

Article

Identifying the Posture of Young Adults in Walking Videos by Using a Fusion Artificial Intelligent Method

Posen Lee ¹, Tai-Been Chen ^{2,3}, Chin-Hsuan Liu ^{1,4,*}, Chi-Yuan Wang ², Guan-Hua Huang ³
and Nan-Han Lu ^{2,5,6}

- ¹ Department of Occupation Therapy, I-Shou University, No. 8, Yida Road, Jiaosu Village, Yanchao District, Kaohsiung 82445, Taiwan; posenlee@isu.edu.tw
- ² Department of Medical Imaging and Radiological Science, I-Shou University, No. 8, Yida Road, Jiaosu Village Yanchao District, Kaohsiung 82445, Taiwan; ctb@isu.edu.tw (T.-B.C.); wang1b011@isu.edu.tw (C.-Y.W.); ed103911@edah.org.tw (N.-H.L.)
- ³ Institute of Statistics, National Yang Ming Chiao Tung University, No. 1001, University Road, Hsinchu 30010, Taiwan; ghuang@stat.nctu.edu.tw
- ⁴ Department of Occupational Therapy, Kaohsiung Municipal Kai-Syuan Psychiatric Hospital, No. 130, Kaisyuan 2nd Road, Lingya District, Kaohsiung City 80276, Taiwan
- ⁵ Department of Pharmacy, Tajen University, No. 20, Weixin Road, Yanpu Township, Pingtung County 90741, Taiwan
- ⁶ Department of Radiology, E-DA Hospital, I-Shou University, No. 1, Yida Road, Jiaosu Village, Yanchao District, Kaohsiung City 82445, Taiwan
- * Correspondence: isu6a394@cloud.isu.edu.tw; Tel.: +886-7-6151100 (ext. 7516)

Abstract: Many neurological and musculoskeletal disorders are associated with problems related to postural movement. Noninvasive tracking devices are used to record, analyze, measure, and detect the postural control of the body, which may indicate health problems in real time. A total of 35 young adults without any health problems were recruited for this study to participate in a walking experiment. An iso-block postural identity method was used to quantitatively analyze posture control and walking behavior. The participants who exhibited straightforward walking and skewed walking were defined as the control and experimental groups, respectively. Fusion deep learning was applied to generate dynamic joint node plots by using OpenPose-based methods, and skewness was qualitatively analyzed using convolutional neural networks. The maximum specificity and sensitivity achieved using a combination of ResNet101 and the naïve Bayes classifier were 0.84 and 0.87, respectively. The proposed approach successfully combines cell phone camera recordings, cloud storage, and fusion deep learning for posture estimation and classification.

Keywords: iso-block postural identity; OpenPose; fusion deep learning



Citation: Lee, P.; Chen, T.-B.; Liu, C.-H.; Wang, C.-Y.; Huang, G.-H.; Lu, N.-H. Identifying the Posture of Young Adults in Walking Videos by Using a Fusion Artificial Intelligent Method. *Biosensors* **2022**, *12*, 295. <https://doi.org/10.3390/bios12050295>

Received: 9 March 2022

Accepted: 2 May 2022

Published: 3 May 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 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

The OpenPose algorithm is a deep learning method in which part affinity fields (PAFs) are used to detect the two-dimensional (2D) postures of humans in images [1]. The relationship between posture stability, motor function, and quality of life has been determined [2,3]. Moreover, the OpenPose algorithm has been used for checking the medication situations of patients and for their physical monitoring [4,5]. The evaluation of the cardinal symptoms of resting tremor and bradykinesia for Parkinson's disease has been conducted using an OpenPose-based deep learning method [6,7]. Furthermore, in [8], the OpenPose framework was used to create a human behavior recognition system for skeleton posture estimation. Quantitative gait (motor) variables can be estimated and recorded using pose tracking systems (e.g., OpenPose, AlphaPose, and Detectron) [9]. These factors are useful for measuring the quality of life of older adults [10–12]. Moreover, parkinsonian motion features have been created using deep-learning-based 2D OpenPose models [13,14]. For people

with autism spectrum disorder, skeleton posture characteristics are correlated with long-term memory in the field of action recognition [15–17]. The physical function of a patient should be assessed according to their health data obtained using a skeleton pose tracking device and gait analysis [18–21]. Many neurological and musculoskeletal disorders are associated with problems related to postural movement, which can be estimated using a pose-capturing device [22]. Therefore, noninvasive tracking devices are used to record, analyze, measure, and detect the postural control of the body, which may indicate health problems in real time. In this study, fusion deep learning was used to generate dynamic joint node plots (DJNPs) by using OpenPose-based methods, and skewness in walking was qualitatively analyzed using convolutional neural networks (CNNs) [23]. An iso-block postural identity (IPI) method was used to perform the quantified analysis of postural control and walking behavior. This proposed approach combines cell phone camera recordings, cloud storage, and fusion deep learning for postural estimation and classification.

2. Materials and Methods

2.1. Research Ethics

All the experimental procedures were approved by the Institutional Review Board of E-DA Hospital [with approval number EMRP52110N (04/11/2021)]. Verbal and written information on all the experimental details was provided to all the participants before they provided informed consent. Written informed consent was obtained from the participants prior to experimental data collection.

2.2. Flow of Research

In this study, videos walking toward and away from a cell phone camera were recorded using the camera (Step 1 in Figure 1). The videos were recorded at 24-bit (RGB), 1080p resolution, and 30 frames per second. The videos were uploaded to Google Cloud through 5G mobile Internet or Wi-Fi (Step 2 in Figure 1). The workstation used in this study downloaded a video, extracted a single frame from the video, and then applied a fusion artificial intelligence (AI) method to this frame (Step 3 in Figure 1). In the aforementioned step, single frames were extracted from an input video (Step 3A), frames with static walking were identified using an OpenPose-based deep learning method (Step 3B), and the joint nodes of the input video were merged into a plot (Step 3C). The obtained DJNP was categorized as representing straight or skewed walking (Step 3D). CNNs were used to classify DJNPs into one of the aforementioned two groups. Two types of deep learning methods were used in the fusion AI method adopted in this study: an OpenPose-based deep learning method and CNN-based methods. The OpenPose-based method is useful for estimating the coordinates of joint nodes from an input image [1]. The adopted CNNs are suitable for the classification of images with high accuracy and robustness.

2.3. Participants

A total of 35 young adults without any health problems were recruited to participate in a walking experiment. The age range was 20.20 ± 1.08 years. The inclusion criteria were healthy adults who were willing to participate and could walk more than 5 m. People with musculoskeletal pain (such as muscle soreness), those who had drunk alcohol or taken sleeping pills within 24 h before the commencement of the experiment, and individuals with limited vision (such as nearsighted people without glasses) were excluded from this study.

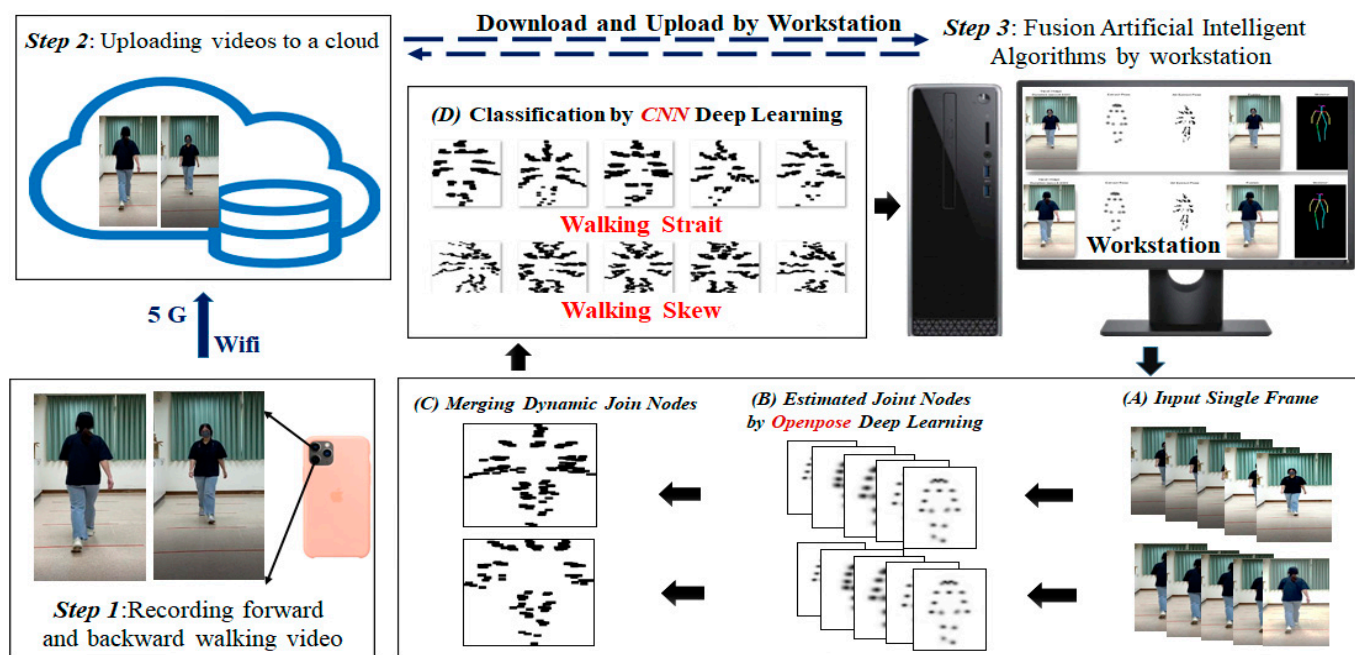


Figure 1. Flow of research.

2.4. Experimental Design

The experimental setup is depicted in Figure 2. The total length of the experimental space was greater than 7 m. The ground was level, free of debris, and smooth to ensure a straight and smooth walking path. The cell phone was placed 1 m above the ground (approximately equal to the height of a medium-sized adult holding a cell phone) and 2 m from the endpoint of the walking path. The entire body of a participant was recorded during the walk. The participants were required to wear walking shoes and not slippers while walking. Participants walked away from the cell phone and then turned back and walked toward the cell phone. The participants walked for 5 m toward and away from the camera three times each. One video was captured for each 5-m walk; thus, six videos were recorded for each participant. A series of single (static) frames was extracted from a video every 0.3 s. For example, for a 3-s input video, 10 frames were extracted to estimate the coordinates of joint nodes. A static frame of one DJNP was extracted per 0.3 s for one video. For example, a 10 s walking video with frame rate 30 (frames/second), the total static frame in one DJNP are 90 frames (i.e., $90 = 10 \text{ (second)} \times 30 \text{ (frames/second)} \times 0.3 \text{ (second)}$). Hence, the DJNP was a variety of frames according to the length of a walking video. The filmmakers are not medical experts but are trained in motion assessment. The video is analyzed by an expert in image analysis and an occupational therapist specializing in rehabilitation Table 1 lists the number of participants and the mean and standard deviation (STD) of velocity (m/s) and time (s) for each group.

Table 1. Information on the number of participants and the mean and standard deviation (STD) of velocity (m/s) and time (s) for each group.

Group	N	Mean Velocity (m/s)	STD Velocity (m/s)	Mean Time (s)	STD Time (s)
Skew	102	0.68	0.08	7.48	0.84
Straight	108	0.69	0.08	7.39	0.91

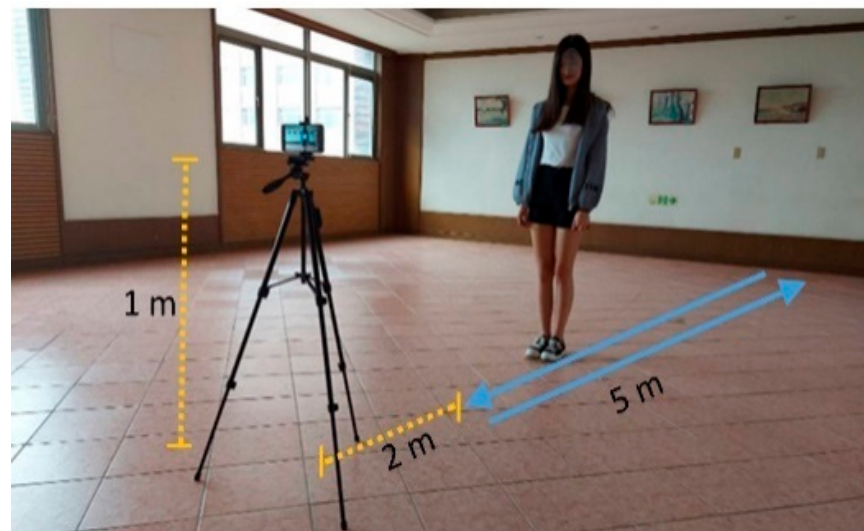


Figure 2. Experimental setup (the cell phone was placed 1 m above the floor and 2 m from the participant).

2.5. Measurement of Joint Nodes through Openpose-Based Deep Learning

OpenPose is a well-known system that uses a bottom-up approach for real-time multiperson body pose estimation. In the proposed OpenPose-based method, PAFs are used to obtain a nonparametric representation for associating body parts with individuals in an image [1]. This bottom-up method achieves high accuracy in real time, regardless of the number of people in the image. It can be used to detect the 2D poses of multiple people in an image and to perform single-person pose estimation for each detection. In this study, the OpenPose algorithm was mainly used to output a heat map of joint nodes (Figure 3). The center coordinates of joint nodes were estimated by using the geometric centroid formula.

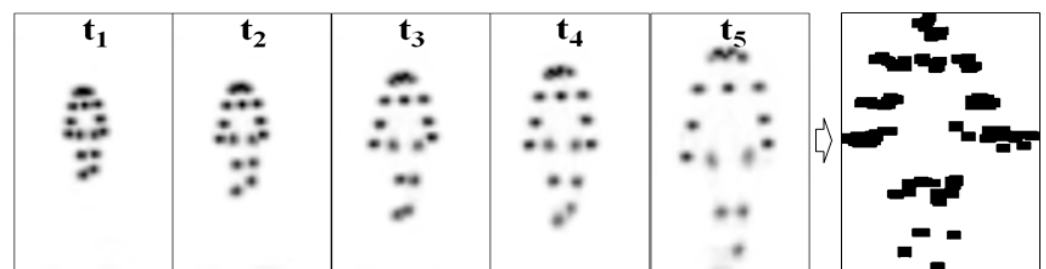


Figure 3. Dynamic joint node plot (DJNP) (right) obtained by merging the heat maps of joint nodes from t_1 to t_5 by using the OpenPose algorithm.

2.6. Definition of the Control and Experimental Groups

The data for the control group comprised DJNPs that indicated straightforward walking toward and away from the camera. The experimental group comprised DJNPs that indicated skewed walking. The data for the control and experimental groups comprised 102 and 108 DJNPs, respectively, which were classified using different CNNs.

2.7. Classification Using Pretrained CNNs and Machine Learning Classifiers

Pretrained CNNs were used to extract the features of DJNPs, and machine learning classifiers were used to construct classification models. The eight pre-trained CNNs used in this study were AlexNet, DenseNet201, GoogleNet, MobileNetV2, ResNet101, ResNet50, VGG16, and VGG19. Moreover, the three machine learning classifiers used in this study were logistic regression (LR), naïve Bayes (NB), and support vector machine (SVM).

CNNs have a high learning capacity, which makes them suitable for image classification. They extract features and learn data according to variations in the breadth and depth of features. Table 2 lists the features that were extracted by CNNs and served as the inputs for the LR, NB, and SVM. A deep CNN network comprises five types of primary layers: a convolutional layer, a pooling layer, a rectified linear unit layer, fully connected layers, and a softmax layer. Information on the pretrained CNNs used in this study is provided in Table 2. The fully connected layers of the CNNs extracted and stored the features of the input image. In the present study, eight CNNs and three classifiers with four batch sizes and 20 random splits were adopted. The four batch sizes selected in this study for the CNNs were 5, 8, 11, and 14. The total number of investigated models was $8 \text{ (CNNs)} \times 3 \text{ (machine learning techniques)} \times 4 \text{ (batch size settings)} \times 20 \text{ (instances of random splitting)} = 1920$. Therefore, the 1920 models represent the 1920 possible combinations of one CNN, classifier, batch size, and random data split. CNNs have demonstrated utility and efficiency in image feature extraction in the fields of biomedicine and biology [23–27].

Table 2. Information on the adopted convolutional neural networks.

CNN	Image Size	Layers	Parametric Size (MB)	Layer of Features
AlexNet	227 × 227	25	227	17th (4096 × 9216)
DenseNet201	224 × 224	709	77	706th (1000 × 1920)
GoogleNet	224 × 224	144	27	142nd (1000 × 1024)
MobileNetV2	224 × 224	154	13	152nd (1000 × 1280)
ResNet101	224 × 224	347	167	345th (1000 × 2048)
ResNet50	224 × 224	177	96	175th (1000 × 2048)
VGG16	224 × 224	41	27	33rd (4096 × 25,088)
VGG19	224 × 224	47	535	39th (4096 × 25,088)

LR is a process of modeling the probability of a discrete outcome when an input variable is given. This process is often used to analyze associations between two or more predictors or variables. LR does not require the existence of a linear relationship between inputs and output variables. This method is useful when the response variable is binary, but the explanatory variables are continuous. LR is also an effective analysis method for classification problems. The LR method is used for the development of classification models in the field of machine learning because of its capacity to provide hierarchical or tree-like structures. Many fields have adopted LR for prediction and classification. LR is suitable for classification problems related to health issues, such as whether a person has a specific ailment or disease when a set of symptoms are given.

NB classifiers are based on Bayes' theorem with a naïve independence hypothesis between the adopted predictors or features. These classifiers are the most suitable ones for solving classification problems in which no dependency exists between a particular feature and other features of a certain class. NB classifiers offer high flexibility for linear or nonlinear relations among variables (features or predictors) in classification problems and provide increased accuracy when combined with kernel density estimation. NB classifiers exhibit higher performance for categorical input data than for numerical input data. These classifiers are easy to implement and computationally inexpensive, perform well on large datasets with high dimensionality, and are extremely sensitive to feature selection.

SVM classifiers are highly powerful classifiers that can be used to solve two-class pattern recognition problems. They transform the original nonlinear data into a higher-dimensional space and then create a separating hyperplane defined by various support vectors in this space to maximize the margin between two datasets. Data can be linearly separated in the higher-dimensional space by using a kernel function. Many useful kernels are available to improve the classification performance and reduce the false rate. SVM is a supervised learning method for the classification of linear and nonlinear data and is generally used for the classification of high-dimensional or nonlinear data.

The computing time of using SVM is in linear time, rather than by expensive iterative approximation, which is performed by many other types of classifiers. The LR, NB, and SVM methods were applied as deep and machine learning methods to extract features of DJNPs and classify the postural control of the straight and skewed walking groups.

2.8. Validation of Classification Performance

The data for the control and experimental groups comprised 102 and 108 DJNPs, respectively. A random splitting schema was employed to separate the training (70%) and testing (30%) sets; 71 and 31 samples from the control group were used for training and testing, respectively, and 76 and 32 samples from the experimental group were used for training and testing, respectively. Testing sets and confusion matrices were used to evaluate the models with respect to the kappa value, accuracy, sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). These indices were sorted in the ascending order of the corresponding kappa value, and a radar plot was then generated to present the aforementioned indices of the adopted models.

3. Results

In this study, 70% of the samples of each group were randomly used to train the adopted classifiers, and the remaining 30% of samples were used to perform validation. Figure 4 shows a scatter plot for the specificity and sensitivity of the 1920 models for the validation dataset. The maximum specificity and sensitivity of 0.84 and 0.87, respectively, were achieved by the ResNet101 and NB classifiers, respectively.

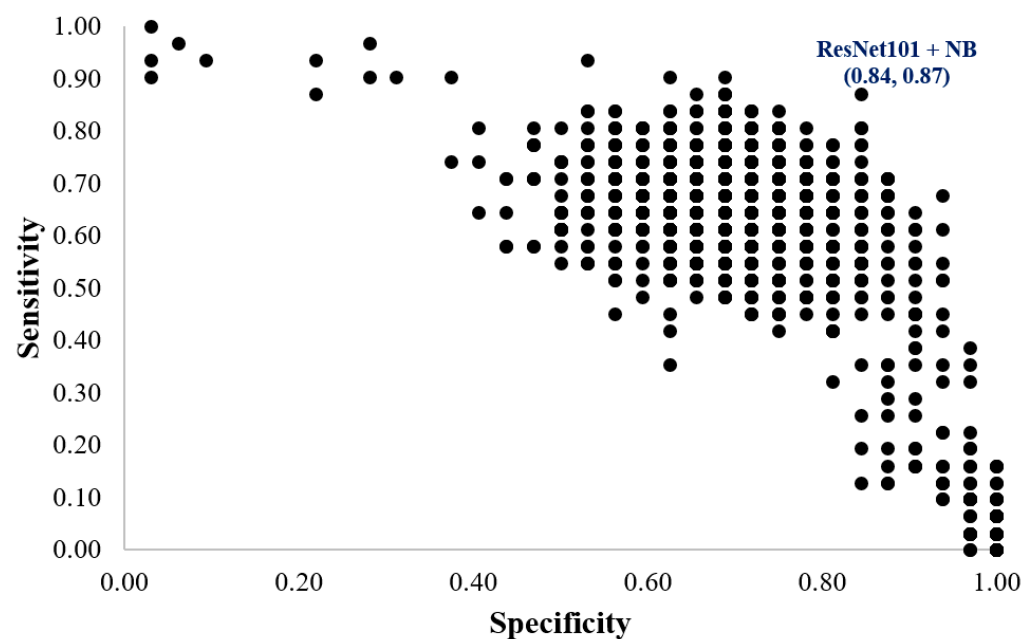


Figure 4. Scatter plot for the specificity and sensitivity of the 1920 models for the validation dataset.

In Figure 5, a radar plot was constructed for six performance indices with the results sorted by the maximum kappa value for 96 models (the abbreviations of the investigated models are written in Appendix A). The best performing model was M53, which is a combination of ResNet101 and naïve Bayes. The kappa, accuracy, sensitivity (Sen), specificity (Spe), PPV, and NPV values were 0.71, 0.86, 0.87, 0.84, and 0.87, respectively. All of the performance indices are over 0.7. The optimized model, ResNet101 with naïve Bayes, had acceptable agreement results and the highest accuracy.

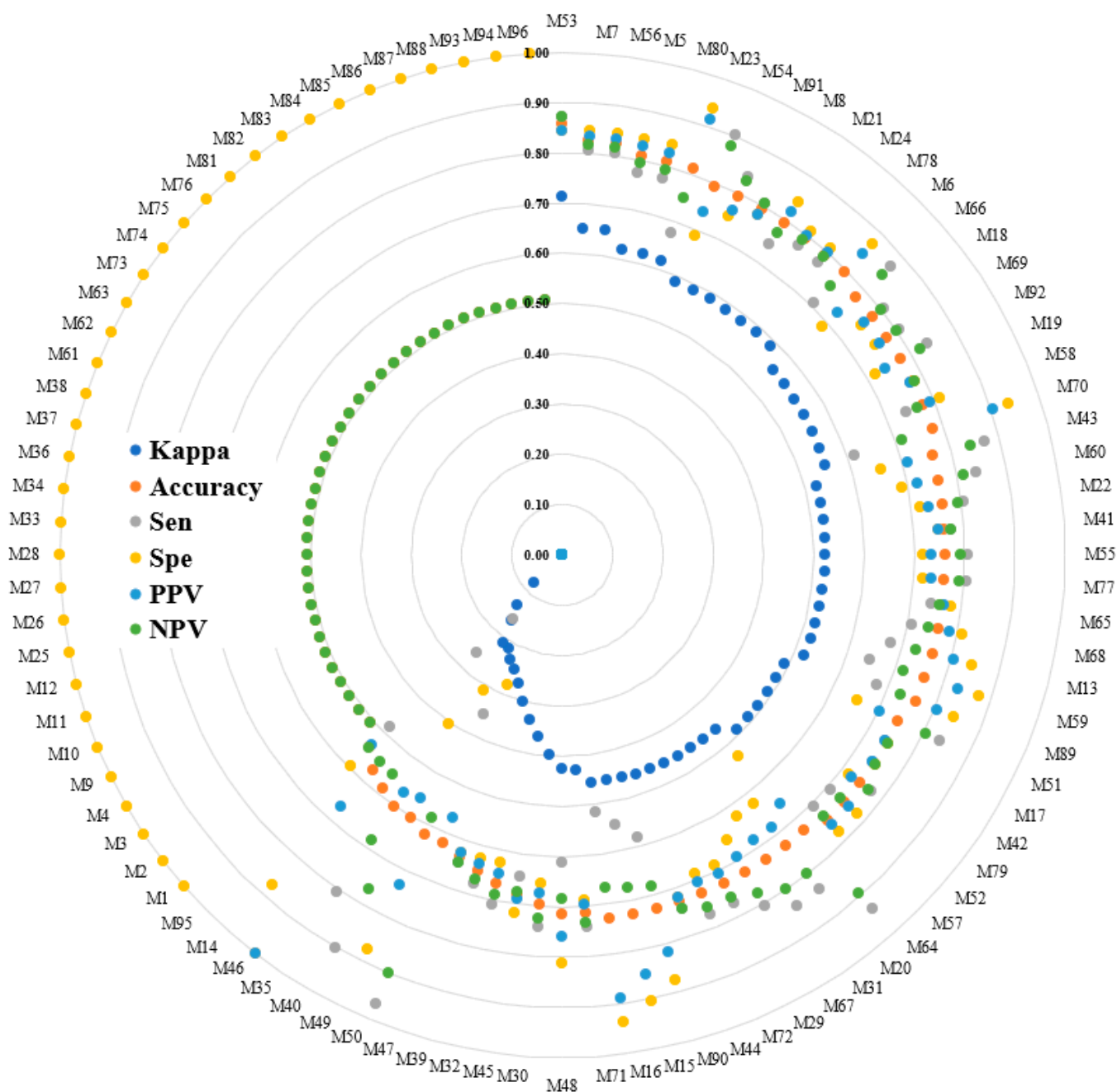


Figure 5. Radar plot of the six performance indices sorted in the ascending order of the kappa value for 96 models (the abbreviations are explained in the Appendix A). Sen represents the sensitivity, and Spe represents the specificity.

Table 3 lists the 13 models with kappa values greater than 0.59. These models comprised four (30.8%) AlexNet models, three DenseNet201 models (23.1%), three ResNet101 models (23.1%), two VGG16 models (15.4%), and one VGG19 model (7.7%). AlexNet, DenseNet201, and ResNet101 accounted for 10 of the aforementioned 13 models (76.9%). SVM and NB were the main machine learning classifiers that performed well in this study. The numbers of the aforementioned 13 models with SVM and NB classifiers were 3 and 10, respectively. Thus, NB performed well. Finally, the batch sizes of the 13 models were 5, 8, 11, and 14 useable in this work.

Table 3. Models with kappa values greater than 0.59.

CNN	Classifier	Batch Size	Model	Kappa	Accuracy	Sen	Spe	PPV	NPV
ResNet101	NB	5	M53	0.71	0.86	0.87	0.84	0.84	0.87
AlexNet	NB	11	M7	0.65	0.83	0.81	0.84	0.83	0.82
ResNet101	NB	14	M56	0.65	0.83	0.81	0.84	0.83	0.82
AlexNet	NB	5	M5	0.62	0.81	0.77	0.84	0.83	0.79
VGG16	NB	14	M80	0.62	0.81	0.77	0.84	0.83	0.79
DenseNet201	SVM	11	M23	0.62	0.81	0.68	0.94	0.91	0.75
ResNet101	NB	8	M54	0.59	0.79	0.90	0.69	0.74	0.88
VGG19	NB	11	M91	0.59	0.79	0.84	0.75	0.77	0.83
AlexNet	NB	14	M8	0.59	0.79	0.81	0.78	0.78	0.81
DenseNet201	SVM	5	M21	0.59	0.79	0.74	0.84	0.82	0.77
DenseNet201	SVM	14	M24	0.59	0.79	0.77	0.81	0.80	0.79
VGG16	NB	8	M78	0.59	0.79	0.77	0.81	0.80	0.79
AlexNet	NB	8	M6	0.59	0.79	0.71	0.88	0.85	0.76

4. Discussion

4.1. Measurement of Postural Control

IPIs were used to measure the skewness or displacement. Figure 6 illustrates the fusion of a DJNP with the IPI generated for a series of time points. In this study, an IPI was created every 0.3 s, and all the IPIs were fused with DJNPs.

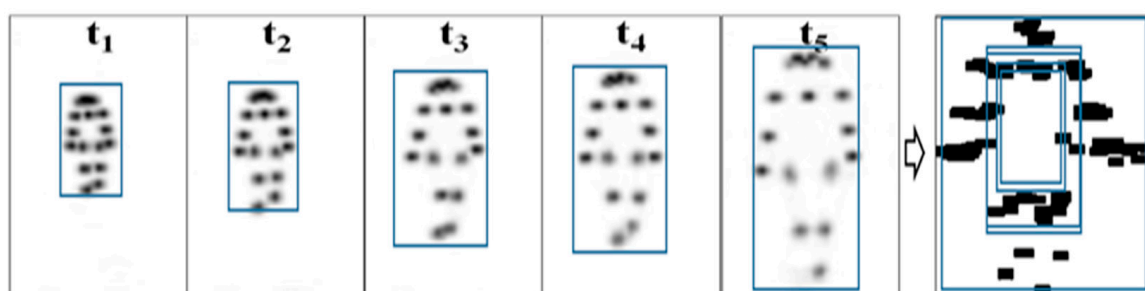


Figure 6. Iso-block postural identity (IPI) generated for a series of times and fusion of the IPI with a DJNP (right).

Figure 7 presents the skewness or displacement for a walking video at three time points (i.e., t_0 , t_1 , and t_2). Figure 7A,C,D,F depict DJNPs and IPIs for skewed walking. Figure 7B,E depict DJNPs and IPIs for straight walking. These DJNPs can be used to measure skewness and horizontal postural movement.

The parameters Θ_r and Θ_l represent the angles of the right and left sides of the body during captured images, respectively (Figure 7E). The ratio of two angles (i.e., $SR = \Theta_l / \Theta_r$) was used to measure the skewness tendency. When this ratio is >1 , the body tends to skew to the right. When $SR = 1$, the body is almost straight. When SR is <1 , the body tends to skew to the left. The displacement of the body between two time points was quantified by estimating the distance covered between these time points. For example, in Figure 7B,E, $D_{r,0,1}$ and $D_{l,0,1}$ represent the displacements of the right and left sides of the body, respectively between t_0 and t_1 . Similarly, $D_{r,1,2}$ and $D_{l,1,2}$ represent the displacements of the right and left sides of the body, respectively, between t_1 and t_2 . Therefore, the ratio of $D_{r,i-1,i}$ to $D_{l,i-1,i}$ (i.e., $MD = D_{r,i-1,i} / D_{l,i-1,i}$, $i = 0, 1, 2$) could be used to determine the dominant side of body displacement. When MD was >1 , the right side was the dominant side of displacement. When MD was 1, the walking posture was almost straight. Moreover, when MD was <1 , the left side was the dominant displacement side.

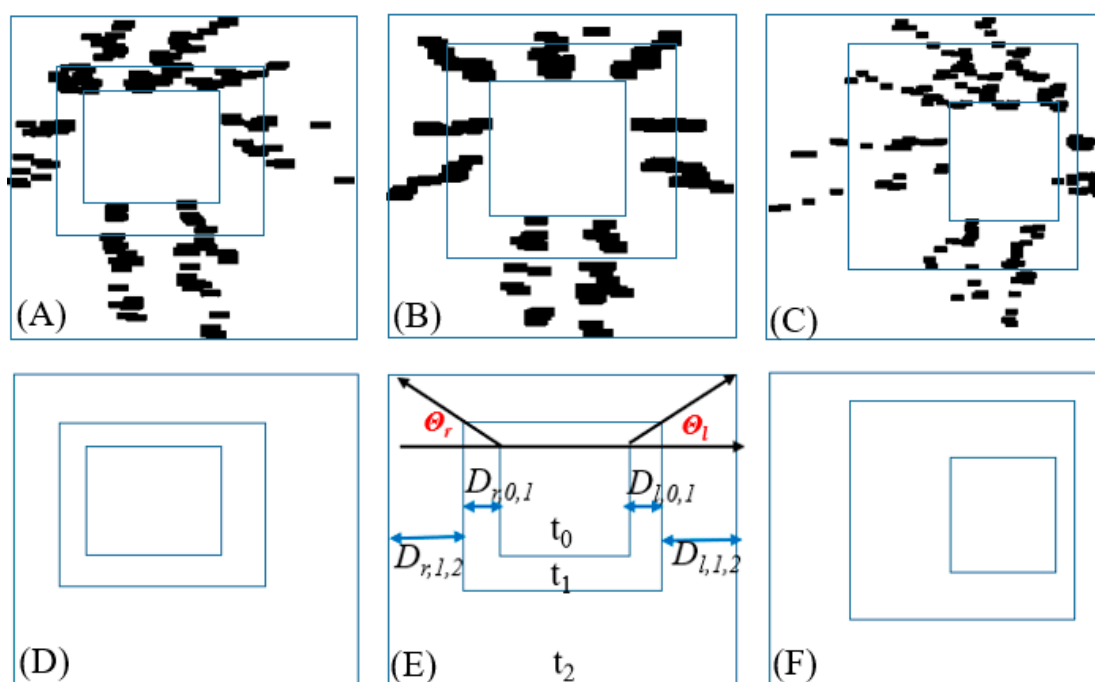


Figure 7. Graphical representation of the skewness or displacement for a walking video at three time points (i.e., t_0 , t_1 , t_2). (A,D), (B,E), and (C,F), respectively, present postural skew to the left, postural balance, and postural skew to the right with participants walking toward the camera.

4.2. Literature for Health Issues and Postural Control during Walking

Poor postural control during walking may indicate health problems. An individual's postural control considerably influences their quality of life [2,3]. Equipping participants with wearable devices that assess their posture can be challenging [4]. Nevertheless, this problem can be overcome by incorporating deep learning into Internet of things monitoring systems to effectively detect motion and posture [5]. Resting tremors and finger tapping have been detected using OpenPose-based deep learning methods [6,7]. Moreover, skeleton normality has been determined through the measurement of angles and velocities by using the aforementioned methods [8–10]. Such methods are useful for not only generating three-dimensional poses [11,12] but also for identifying the relationship between postural behavior and functional diseases, such as Parkinson's disease [6,13,14], autism spectrum disorder [15], and metatarsophalangeal joint flexions [16]. OpenPose-based deep learning methods can be used for skeleton, ankle, and foot motion [8,17] detection; physical function assessment [18,19]; and poststroke study [20].

Thus, noninvasive tracking devices play crucial roles in the recording [21], analysis, measurement, and detection of body posture, which may indicate health issues in real time.

5. Conclusions

In this study, fusion deep learning was applied to generate DJNPs by using an OpenPose-based method and quantify skewness by using CNNs. The adopted approach successfully incorporates cell phone camera recording, cloud storage, and fusion deep learning for posture estimation and classification. Moreover, the adopted IPI method can be used to perform a quantified analysis of postural control and walking behavior.

The research conducted in the present study can be considered preliminary. We developed the IPI method and attempted a quantified analysis of postural control and walking behavior to identify factors indicative of possible clinical gait disorders. However, at the time of writing, the research is in the preliminary phase and will remain as such until the automated analysis is completed through the IPI method. The highlights of our proposed method include its suitability for use with computer vision for identifying signs

of gait problems for clinical application, as well as its replacement of a dynamic joint node plot. In addition, the IPI method is straightforward and allows for real-time monitoring. A video of walking behavior can be conveniently recorded in real-time by using a mobile device. A user can easily remove the background from the video and generate dynamic joint node coordinates through fusion AI methods. The developed IPI method allows for use with computer vision to identify postural characteristics for clinical applications.

Future studies can apply the proposed approach to individuals with health problems to validate this approach.

Author Contributions: Conceptualization, P.L., T.-B.C. and C.-H.L.; Data curation, T.-B.C., G.-H.H. and N.-H.L.; Formal analysis, P.L., T.-B.C. and C.-H.L.; Investigation, C.-Y.W.; Methodology, P.L., T.-B.C. and C.-H.L.; Project administration, P.L.; Software, T.-B.C.; Supervision, P.L. and C.-H.L.; Writing—original draft, T.-B.C. and C.-H.L.; Writing—review and editing, P.L. and C.-H.L. All authors have read and agreed to the published version of the manuscript.

Funding: Ministry of Science and Technology of Taiwan funded this research under the grant numbers MOST 110-2118-M-214-001.

Institutional Review Board Statement: The study was conducted in accordance with the guidelines of the Declaration of Helsinki. All experimental procedures were approved by the Institutional Review Board of the E-DA Hospital, Kaohsiung, Taiwan (approval number EMRP52110N).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

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

Appendix A

Table A1. The 96 combinations of investigated models with abbreviation are listed below.

CNN	Classifier	Batch Size	Model	CNN	Classifier	Batch Size	Model	CNN	Classifier	Batch Size	Model
AlexNet	LR	5	M1	GoogleNet	SVM	5	M33	ResNet50	NB	5	M65
AlexNet	LR	8	M2	GoogleNet	SVM	8	M34	ResNet50	NB	8	M66
AlexNet	LR	11	M3	GoogleNet	SVM	11	M35	ResNet50	NB	11	M67
AlexNet	LR	14	M4	GoogleNet	SVM	14	M36	ResNet50	NB	14	M68
AlexNet	NB	5	M5	MobileNetV2	LR	5	M37	ResNet50	SVM	5	M69
AlexNet	NB	8	M6	MobileNetV2	LR	8	M38	ResNet50	SVM	8	M70
AlexNet	NB	11	M7	MobileNetV2	LR	11	M39	ResNet50	SVM	11	M71
AlexNet	NB	14	M8	MobileNetV2	LR	14	M40	ResNet50	SVM	14	M72
AlexNet	SVM	5	M9	MobileNetV2	NB	5	M41	VGG16	LR	5	M73
AlexNet	SVM	8	M10	MobileNetV2	NB	8	M42	VGG16	LR	8	M74
AlexNet	SVM	11	M11	MobileNetV2	NB	11	M43	VGG16	LR	11	M75
AlexNet	SVM	14	M12	MobileNetV2	NB	14	M44	VGG16	LR	14	M76
DenseNet201	LR	5	M13	MobileNetV2	SVM	5	M45	VGG16	NB	5	M77
DenseNet201	LR	8	M14	MobileNetV2	SVM	8	M46	VGG16	NB	8	M78
DenseNet201	LR	11	M15	MobileNetV2	SVM	11	M47	VGG16	NB	11	M79
DenseNet201	LR	14	M16	MobileNetV2	SVM	14	M48	VGG16	NB	14	M80
DenseNet201	NB	5	M17	ResNet101	LR	5	M49	VGG16	SVM	5	M81
DenseNet201	NB	8	M18	ResNet101	LR	8	M50	VGG16	SVM	8	M82
DenseNet201	NB	11	M19	ResNet101	LR	11	M51	VGG16	SVM	11	M83
DenseNet201	NB	14	M20	ResNet101	LR	14	M52	VGG16	SVM	14	M84
DenseNet201	SVM	5	M21	ResNet101	NB	5	M53	VGG19	LR	5	M85
DenseNet201	SVM	8	M22	ResNet101	NB	8	M54	VGG19	LR	8	M86
DenseNet201	SVM	11	M23	ResNet101	NB	11	M55	VGG19	LR	11	M87
DenseNet201	SVM	14	M24	ResNet101	NB	14	M56	VGG19	LR	14	M88
GoogleNet	LR	5	M25	ResNet101	SVM	5	M57	VGG19	NB	5	M89
GoogleNet	LR	8	M26	ResNet101	SVM	8	M58	VGG19	NB	8	M90
GoogleNet	LR	11	M27	ResNet101	SVM	11	M59	VGG19	NB	11	M91
GoogleNet	LR	14	M28	ResNet101	SVM	14	M60	VGG19	NB	14	M92
GoogleNet	NB	5	M29	ResNet50	LR	5	M61	VGG19	SVM	5	M93
GoogleNet	NB	8	M30	ResNet50	LR	8	M62	VGG19	SVM	8	M94
GoogleNet	NB	11	M31	ResNet50	LR	11	M63	VGG19	SVM	11	M95
GoogleNet	NB	14	M32	ResNet50	LR	14	M64	VGG19	SVM	14	M96

References

1. 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.* **2019**, *43*, 172–186. [\[CrossRef\]](#)
2. Ali, M.S. Does spasticity affect the postural stability and quality of life of children with cerebral palsy? *J. Taibah Univ. Med Sci.* **2021**, *16*, 761–766. [\[CrossRef\]](#)
3. Park, E.-Y. Path analysis of strength, spasticity, gross motor function, and health-related quality of life in children with spastic cerebral palsy. *Health Qual. Life Outcomes* **2018**, *16*, 70. [\[CrossRef\]](#) [\[PubMed\]](#)
4. Roh, H.; Shin, S.; Han, J.; Lim, S. A deep learning-based medication behavior monitoring system. *Math. Biosci. Eng.* **2021**, *18*, 1513–1528. [\[CrossRef\]](#)
5. Manogaran, G.; Shakeel, P.M.; Fouad, H.; Nam, Y.; Baskar, S.; Chilamkurti, N.; Sundarasekar, R. Wearable IoT Smart-Log Patch: An Edge Computing-Based Bayesian Deep Learning Network System for Multi Access Physical Monitoring System. *Sensors* **2019**, *19*, 3030. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Park, K.W.; Lee, E.-J.; Lee, J.S.; Jeong, J.; Choi, N.; Jo, S.; Jung, M.; Do, J.Y.; Kang, D.-W.; Lee, J.-G.; et al. Machine Learning-Based Automatic Rating for Cardinal Symptoms of Parkinson Disease. *Neurology* **2021**, *96*, e1761–e1769. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Heldman, D.A.; Espay, A.; LeWitt, P.A.; Giuffrida, J.P. Clinician versus machine: Reliability and responsiveness of motor endpoints in Parkinson's disease. *Park. Relat. Disord.* **2014**, *20*, 590–595. [\[CrossRef\]](#)
8. Lin, F.-C.; Ngo, H.-H.; Dow, C.-R.; Lam, K.-H.; Le, H. Student Behavior Recognition System for the Classroom Environment Based on Skeleton Pose Estimation and Person Detection. *Sensors* **2021**, *21*, 5314. [\[CrossRef\]](#)
9. Mehdizadeh, S.; Nabavi, H.; Sabo, A.; Arora, T.; Iaboni, A.; Taati, B. Concurrent validity of human pose tracking in video for measuring gait parameters in older adults: A preliminary analysis with multiple trackers, viewing angles, and walking directions. *J. Neuroeng. Rehabilitation* **2021**, *18*, 1–16. [\[CrossRef\]](#)
10. Ota, M.; Tateuchi, H.; Hashiguchi, T.; Ichihashi, N. Verification of validity of gait analysis systems during treadmill walking and running using human pose tracking algorithm. *Gait Posture* **2021**, *85*, 290–297. [\[CrossRef\]](#)
11. Rapczyński, M.; Werner, P.; Handrich, S.; Al-Hamadi, A. A Baseline for Cross-Database 3D Human Pose Estimation. *Sensors* **2021**, *21*, 3769. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Pagnon, D.; Domalain, M.; Reveret, L. Pose2Sim: An End-to-End Workflow for 3D Markerless Sports Kinematics—Part 1: Robustness. *Sensors* **2021**, *21*, 6530. [\[CrossRef\]](#) [\[PubMed\]](#)
13. 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. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Ruppel, S.; Morinan, G.; Peng, Y.; Foltyniec, T.; Sibley, K.; Weil, R.S.; Leyland, L.-A.; Baig, F.; Morgante, F.; Gilron, R.; et al. A Clinically Interpretable Computer-Vision Based Method for Quantifying Gait in Parkinson's Disease. *Sensors* **2021**, *21*, 5437. [\[CrossRef\]](#)
15. Zhang, Y.; Tian, Y.; Wu, P.; Chen, D. Application of Skeleton Data and Long Short-Term Memory in Action Recognition of Children with Autism Spectrum Disorder. *Sensors* **2021**, *21*, 411. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Takeda, I.; Yamada, A.; Onodera, H. Artificial Intelligence-Assisted motion capture for medical applications: A comparative study between markerless and passive marker motion capture. *Comput. Methods Biomech. Biomed. Eng.* **2020**, *24*, 864–873. [\[CrossRef\]](#)
17. Kobayashi, T.; Orendurff, M.S.; Hunt, G.; Gao, F.; LeCursi, N.; Lincoln, L.S.; Foreman, K.B. The effects of an articulated ankle-foot orthosis with resistance-adjustable joints on lower limb joint kinematics and kinetics during gait in individuals post-stroke. *Clin. Biomech.* **2018**, *59*, 47–55. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Clark, R.A.; Mentiplay, B.F.; Hough, E.; Pua, Y.H. Three-dimensional cameras and skeleton pose tracking for physical function assessment: A review of uses, validity, current developments and Kinect alternatives. *Gait Posture* **2019**, *68*, 193–200. [\[CrossRef\]](#)
19. Albert, J.A.; Owolabi, V.; Gebel, A.; Brahms, C.M.; Granacher, U.; Arnrich, B. Evaluation of the Pose Tracking Performance of the Azure Kinect and Kinect v2 for Gait Analysis in Comparison with a Gold Standard: A Pilot Study. *Sensors* **2020**, *20*, 5104. [\[CrossRef\]](#)
20. Ferraris, C.; Cimolin, V.; Vismara, L.; Votta, V.; Amprimo, G.; Cremascoli, R.; Galli, M.; Nerino, R.; Mauro, A.; Priano, L. Monitoring of Gait Parameters in Post-Stroke Individuals: A Feasibility Study Using RGB-D Sensors. *Sensors* **2021**, *21*, 5945. [\[CrossRef\]](#)
21. Han, K.; Yang, Q.; Huang, Z. A Two-Stage Fall Recognition Algorithm Based on Human Posture Features. *Sensors* **2020**, *20*, 6966. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Kidziński, Ł.; Yang, B.; Hicks, J.L.; Rajagopal, A.; Delp, S.L.; Schwartz, M.H. Deep neural networks enable quantitative movement analysis using single-camera videos. *Nat. Commun.* **2020**, *11*, 1–10. [\[CrossRef\]](#)
23. Lee, P.; Chen, T.-B.; Wang, C.-Y.; Hsu, S.-Y.; Liu, C.-H. Detection of Postural Control in Young and Elderly Adults Using Deep and Machine Learning Methods with Joint-Node Plots. *Sensors* **2021**, *21*, 3212. [\[CrossRef\]](#) [\[PubMed\]](#)
24. Bakator, M.; Radosav, D. Deep Learning and Medical Diagnosis: A Review of Literature. *Multimodal Technol. Interact.* **2018**, *2*, 47. [\[CrossRef\]](#)
25. Lee, J.-G.; Jun, S.; Cho, Y.-W.; Lee, H.; Kim, G.B.; Seo, J.B.; Kim, N. Deep Learning in Medical Imaging: General Overview. *Korean J. Radiol.* **2017**, *18*, 570–584. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Suzuki, K. Overview of deep learning in medical imaging. *Radiol. Phys. Technol.* **2017**, *10*, 257–273. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Ravi, D.; Wong, C.; Deligianni, F.; Berthelot, M.; Andreu-Perez, J.; Lo, B.; Yang, G.-Z. Deep Learning for Health Informatics. *IEEE J. Biomed. Health Informatics* **2016**, *21*, 4–21. [\[CrossRef\]](#)