda_k^{-\frac{1}{2}} \bar{A}_k \Lambda_k^{-\frac{1}{2}} \tag{4}$$

where, k in Equation (3) represents the amount of convolutional kernel, which is 3 according to the mapping strategy; A_k is an $N \times N$ normalized adjacency matrix; $\Lambda_k^{-\frac{1}{2}}$ is a normalized

diagonal matrix. ω is a 1×1 convolution operation, which represents the weight function in Equation (2).

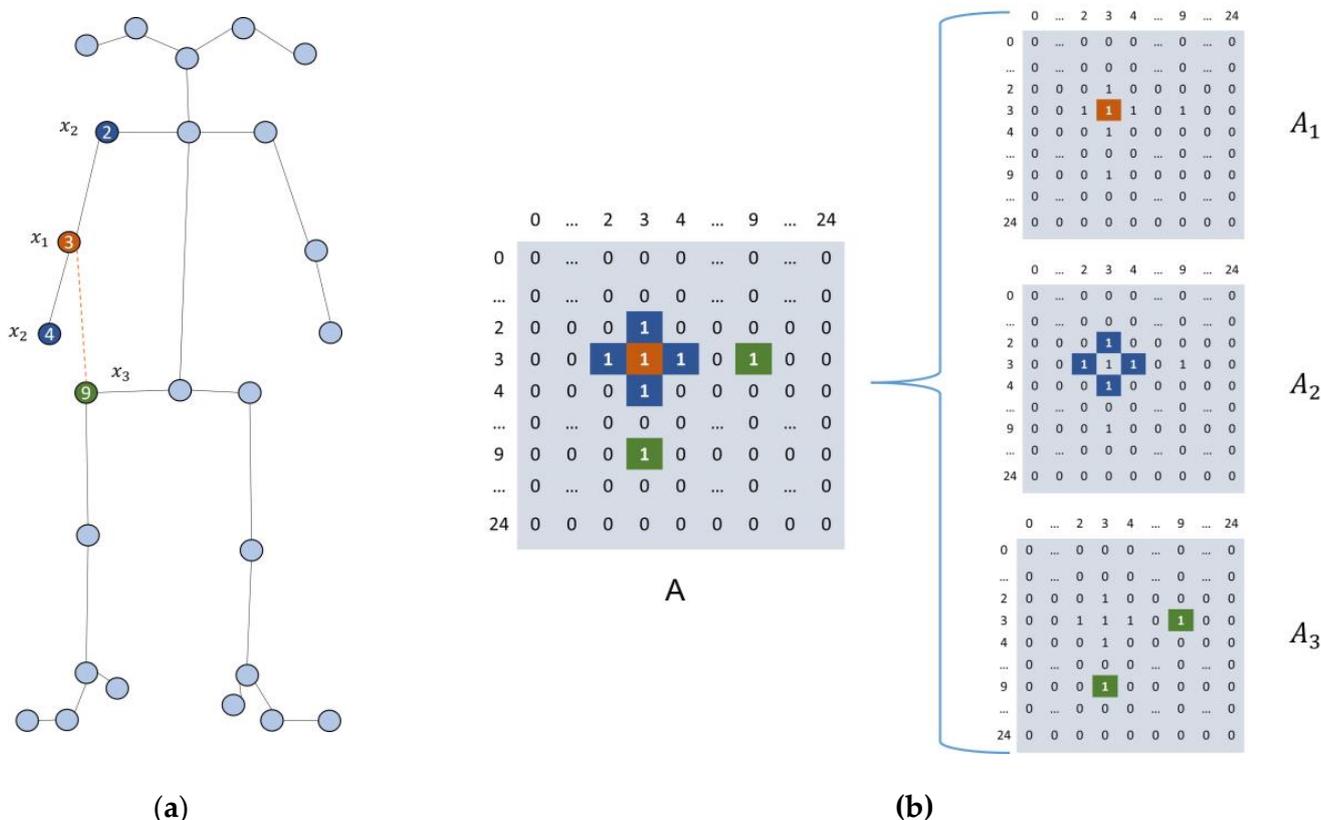


Figure 11. (a) A graph of the input skeleton sequence; (b) The three submatrices.

In the spatial domain, the input is represented as $\mathcal{G}_{in} \in R^{T \times V \times C_{in}}$; upon applying the spatial graph convolution, the resulting output feature map is denoted $\mathcal{G}_{out} \in R^{T \times V \times C_{out}}$.

3.4.3. Weighted Adjacency Matrix with Virtual Connection

The spatial structure of the skeleton is represented by an artificial, predefined adjacency matrix, which represents the a priori knowledge of the connections of the human skeleton. However, it cannot generate new connections between non-adjacent joints during training, which means that the learning ability of the graph convolutional network is limited and that such an adjacency matrix is not an optimal choice.

To address the above problems, we design a novel adjacency matrix, which has the following two features:

Virtual connection. We combined some unique features of Parkinson’s gait compared to normal gait (including small amplitude of arm swing, fast frequency and small stride length of foot movement, and random little steps) and introduced some virtual connections, i.e., unnaturally connected joints.

Weighted adjacent matrix. We used a scalar to multiply with the original adjacency matrix to get a new adjacency matrix, which makes distinct kinds of joints with different weights.

With these new designs, we make it possible to generate connections between non-adjacent joints, and give different weights for physical connections, virtual connections and self-connections. We design a new adjacency matrix and obtain a skeletal space structure that is more suitable for describing the Parkinson samples, thus enabling better gait recognition. Specifically, a_{ij} is a scalar:

$$a_{ij} = \begin{cases} \alpha, & \text{if } i = j \\ \beta, & \text{if joint } i \text{ and joint } j \text{ are connected physically} \\ \gamma, & \text{if joint } i \text{ and joint } j \text{ are connected virtually} \end{cases} \quad (5)$$

If we set the value of $a_{ii} = 0$, this indicates that we eliminate the self-connection of each joint. Additionally, we distinguish between physical and virtual dependencies between joints. The physical dependency, represented by β and depicted as blue solid lines in Figure 12a, captures the intrinsic connection between joints. The virtual dependency, depicted as orange dashed lines in Figure 12a, represents the extrinsic connection between two joints, which is also crucial for gait recognition. We use the parameter γ to model this virtual relationship. For example, although the left hip and left hand are physically disconnected, their relationship is essential in identifying Parkinsonian gait.

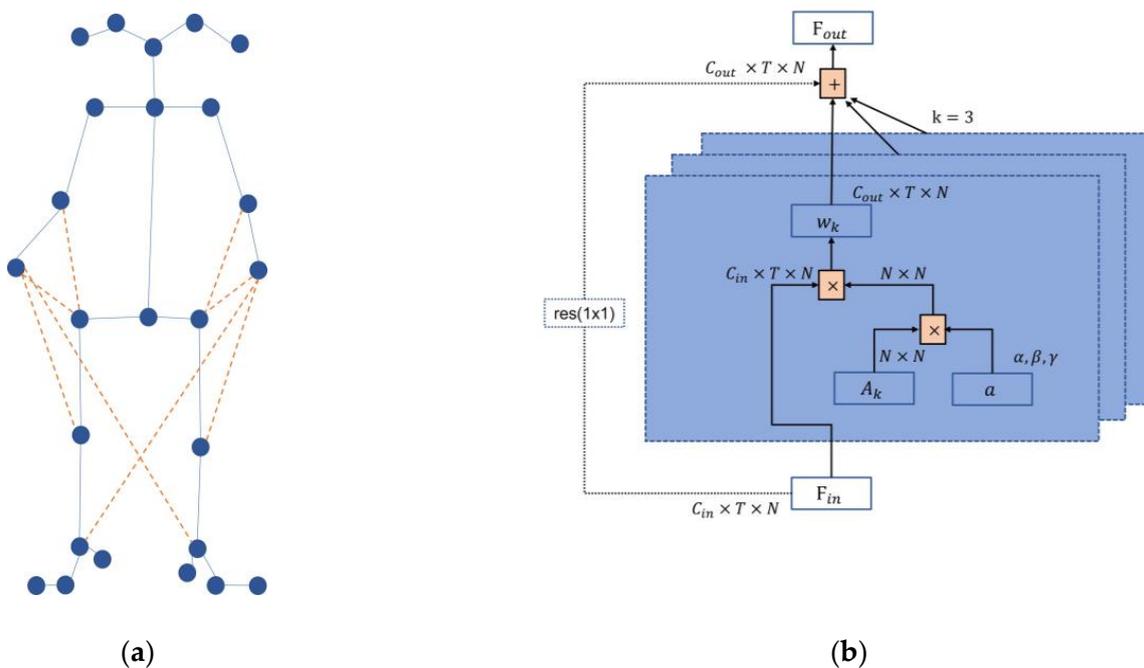


Figure 12. (a) Virtual connection; (b) Diagram of the graph convolutional layer with weights.

After adding weights, the graph convolution formula in spatial dimension can be transformed from Equation (3) to the following:

$$F_{out} = \sum_k \omega_k F_{in}(A_k \cdot a) \quad (6)$$

Figure 12b shows the process of weight addition, where the adjacency matrix of each layer consists of A_k and weight a together, k denotes the number of subsets, and the dashed line indicates that the residual convolution operation is required only when C_{in} is different from C_{out} .

For the experiment, we tested 4 cases: ① $\alpha = 1, \beta = 1, \gamma = 0$; ② $\alpha = 1, \beta = 1, \gamma = 0.5$; ③ $\alpha = 0, \beta = 1, \gamma = 0.5$; ④ $\alpha = 0.2, \beta = 1, \gamma = 0.5$. This means that we tested the performance of the model with self-connection, 0.5 weight virtual connection, without self-connection and 0.2 weight self-connection and 0.5 weight virtual connection.

Figure 13 shows the corresponding weighted adjacency matrixes. The red box marks the representation of the virtual connection in the matrix.

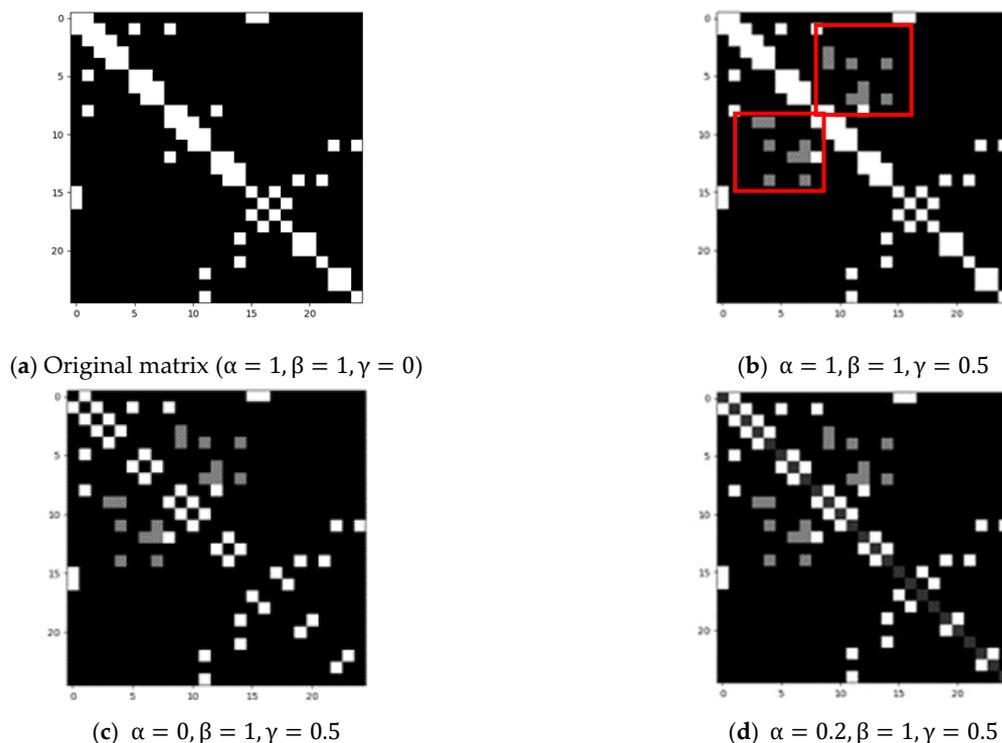


Figure 13. Different parameters for weighted matrix.

3.4.4. Temporal Module \mathcal{T} : Graph Convolution in Temporal Domain

\mathcal{G} captures the spatial dependencies between adjacent joints, and to model the temporal changes of these features, we employed a multi-scale temporal convolution network (MS-TCN). Unlike many existing works that employ temporal convolution networks with fixed kernel sizes $k_t \times 1$ throughout the architecture, we designed a MS-TCN, as shown in Figure 14, to promote the flexibility and temporal modeling capability by using multi-group convolution.

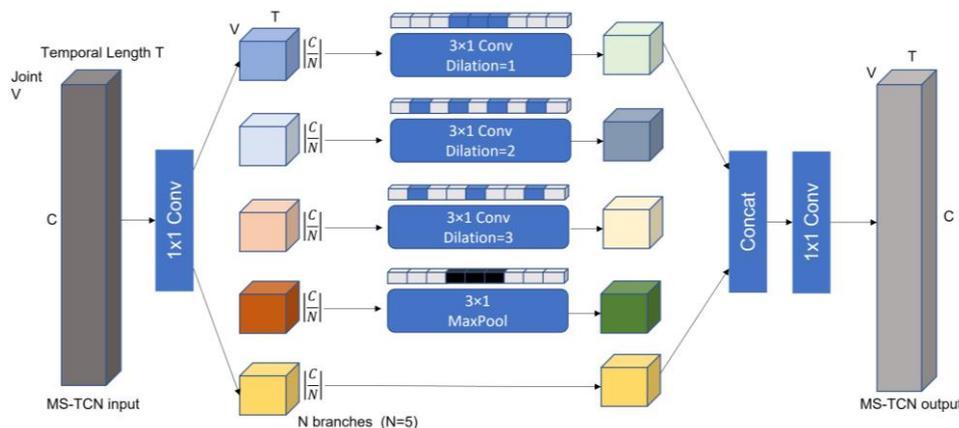


Figure 14. Multi-scale temporal convolution network.

The adopted multi-scale TCN contains five branches: a 1×1 convolution branch, a Max-pooling branch, and three temporal convolutions with kernel size 5 and dilations from (1 to 3). Every branch contains a 1×1 convolution, which is used to reduce channel dimension before the expensive convolution 3×1 . Additionally, the 1×1 convolution

introduces additional nonlinearity via a nonlinear activation function, thereby increasing the network's complexity, and enabling it to be deeper. This output continues to be fed into the spatial graph convolution, as shown in Figure 9, and it is fed into the fully connected layer only in the last GCN block.

The MS-TCN enhances vanilla temporal convolution layer's receptive fields, and improves the temporal aggregation capability. At the same time, it reduces computational cost and parameters through reduced channel width for each branch.

4. Experiments

4.1. Implementation Details

We used NVIDIA GeForce RTX 2080Ti GPU with 12 GB memory, Intel(R) Core(TM) i9-10900 CPU with 2.80 GHz 64 GB RAM to build the deep learning framework using PyTorch in Windows 10 environment. We used CUDA, Cudnn, OpenCV, and other required libraries to train and test the Parkinsonian gait recognition model. The batch size during training and testing was 16. The base learning rate was 0.1. We chose SGD as optimizer with step [20,30,40,50]. Following data preprocessing, we obtained 160 normal samples, and 150 Parkinsonian samples. We split our dataset into a training set and a test set, with a ratio of 80% and 20%, respectively. The test set comprised 32 normal samples and 30 Parkinsonian samples.

4.2. Evaluation Metric

In this study, we defined Parkinsonian gait samples as positive and normal gait samples as negative. We utilized widely accepted evaluation metrics, including True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN), to accurately classify samples into these categories. To evaluate the performance of our method, we selected accuracy, precision, sensitivity, specificity, false alarm, miss rate, and F1 score as our evaluation metrics. A higher value for accuracy, precision, sensitivity, specificity, and F1 score indicates better model performance. In contrast, a smaller value for false alarm and miss rate indicates better performance.

Accuracy reflects the ability of the model to correctly judge the overall sample, i.e., the ability to correctly classify Parkinsonian samples as positive, and normal samples as negative.

Precision reflects the ability of the model to correctly predict the positive samples, i.e., how many of the predicted Parkinsonian samples are true Parkinsonian samples.

Sensitivity is defined as the proportion of Parkinsonian samples predicted to be Parkinsonian samples to the total number of Parkinsonian samples. Specificity reflects the proportion of normal samples that are predicted as normal samples to the total normal samples.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (7)$$

$$precision = \frac{TP}{TP + FP} \quad (8)$$

$$sensitivity = TPR = \frac{TP}{TP + FN} \quad (9)$$

$$specificity = TNR = \frac{TN}{FP + TN} \quad (10)$$

False alarm, also known as false positive rate or false detection rate, is obtained by calculating the proportion of normal samples predicted as Parkinsonian samples to the total normal samples. Miss rate is obtained by calculating the proportion of Parkinsonian samples that are predicted as normal samples to the total Parkinsonian samples.

$$false\ alarm = FPR = \frac{FP}{FP + TN} \quad (11)$$

$$\text{miss rate} = \text{FNR} = \frac{FN}{TP + FN} \quad (12)$$

Furthermore, *F1* score is widely used in model evaluation. This is the harmonic mean of the precision and recall, which can reflect the performance of the model in a balanced way.

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (13)$$

4.3. Results and Discussion

We experimented with different parameters of Gaussian noise augmentation with $\mu = 0$, and $\sigma = (0.01, 0.05, 0.1)$. In Table 2 and Figure 15, the experimental results show that the model had the highest accuracy of 85.48% for $\sigma = 0.1$. Although the precision was 4.87% lower compared to the group with $\sigma = 0.01$, the sensitivity increased from 60% to 80%, improving the performance of predicting Parkinsonian samples as positive, which was the best of the three experimental groups. Meanwhile, the miss rate was only 20%, which was much lower than the 40% at $\sigma = 0.05$. Overall, the model showed the best performance for detecting Parkinson's samples at $\sigma = 0.1$. Figure 16 shows the accuracy during training based on several groups of Gaussian noise.

Table 2. Different parameter results.

Group	Accuracy	Precision	Sensitivity	Specificity	False Alarm/FPR	Miss Rate /FNR
Gaussian noise ($\mu = 0, \sigma = 0.01$)	74.19%	93.75%	50.0%	96.87%	3.12%	50.0%
Gaussian noise ($\mu = 0, \sigma = 0.05$)	75.81%	85.71%	60.0%	90.62%	9.37%	40.0%
Gaussian noise ($\mu = 0, \sigma = 0.1$)	85.48%	88.88%	80.0%	90.62%	9.37%	20.0%

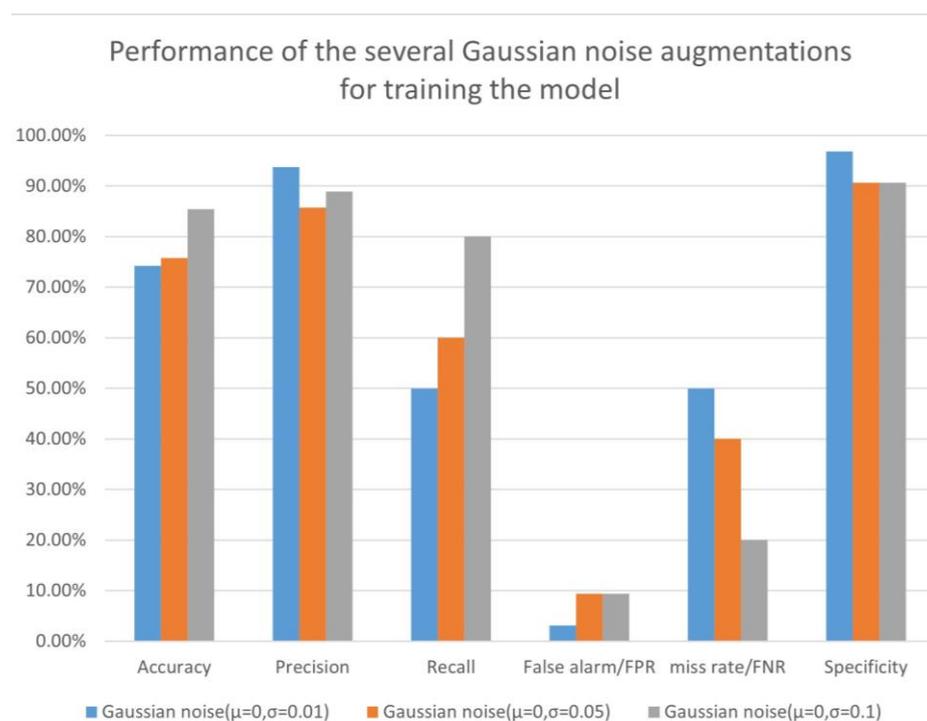


Figure 15. Performance of the several Gaussian noise augmentations.

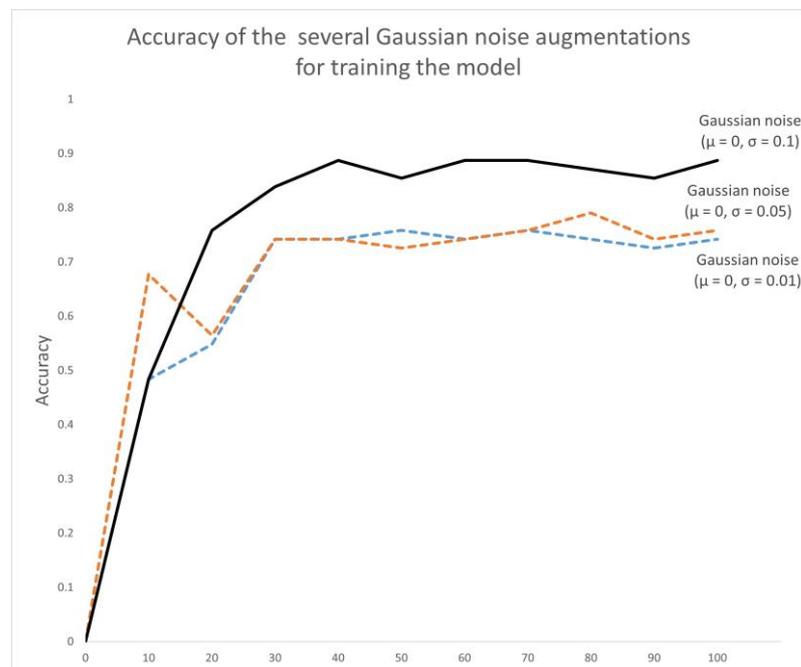


Figure 16. Accuracy of the several Gaussian noise augmentations.

For the different weight adjacencies, we tested four cases. When $\alpha = 1$, $\beta = 1$, $\gamma = 0$, which is the original matrix containing only self-connections and physical connections. In Table 3, the experimental results showed that the accuracy reached 72.58%, and the recognition miss rate of Parkinson's gait was 46.67%, the lowest among the four groups. When adding $\gamma = 0.5$, i.e., 0.5 weight of virtual connections, we found that although the accuracy rate decreased slightly from 72.58% to 70.97%, the sensitivity and miss rate increased.

Table 3. Results of different weight parameters.

Weight Parameters	Accuracy	Precision	Sensitivity	Specificity	False Alarm/FPR	Miss Rate /FNR
Original ($\alpha = 1$, $\beta = 1$, $\gamma = 0$)	72.58%	84.21%	53.33%	90.62%	9.38%	46.67%
$\alpha = 1$, $\beta = 1$, $\gamma = 0.5$	70.97%	77.27%	56.67%	84.38%	15.63%	43.33%
$\alpha = 0$, $\beta = 1$, $\gamma = 0.5$	85.48%	88.88%	80.0%	90.62%	9.38%	20.0%
$\alpha = 0.2$, $\beta = 1$, $\gamma = 0.5$	87.10%	86.67%	86.67%	87.50%	12.50%	13.33%

After removing the self-connection, we found that the accuracy increased by 14.51% and sensitivity increased by 23.33%, while the miss rate decreased from 43.33% to 20%. This indicates that removing the effect of joint self-connection aids the correct recognition of gait.

Finally, we achieved the best results with 0.2 weight of the joint self-connections and 0.5 weight of the virtual joint, where the accuracy was 87.10%, the sensitivity was 86.67%, and the miss rate was the smallest, at 13.33%. Figures 17a and 17b show the confusion matrix and loss function, respectively.

Through our experiments, our best result showed an accuracy of 87.10%. Table 4 compares the performance with the other well-known machine learning models of LSTM, KNN, Decision Tree, AdaBoost, and ST-GCN. In particular, Lstm-layer1 means a one-layer network, layer2 means a two-layer network, and the weak learner model in the AdaBoost classifier is 50 decision trees of depth 1.

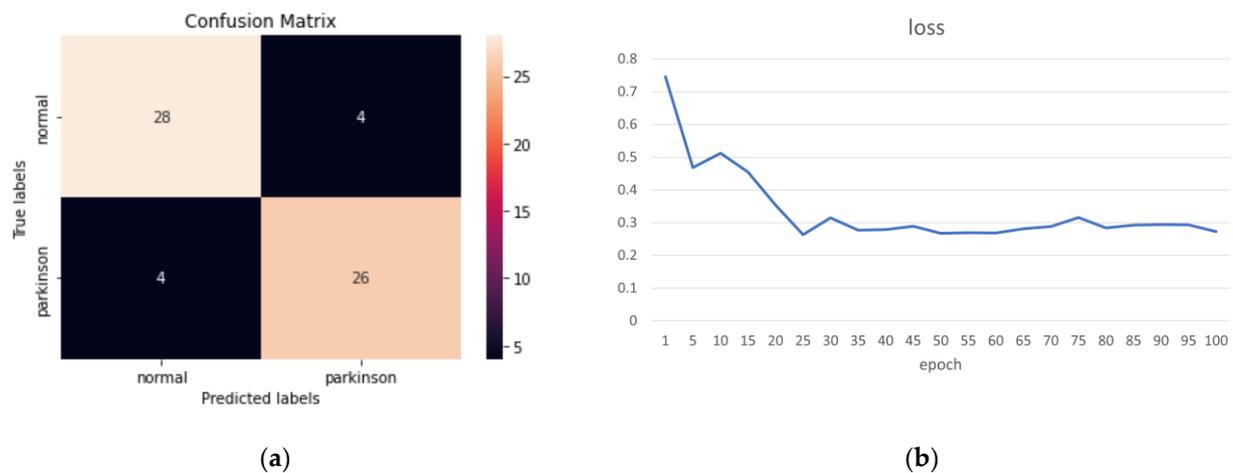


Figure 17. (a) Confusion matrix; (b) Loss function.

Table 4. Comparison with other models.

Methods	Accuracy	Precision	Sensitivity	F1 Score	Specificity	False Alarm /FPR	Miss Rate /FNR
Lstm-layer1	82.25%	85.19%	76.67%	0.8679	87.5%	12.5%	23.33%
Lstm-layer2	69.35%	100%	63.33%	0.7755	100%	0%	36.67%
KNN	83.87%	85.71%	80%	0.8276	87.5%	12.5%	20%
Decision tree	79.03%	81.48%	73.33%	0.8461	84.38%	15.63%	26.67%
AdaBoost	75.81%	85.71%	60%	0.7059	90.63%	9.38%	40%
ST-GCN	77.42%	90%	56.25%	0.72	93.75%	6.25%	40%
Proposed method	87.10%	86.67%	86.67%	0.9285	87.5%	12.5%	13.33%

We conducted an analysis to investigate the superior performance of WM-STGCN in comparison to other models based on the following factors. The first factor is the utilization of a weighted adjacent matrix with virtual connections. The weighted adjacency matrix with virtual connections plays a crucial role in WM-STGCN. While an adjacency matrix without weights can be used to represent adjacency information, a weighted adjacency matrix allows for a more sophisticated representation of adjacency information. Moreover, weights can reflect the structure of the graph in a more granular way, for example, by adjusting weights based on the connection types to emphasize relationships with physical connections or virtual connections. Therefore, using a weighted adjacency matrix enables WM-STGCN models to reflect more detailed graph structures and make better predictions. The second factor is the integration of a multi-scale temporal convolutional network. The multi-scale temporal convolutional network used in this study can enhance the receptive field of temporal convolution, improve time aggregation ability, and extract features from various time intervals. At the same time, it can reduce the computational cost and parameters by reducing the channel width of each branch. Finally, we use a separately designed data augmentation method for both raw video and skeletal data, which also effectively improves the performance of the model.

These advantages enable effective recognition of Parkinson's disease from gait data. However, there are also some shortcomings. For example, due to equipment limitations, we focused on the RGB color video of the front view, but users cannot guarantee to record high-quality video when using it, which will affect the recognition accuracy. At the same time, our model performance can be further improved by using multi-modal analysis methods, such as adding sensor data. In the future, our WM-STGCN model is expected to be applied to research on gait-related diseases in the elderly, including not only Parkinson's disease but also dementia, stroke, and other related conditions.

5. Conclusions

In this paper, we proposed a novel spatiotemporal modeling approach, known as WM-STGCN, which employs a weighted adjacent matrix with virtual connections and multi-scale temporal convolutional networks to recognize Parkinsonian gait from forward walking videos. Our experimental results demonstrated the effectiveness of the proposed method, which outperformed the machine learning-based methods such as LSTM, KNN, Decision Tree, AdaBoost, and ST-GCN. This method could provide a promising solution for PD gait recognition, which is crucial for the early and accurate diagnosis of PD. We believe that our method can be further improved by integrating it with other advanced deep learning techniques and can be extended to the fields of healthcare and biomedicine.

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

Funding: This work was supported by the Institute of Information & Communications Technology Planning & Evaluation (IITP) grant funded by the Korean government (MSIT) (No. 2020-0-00990, Platform Development and Proof of High Trust & Low Latency Processing for Heterogeneous-Atypical-Large Scaled Data in 5G-IoT Environment).

Institutional Review Board Statement: This study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board of Seoul Boramae Medical Center (Approval number: 20-2021-4. Approval Date 16 March 2022).

Informed Consent Statement: Normal person informed consent was obtained from all subjects involved in the study. Patient informed consent was waived due to the reason that the data were obtained from online public records.

Data Availability Statement: The data presented in this study are available on reasonable request from the corresponding author.

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

References

1. Mc Ardle, R.; Galna, B.; Donaghy, P.; Thomas, A.; Rochester, L. Do Alzheimer's and Lewy Body Disease Have Discrete Pathological Signatures of Gait? *Alzheimer's Dement.* **2019**, *15*, 1367–1377. [[CrossRef](#)]
2. Beauchet, O.; Blumen, H.M.; Callisaya, M.L.; De Cock, A.M.; Kressig, R.W.; Srikanth, V.; Steinmetz, J.P.; Verghese, J.; Allali, G. Spatiotemporal gait characteristics associated with cognitive impairment: A multicenter cross-sectional study, the intercontinental. *Curr. Alzheimer Res.* **2018**, *15*, 273–282. [[CrossRef](#)] [[PubMed](#)]
3. Mirelman, A.; Bonato, P.; Camicioli, R.; Ellis, T.D.; Giladi, N.; Hamilton, J.L.; Hass, C.J.; Hausdorff, J.M.; Pelosin, E.; Almeida, Q.J. Gait impairments in Parkinson's disease. *Lancet Neurol.* **2019**, *18*, 697–708. [[CrossRef](#)]
4. Goetz, C.G.; Tilley, B.C.; Shaftman, S.R.; Stebbins, G.T.; Fahn, S.; Martinez-Martin, P.; Poewe, W.; Sampaio, C.; Stern, M.B.; Dodel, R.; et al. Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale Presentation and Clinimetric Testing Results. *Mov. Disord.* **2008**, *23*, 2129–2170. [[CrossRef](#)] [[PubMed](#)]
5. Simpson, G.M.; Angus, J.W.S. A Rating Scale for Extrapyrarnidal Side Effects. *Acta Psychiatr. Scand.* **1970**, *45*, 11–19. [[CrossRef](#)] [[PubMed](#)]
6. Abdul Razak, A.H.; Zayegh, A.; Begg, R.K.; Wahab, Y. Foot Plantar Pressure Measurement System: A Review. *Sensors* **2012**, *12*, 9884–9912. [[CrossRef](#)]
7. Shull, P.B.; Jirattigalachote, W.; Hunt, M.A.; Cutkosky, M.R.; Delp, S.L. Quantified Self and Human Movement: A Review on the Clinical Impact of Wearable Sensing and Feedback for Gait Analysis and Intervention. *Gait Posture* **2014**, *40*, 11–19. [[CrossRef](#)]
8. Stone, E.E.; Skubic, M. Passive in-home measurement of stride-to-stride gait variability comparing vision and Kinect sensing. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 6491–6494. [[CrossRef](#)]
9. Rocha, A.P.; Choupina, H.; Fernandes, J.M.; Rosas, M.J.; Vaz, R.; Cunha, J.P.S. Parkinson's disease assessment based on gait analysis using an innovative RGB-D camera system. In Proceedings of the 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, IL, USA, 27–31 August 2014; pp. 3126–3129. [[CrossRef](#)]

10. Pfister, A.; West, A.M.; Bronner, S.; Noah, J.A. Comparative Abilities of Microsoft Kinect and Vicon 3D Motion Capture for Gait Analysis. *J. Med. Eng. Technol.* **2014**, *38*, 274–280. [CrossRef]
11. Geerse, D.J.; Roerdink, M.; Marinus, J.; van Hilten, J.J. Assessing Walking Adaptability in Parkinson’s Disease: “The Interactive Walkway”. *Front. Neurol.* **2018**, *9*, 1096. [CrossRef]
12. Dranca, L.; de Abetxuko Ruiz de Mendarozketa, L.; Goñi, A.; Illarramendi, A.; Navalpotro Gomez, I.; Delgado Alvarado, M.; Rodríguez-Oroz, M.C. Using Kinect to Classify Parkinson’s Disease Stages Related to Severity of Gait Impairment. *BMC Bioinform.* **2018**, *19*, 471. [CrossRef]
13. Kim, H.N. Ambient intelligence: Placement of Kinect sensors in the home of older adults with visual disabilities. *Technol. Disabil.* **2020**, *32*, 271–283. [CrossRef]
14. Müller, B.; Ilg, W.; Giese, M.A.; Ludolph, N. Validation of enhanced kinect sensor based motion capturing for gait assessment. *PLoS ONE* **2017**, *12*, e0175813. [CrossRef]
15. Cao, Z.; Hidalgo, G.; Simon, T.; Wei, S.-E.; Sheikh, Y. OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2021**, *43*, 172–186. [CrossRef]
16. Girshick, R.; Radosavovic, I.; Gkioxari, G.P.; Doll, A.R.; He, K. Detectron. 2018. Available online: <https://github.com/facebookresearch/detectron> (accessed on 3 April 2023).
17. Fang, H.-S.; Li, J.; Tang, H.; Xu, C.; Zhu, H.; Xiu, Y.; Li, Y.-L.; Lu, C. AlphaPose: Whole-Body Regional Multi-Person Pose Estimation and Tracking in Real-Time. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 7157–7173. [CrossRef] [PubMed]
18. Sato, K.; Nagashima, Y.; Mano, T.; Iwata, A.; Toda, T. Quantifying Normal and Parkinsonian Gait Features from Home Movies: Practical Application of a Deep Learning–Based 2D Pose Estimator. *PLoS ONE* **2019**, *14*, e0223549. [CrossRef] [PubMed]
19. Sabo, A.; Mehdizadeh, S.; Iaboni, A.; Taati, B. Estimating Parkinsonism Severity in Natural Gait Videos of Older Adults With Dementia. *IEEE J. Biomed. Health Inform.* **2022**, *26*, 2288–2298. [CrossRef]
20. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Vision-Based Assessment of Parkinsonism and Levodopa-Induced Dyskinesia with Pose Estimation. *J. NeuroEng. Rehabil.* **2018**, *15*, 97. [CrossRef]
21. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Automated Assessment of Levodopa-Induced Dyskinesia: Evaluating the Responsiveness of Video-Based Features. *Park. Relat. Disord.* **2018**, *53*, 42–45. [CrossRef]
22. Lu, M.; Poston, K.; Pfefferbaum, A.; Sullivan, E.V.; Fei-Fei, L.; Pohl, K.M.; Niebles, J.C.; Adeli, E. Vision-Based Estimation of MDS–UPDRS Gait Scores for Assessing Parkinson’s Disease Motor Severity. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference 2020, Lima, Peru, 4–8 October 2020; Martel, A.L., Abolmaesumi, P., Stoyanov, D., Mateus, D., Zuluaga, M.A., Zhou, S.K., Racocanu, D., Joskowicz, L., Eds.; Lecture Notes in Computer Science. Springer International Publishing: Cham, Switzerland, 2020; pp. 637–647. [CrossRef]
23. Chen, R.J.; Lu, M.Y.; Chen, T.Y.; Williamson, D.F.K.; Mahmood, F. Synthetic Data in Machine Learning for Medicine and Healthcare. *Nat. Biomed. Eng.* **2021**, *5*, 493–497. [CrossRef]
24. Emam, K.E.; Mosquera, L.; Hoptroff, R. *Practical Synthetic Data Generation: Balancing Privacy and the Broad Availability of Data*; O’Reilly Media: Sebastopol, CA, USA, 2020.
25. Nikolenko, S.I. Synthetic Data for Deep Learning. *arXiv* **2019**. [CrossRef]
26. Rankin, D.; Black, M.; Bond, R.; Wallace, J.; Mulvenna, M.; Epelde, G. Reliability of Supervised Machine Learning Using Synthetic Data in Health Care: Model to Preserve Privacy for Data Sharing. *JMIR Med. Inform.* **2020**, *8*, e18910. [CrossRef]
27. Yan, S.; Xiong, Y.; Lin, D. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, New Orleans, LA, USA, 2–7 February 2018; Volume 32. [CrossRef]
28. Guo, R.; Shao, X.; Zhang, C.; Qian, X. Sparse Adaptive Graph Convolutional Network for Leg Agility Assessment in Parkinson’s Disease. *IEEE Trans. Neural Syst. Rehabil. Eng.* **2020**, *28*, 2837–2848. [CrossRef]
29. Lou, C.; Wang, S.; Liang, T.; Pang, C.; Huang, L.; Run, M.; Liu, X. A Graphene-Based Flexible Pressure Sensor with Applications to Plantar Pressure Measurement and Gait Analysis. *Materials* **2017**, *10*, 1068. [CrossRef]
30. Camps, J.; Samà, A.; Martín, M.; Rodríguez-Martín, D.; Pérez-López, C.; Moreno Arostegui, J.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]
31. Seifert, A.-K.; Amin, M.G.; Zoubir, A.M. Toward Unobtrusive In-Home Gait Analysis Based on Radar Micro-Doppler Signatures. *IEEE Trans. Biomed. Eng.* **2019**, *66*, 2629–2640. [CrossRef]
32. Prakash, C.; Gupta, K.; Mittal, A.; Kumar, R.; Laxmi, V. Passive Marker Based Optical System for Gait Kinematics for Lower Extremity. *Procedia Comput. Sci.* **2015**, *45*, 176–185. [CrossRef]
33. Seifollahi, M.; Soltanzadeh, H.; Hassani Mehraban, A.; Khamseh, F. Alzheimer’s Disease Detection Using Skeleton Data Recorded with Kinect Camera. *Clust. Comput.* **2020**, *23*, 1469–1481. [CrossRef]
34. Nguyen, T.-N.; Huynh, H.-H.; Meunier, J. Estimating Skeleton-Based Gait Abnormality Index by Sparse Deep Auto-Encoder. In Proceedings of the 2018 IEEE Seventh International Conference on Communications and Electronics (ICCE), Hue, Vietnam, 18–20 July 2018; pp. 311–315. [CrossRef]
35. Jun, K.; Lee, D.-W.; Lee, K.; Lee, S.; Kim, M.S. Feature Extraction Using an RNN Autoencoder for Skeleton-Based Abnormal Gait Recognition. *IEEE Access* **2020**, *8*, 19196–19207. [CrossRef]
36. Shalin, G.; Pardoel, S.; Lemaire, E.D.; Nantel, J.; Kofman, J. Prediction and Detection of Freezing of Gait in Parkinson’s Disease from Plantar Pressure Data Using Long Short-Term Memory Neural-Networks. *J. NeuroEng. Rehabil.* **2021**, *18*, 167. [CrossRef]

37. Zhang, S.; Yang, Y.; Xiao, J.; Liu, X.; Yang, Y.; Xie, D.; Zhuang, Y. Fusing Geometric Features for Skeleton-Based Action Recognition Using Multilayer LSTM Networks. *IEEE Trans. Multimed.* **2018**, *20*, 2330–2343. [[CrossRef](#)]
38. Zhu, K.; Wang, R.; Zhao, Q.; Cheng, J.; Tao, D. A Cuboid CNN Model With an Attention Mechanism for Skeleton-Based Action Recognition. *IEEE Trans. Multimed.* **2020**, *22*, 2977–2989. [[CrossRef](#)]
39. Li, C.; Zhong, Q.; Xie, D.; Pu, S. Skeleton-Based Action Recognition with Convolutional Neural Networks. In Proceedings of the 2017 IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Hong Kong, China, 10–14 July 2017; pp. 597–600. [[CrossRef](#)]
40. Wen, Y.-H.; Gao, L.; Fu, H.; Zhang, F.-L.; Xia, S. Graph CNNs with Motif and Variable Temporal Block for Skeleton-Based Action Recognition. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; Volume 33, pp. 8989–8996. [[CrossRef](#)]
41. Shi, L.; Zhang, Y.; Cheng, J.; Lu, H. Two-Stream Adaptive Graph Convolutional Networks for Skeleton-Based Action Recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 16–17 June 2019; pp. 12018–12027. [[CrossRef](#)]
42. Singh, J.P.; Jain, S.; Arora, S.; Singh, U.P. Vision-Based Gait Recognition: A Survey. *IEEE Access* **2018**, *6*, 70497–70527. [[CrossRef](#)]
43. Li, S.; Liu, W.; Ma, H. Attentive Spatial–Temporal Summary Networks for Feature Learning in Irregular Gait Recognition. *IEEE Trans. Multimed.* **2019**, *21*, 2361–2375. [[CrossRef](#)]
44. Ye, M.; Yang, C.; Stankovic, V.; Stankovic, L.; Cheng, S. Distinct Feature Extraction for Video-Based Gait Phase Classification. *IEEE Trans. Multimed.* **2020**, *22*, 1113–1125. [[CrossRef](#)]
45. Li, M.H.; Mestre, T.A.; Fox, S.H.; Taati, B. Automated Vision-Based Analysis of Levodopa-Induced Dyskinesia with Deep Learning. In Proceedings of the 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Republic of Korea, 11–15 July 2017; pp. 3377–3380. [[CrossRef](#)]
46. Hu, K.; Wang, Z.; Mei, S.; Ehgoetz Martens, K.A.; Yao, T.; Lewis, S.J.G.; Feng, D.D. Vision-Based Freezing of Gait Detection with Anatomic Directed Graph Representation. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 1215–1225. [[CrossRef](#)] [[PubMed](#)]
47. Neurology-Topic 13-Parkinson’s Disease Female Patient. Available online: <https://www.youtube.com/watch?v=kXMydIXQYpY> (accessed on 6 March 2023).
48. Gait Impairments in Parkinson’s Disease. Available online: <https://www.youtube.com/watch?v=pFLC9C-xH8E> (accessed on 6 March 2023).
49. Freezing of Gait. Available online: <https://www.youtube.com/watch?v=3-wrNhyVTNE> (accessed on 6 March 2023).
50. Moderate and Severe Parkinsonian Gait. Available online: <https://www.youtube.com/watch?v=t1IkEakBSz4> (accessed on 6 March 2023).
51. Parkinson’s Disease Gait—Moderate Severity. Available online: <https://www.youtube.com/watch?v=pu5Vwf1CBO0> (accessed on 6 March 2023).
52. A 66-Year-Old Man with Parkinson’s Disease Taught to Improve Walking Gait and Running Gait. Available online: <https://www.youtube.com/watch?v=JUMhhwFANKE> (accessed on 6 March 2023).

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.