Towards a diagnostic tool for neurological gait disorders in childhood combining 3D gait kinematics and deep learning

Zhengyang Lan, Mathieu Lempereur, Gwenael Gueret, Laetitia Houx, Marine Cacioppo, Christelle Pons, Johanne Mensah, Olivier Rémy-Néris, Abdeldjalil Aissa El Bey, François Rousseau, et al.

▶ **To cite this version:**

Zhengyang Lan, Mathieu Lempereur, Gwenael Gueret, Laetitia Houx, Marine Cacioppo, et al.. Towards a diagnostic tool for neurological gait disorders in childhood combining 3D gait kinematics and deep learning. Computers in Biology and Medicine, 2024, 171, pp.108095. 10.1016/j.compbiomed.2024.108095 . hal-04464978

HAL Id: hal-04464978
https://hal.science/hal-04464978

Submitted on 19 Feb 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

Computers in Biology and Medicine 171 (2024) 108095

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Towards a diagnostic tool for neurological gait disorders in childhood combining 3D gait kinematics and deep learning

Zhengyang Lan [a,d,1,2], Mathieu Lempereur [a,b,c,1,2,*], Gwenael Gueret [c], Laetitia Houx [a,b,c,f], Marine Cacioppo [a,b,c], Christelle Pons [a,b,c,f], Johanne Mensah [a,b,c,f], Olivier Rémy-Néris [a,b,c], Abdeldjalil Aïssa-El-Bey [e], François Rousseau [a,d,2], Sylvain Brochard [a,b,c,f,2]

[a] Laboratoire de Traitement de l'Information Médicale INSERM U1101, Brest, France
[b] Université de Bretagne Occidentale, Brest, France
[c] CHU de Brest, Hôpital Morvan, service de médecine physique et de réadaptation, Brest, France
[d] IMT Atlantique, LaTIM U1101 INSERM, Brest, France
[e] IMT Atlantique, UMR CNRS 6285, Lab-STICC, Brest, France
[f] Fondation Ildys, Brest, France

ARTICLE INFO

Keywords:
Clinical gait analysis
Children
Cerebral palsy
Times series classification
Deep learning
Diagnostic tool

ABSTRACT

Gait abnormalities are frequent in children and can be caused by different pathologies, such as cerebral palsy, neuromuscular disease, toe walker syndrome, etc. Analysis of the "gait pattern" (i.e., the way the person walks) using 3D analysis provides highly relevant clinical information. This information is used to guide therapeutic choices; however, it is underused in diagnostic processes, probably because of the lack of standardization of data collection methods. Therefore, 3D gait analysis is currently used as an assessment rather than a diagnostic tool. In this work, we aimed to determine if deep learning could be combined with 3D gait analysis data to diagnose gait disorders in children.

We tested the diagnostic accuracy of deep learning methods combined with 3D gait analysis data from 371 children (148 with unilateral cerebral palsy, 60 with neuromuscular disease, 19 toe walkers, 60 with bilateral cerebral palsy, 25 stroke, and 59 typically developing children), with a total of 6400 gait cycles. We evaluated the accuracy, sensitivity, specificity, F1 score, Area Under the Curve (AUC) score, and confusion matrix of the predictions by ResNet, LSTM, and InceptionTime deep learning architectures for time series data.

The deep learning-based models had good to excellent diagnostic accuracy (ranging from 0.77 to 0.99) for discrimination between healthy and pathological gait, discrimination between different etiologies of pathological gait (binary and multi-classification); and determining stroke onset time. LSTM performed best overall.

This study revealed that the gait pattern contains specific, pathology-related information. These results open the way for an extension of 3D gait analysis from evaluation to diagnosis. Furthermore, the method we propose is a data-driven diagnostic model that can be trained and used without human intervention or expert knowledge. Furthermore, the method could be used to distinguish gait-related pathologies and their onset times beyond those studied in this research.

1. Introduction

Gait disorders are frequent in children and can result from various pathologies. Cerebral palsy (CP) is the leading cause of gait disorders [1] (3 out of 1000 births) [2]; however, many other childhood gait disorders exist [3], caused by pathologies such as neuromuscular diseases (Duchenne muscular dystrophy, Charcot Marie-Tooth disease, etc.), genetic or orthopedic diseases, toe walker syndrome, etc. At present, the final medical diagnosis is based on the child's development, clinical examination, medical imaging (brain and/or muscle imaging) and sometimes electromyography, muscle biopsy or genetic testing. Although the clinical observation of a child's gait is often very informative, it only serves as a guide and is rarely a primary contributor

* Corresponding author at: CHU de Brest, Hôpital Morvan, service de médecine physique et de réadaptation, Brest, France.
E-mail address: mathieu.lempereur@univ-brest.fr (M. Lempereur).
[1] Co-first author.
[2] Contributed equally.

https://doi.org/10.1016/j.compbiomed.2024.108095
Received 28 July 2023; Received in revised form 31 January 2024; Accepted 1 February 2024
Available online 3 February 2024
0010-4825/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

Table 1
Literature survey on Machine and Deep Learning methods for pathological gait analysis.

Authors	Year	ML/DL	Population	Input	Performance
Eskofier et al. [20]	2013	ML	24 young & 24 elderly females	37 3D marker trajectories	acc.: 95.8%
Begg & Kamruzzaman [32]	2005	ML	12 young & 12 elderly adults	24 gait features	acc.: 91.7%
Wu et al. [22]	2007	ML	24 young & 24 elderly adults	36 gait features	acc.: 91%
Zhou et al. [23]	2020	ML & DL	239 participants (healthy young-middle aged, older adults and geriatric patients)	23 gait features	acc.: 90%
Laroche et al. [24]	2014	ML	20 patients with hip osteoarthritis & 20 controls	12 kinematic angles	acc.:88%
Aich et al. [26]	2018	ML	20 patients with parkinson's disease & 20 older adults	12 gait features	acc.: 98.54%
Rehman et al. [5]	2019	ML	119 patients with parkinson's disease & 184 controls	16 gait features	acc.: 97%
Alharti et al. [27]	2019	DL	review literature		
Medeiros et al. [28]	2016	ML	93 patients with parkinson's disease & 73 controls	vertical GRF	acc.: 81%
Bajpai et al. [25]	2022	DL	356 children with cerebral palsy & 41 controls	Knee joint in sagittal plane using Automated Gait Assessment Score	acc. 98%
Alharti et al. [30]	2021	DL	93 patients with parkinson's disease & 73 controls	vertical GRF	acc.: 95.5%

to the final diagnosis. The information contained in the "gait pattern" (i.e., the way the individual walks) is under used in the diagnostic process, probably because of the lack of standardization of gait analysis methods between clinicians. However, clinical gait analysis has the potential to differentiate between bilateral CP and idiopathic toe walking [4–7], and in certain cases, changes in the gait pattern may lead to questioning of the initial diagnosis [4–7].

In the field of pediatrics, instrumented 3D gait analysis (3DGA) allows a comprehensive analysis of gait data through the calculation of biomechanical parameters from marker-based motion capture system data [8]. In the clinical setting, 3DGA is used to calculate 2D kinematic curves that provide quantitative information about the gait pattern. This information is then used to guide rehabilitation [9], to assess the effect of treatments [10–14], or to help surgical decision making [15–19]. Despite its great potential, 3DGA is currently used as an assessment tool rather than a diagnostic tool. In research, instrumented 3DGA has recently been combined with machine learning (ML) or deep learning (DL) (Table 1) to discriminate between the gait patterns of people of different ages [20–23], to assess gait ability [24,25], to detect Parkinson's disease [5,26–28], to help clinical decision making [29], and to detect disease stage [30]. Few preliminary studies in adults combining 3D gait data and deep learning are carried out for diagnosis purposes.

We wished to determine if 3D gait kinematics considered as unique representation of the gait [31] combined with to DL methods (3DGK+DL) could provide a new diagnostic tool for gait disorders in children. To answer this question, we set 3 aims: (1) to determine how accurately 3DGK+DL could distinguish between pathological (children with CP) and typical (typically developing children) gait; (2) to determine how accurately the method could distinguish between different pathologies (CP and neuromuscular disease; and CP and toe walkers); and (3) to determine how accurately the method could distinguish between stroke that occurred before or after the age of 2 years to assess the potential of dating a medical event based on the gait pattern.

To fulfill these aims, we also evaluated the performance of different deep learning networks, which were selected to deal with the multi-dimensional time series signals produced by gait signals. We assessed the diagnostic accuracy for a single gait cycle of a single child and for multiple gait cycles recorded for a single child, as is usual in clinical practice. Our main hypothesis was that 3DGK+DL would provide sufficiently accurate results for use as a new diagnostic tool for children's gait disorders.

2. Method

2.1. Gait dataset

The 3D gait data were collected between 2006 and 2022 in the motion analysis laboratory of CHU Brest (Brest University Hospital), France. Initially (from 2006 to 2007), 7 infrared cameras, then (from 2007 to 2016) 9 cameras and finally (from 2016 to 2022), 15 cameras were used to track the markers. The data were filtered using a second order, zero-lag Butterworth low-pass filter with a cut-off frequency of 6 Hz. The markers were placed in accordance with Kadaba et al. (1990) [8]. Lower limb kinematics were computed using the PlugIn-Gait from Vicon Nexus according to the conventional gait model and were time normalized to the gait cycle [33]. The input data for each diagnostic task was a concatenated version of the kinematic signals from both limbs. The data used were pelvic tilt, obliquity, and rotation; hip flexion/extension, abduction/adduction, and internal/external rotation; knee flexion/extension, valgus/varus, and internal/external rotation; ankle dorsi/plantar flexion; and foot progression for both legs.

The dataset used contained lower limb kinematics from children with gait disorders walking unassisted and from typically developing (TD) children, all walking at their self-selected speed over a walkway of approximately 10 m in length. The kinetics, the electromyographic signals and the spatio-temporal parameters were not used to due to the difficulty that children have to placing one foot on each force platform, the choice of the normalization methods of electromyographic data and the concatenation of time series signals and scalar values.

The cycles of both legs were concatenated to feed into the diagnostic tool. The final gait signal was a 22×101 matrix (gait cycles normalized from 0 to 100%), where the first dimension contained 11 channels for one leg out of the 22 channels for two legs. A minimum of 20 gait cycles per session were used for each child. Examples of data visualization are shown in Appendix A.2. For the TD children, kinematic data from the left and right limbs were randomly used since TD children have symmetrical gait. For the children with gait disorders, kinematic data for the more affected side were inputted first followed by kinematic for the less affected. The side with a higher Gait Profile Score (GPS) was considered as the more affected side. [34]. The order of the concatenation is shown in Table 2.

If a child had more than one gait analysis session, only data from the first session were included in the dataset without any criteria on treatments (medication, rehabilitation or surgery) that the child may have undergone.

All parents provided written informed consent on a form that had been approved by the CHU Brest and the study was conducted according to current French legislation [35]. The study was approved by our local ethical committee (2018CE.31) under the number 29BRC18.0109. Six datasets from the database were used:

- TDvsCPu included the gait kinematics of children with unilateral CP and TD children and was used for the pathology detection task: 148 children (64 girls, 84 boys; 69 children with left side affected, 79 children with right side affected; mean age at first

Table 2
Description of the six datasets used in this study.

Number / Item / Dataset	Pathology	Concatenation order	Child *(n=)*	Cycle *(n=)*
TDvsCPu	TD	Random	59	650
	CPu	Affected Side, Unaffected Side	148	2729
TDvsNM	TD	Random	59	650
	NM	More affected Side, Less affected Side	60	1171
ToeWalkervsCPb	Toe walker	Left Side, Right Side	19	328
	CPb	More affected Side, Less affected Side	60	1095
StrokeBeforeAfter	Before	Left Side, Right Side	20	334
	After	Left Side, Right Side	5	93
NMvsCPb	CPb	More affected Side, Less unaffected Side	60	1095
	NM	More affected Side, Less affected Side	60	1171
TDvsNMvsCPb	TD	Random	59	650
	CPb	More Affected Side, Less unaffected Side	60	1095
	NM	More affected Side, Less affected Side	60	1171

CGA was 7.4 years; 103 children categorized as GMFCS level I, 42 children as GMFCS level II and 3 children as GMFCS level III) with unilateral CP mainly caused by perinatal stroke, prematurity, perinatal hypoxia or cerebral malformation.

- TDvsNM included the gait kinematics of children with neuromuscular (NM) diseases and TD children and was used for the pathology detection task: 60 children (18 girls, 42 boys; mean age at first CGA is 8.8 years; 32 children categorized as GMFCS level I, 27 children as GMFCS level II and 1 child as GMFCS level IV) with NM disease mainly caused by Duchenne muscular dystrophy, Charcot-Marie-Tooth disease and Steinert disease.
- ToeWalkervsCPb included the gait kinematics of children with bilateral CP and toe walkers and was used to evaluate the ability of the deep learning algorithms to discriminate between two distinct pathologies with similar gait patterns: 19 children (11 girls, 8 boys; mean age at first CGA is 9 years; 18 children categorized as GMFCS level I, 1 categorized as GMFCS level II) with a toe walker syndrome; 60 children (36 girls, 24 boys; mean age at first CGA is 8.5 years; 22 children categorized as GMFCS level I, 20 children as GMFCS level II, 16 children as GMFCS level III and 2 children as GMFCS level (IV)) with a bilateral CP mainly caused by prematurity and perinatal hypoxia.
- StrokeBeforeAfter included the gait kinematics of children with stroke with the date of the stroke and was used to detect the onset time of the stroke: 20 children (13 girls, 7 boys; mean age at first CGA is 5.7 years; 14 children categorized as GMFCS level I and 6 children as GMFCS level II) with left hemiparesis caused by stroke before the age of 2 years; 5 children (3 girls, 2 boys; mean age at first CGA is 10.4 years; 2 children categorized as GMFCS level I and 3 children as GMFCS level II) who experienced right hemisphere stroke after the age of 2 years. The gait kinematics were labeled according to two categories: stroke before or after the age of two years.
- NMvCPb included the gait kinematics of children with bilateral CP and NM disease and was used to assess the ability of deep learning algorithms to detect whether the disorder was caused by cerebral or muscular pathology.
- TDvsNMvsCPb included the gait kinematics of children with bilateral CP, NM disease and TD children and was used to evaluate the ability of the deep learning algorithms to differentiate between multiple etiologies of pathological gait (multi-classification).

An overview of the children and gait kinematics included in the database is provided in Table 2. All diagnoses were double-checked by the medical teams and a physician dedicated to the study who reviewed the medical histories and diagnosis records of all 371 children.

2.2. Deep learning methods proposed

We used DL time series classification (TSC) to diagnose the pathology (or date of onset for stroke) from the 3D clinical gait signals. After reviewing DL methods for TSC [36], we chose to test the ability of three DL architectures for TSC : ResNet [37,38], LSTM [39] and InceptionTime [40].

2.2.1. ResNet

The ResNet architecture is described in [37] and has previously been used for TSC [38], as shown in Fig. 1.

The ResNet contains three building blocks. The first two building blocks consist of three convolutional layers with a kernel size of [8, 5, 3], and each convolutional layer is followed by a batch normalization layer and a ReLU activation. In addition, there is a residual connection from the input to the output of the first two building blocks, with a convolutional layer of kernel size 1, followed by a batch normalization layer. The last building block differs from the first two in that there is no convolutional layer or batch normalization layer in the residual connection part. Finally, the output of the features from the last building block are fed into a dense layer with softmax activation, after passing through GlobalAveragePooling.

2.2.2. LSTM

Similarly to ResNet, the Long Short Term Memory model (LSTM) [39] has been shown to successfully extract time series features form a dataset [41–44]. The main component of this network is the LSTM unit, as shown in Fig. 2.

x_t is the input vector with dimension d. The f_t, i_t and o_t are three activation vectors for which the dimension and range of the number of hidden layers h is between 0 and 1. The activation vectors correspond to forget gate, input gate, and output gate. The h_t is the output vector of the LSTM unit and has the same dimension as the number of hidden layers. The $\tilde{c}_t$ is referred to as the cell input activation vector with dimension h. The activation vectors correspond to forget gate, input gate, and output gate. σ_g is a sigmoid function that scales the input value between 0 and 1. σ_c and σ_h are hyperbolic tangent

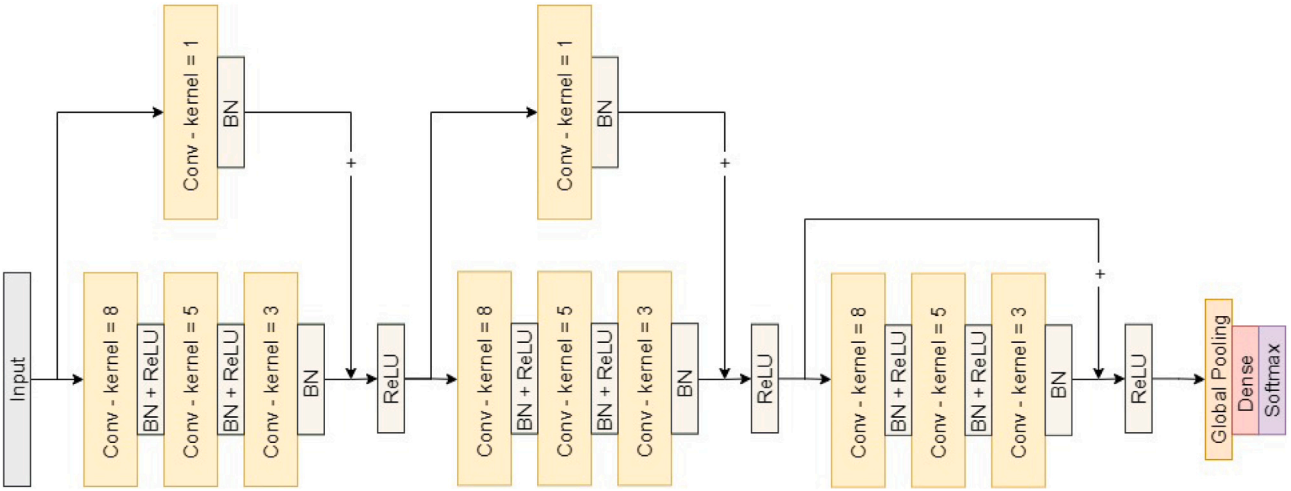


Fig. 1. Architecture of the ResNet for TSC [38]. The yellow blocks represent the convolutional layer, the beige blocks represent the activation function ReLU and/or batch normalization, the orange block represents global average pooling operation, the red block represents the dense layer with two output features, the purple block represents the softmax function, the diverging arrows represent duplication, and the converging arrows represent residual connection.

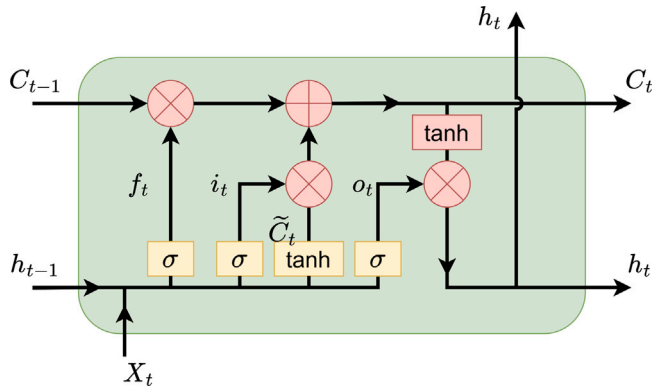


Fig. 2. LSTM unit at time t [39]. The yellow blocks represent activation functions, the red blocks represent component-wise operations, the diverging arrows represent duplication, and the converging arrows represent concatenation.

functions, which generate an output value between −1 and 1. The $W \in \mathbb{R}^{h\times d}, U \in \mathbb{R}^{h\times h}$ and $b \in \mathbb{R}^h$ are trainable weight matrices and bias vector parameters.

The LSTM unit is modeled by the following equations:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (1)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (4)$$

$$c_t = f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \quad (5)$$

$$h_t = o_t \odot \sigma_h(c_t) \quad (6)$$

In practice, we cascade C_t and H_t in the order of time t. The number of LSTM units is equal to the time series length n. The LSTM network we used in this study was inspired by [45], and is shown in Fig. 3.

The LSTM network used in this work is bidirectional. Compared to the traditional LSTM, which only considers the forward state features (in the same direction as the time flow), the bidirectional LSTM also needs to feed the input signal into another sequence of LSTM units connected in the reverse direction of the time flow to allow the network to consider the backward state features. The architecture consists of three bidirectional LSTM layers with a hidden state of 400 (Fig. 3). The first two bidirectional LSTM layers output all hidden states, and the last bidirectional LSTM layer outputs only the last hidden state. In addition, the outputs of the bidirectional LSTM layers in different directions are concatenated and passed to a dropout with a probability of 0.2, before being fed to the next layer. After the last bidirectional LSTM layer, the dense layer with softmax activation is used as a linear classifier to make the prediction.

2.2.3. InceptionTime

The InceptionTime network [40] consists of six stacked connected Inception modules, some of which are also connected by a shortcut linear connection layer.

The architecture of the Inception network is shown in Fig. 4. The first layer in the upper right corner is a convolutional layer applied to all channels with a kernel of size 1 and is called the bottleneck layer. Its function is to reduce the dimensionality of the input data to mitigate overfitting and complexity of the model. The output of the bottleneck layer is fed into three convolutional layers with kernels of size 10, 20, and 40, to use different receptive fields for better feature extraction from time series data with different lengths. The bottom part of the Inception module is a Max pooling of size 3, followed by a bottleneck layer. The function of this part is to make the model more robust to small perturbations. Finally, the above four outputs are concatenated as the output features of the Inception module.

The architecture of the Inception network is shown in Fig. 4. The output of the Inception module set is fed into a Global Average Pooling, then passed to a dense layer with two output values, then a softmax is applied to obtain the prediction of the network. The InceptionTime's prediction is the class with the highest probability estimation in the average probability estimation of the five Inception network model predictions. This ensemble strategy mitigates the problem of the high standard deviation of the results caused by randomly initialized weights and stochastic optimization.

2.3. Experimental process

To preprocess the data, the dataset was split into three parts: training (60%), validation (10%), and test set (30%). The training set was used to train the models and the validation set to verify the outcomes of the training process. The test set was used to evaluate the performance of the trained models in our experiment. The details of the data split are shown in Appendix A.1.

All three DL models were implemented using PyTorch Lightning [46]. They were trained by minimizing the categorical cross-entropy

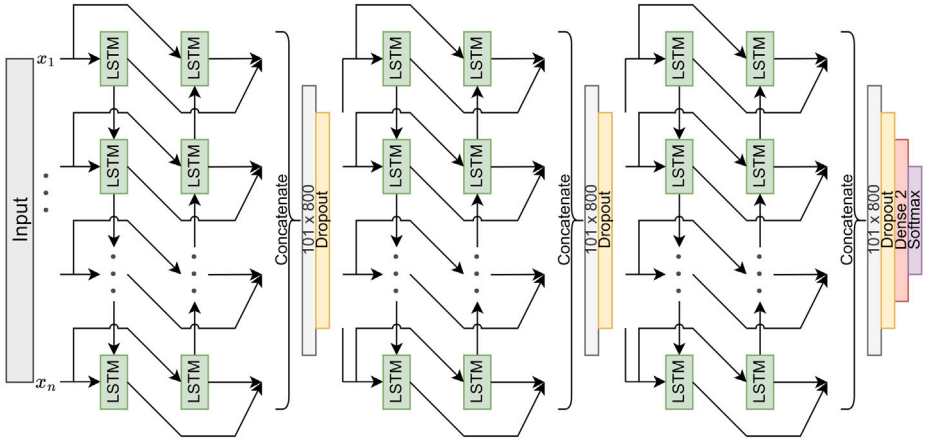


Fig. 3. Architecture of the LSTM network for TSC proposed in this paper. The green blocks represent LSTM units, the gray blocks represent the data, the yellow blocks represent dropout operations, the red block represents the dense layer with two output features, the purple block represents the softmax function, the diverging arrows represent duplication, and the converging arrows represent concatenation.

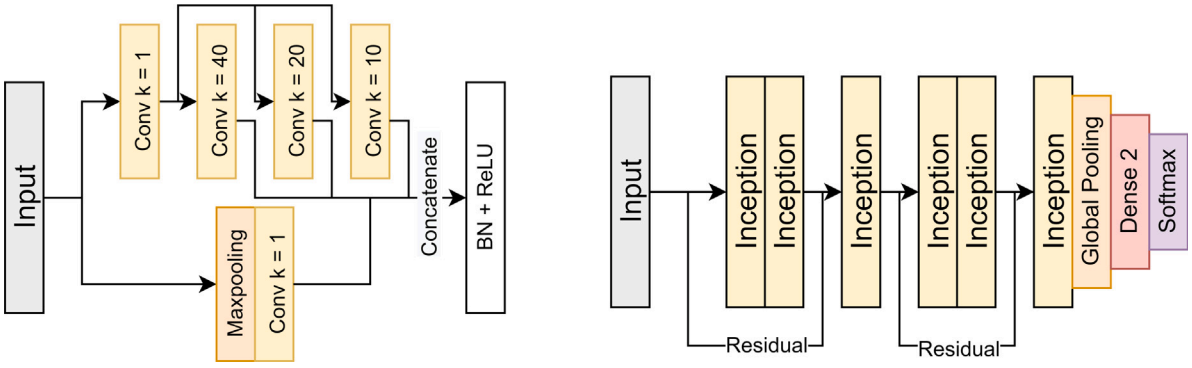


Fig. 4. Architecture of the Inception module (left) and the InceptionTime network (right) [40]. The yellow blocks represent the convolutional layer (left) or Inception module (right), The white block represents the activation function ReLU and batch normalization, the orange block represents the max pooling operation (left) or global average pooling (right), the red block represents the dense layer with two output features, the purple block represents the softmax function, the diverging arrows represent duplication, and the converging arrows represent concatenation (left) or residual connection (right).

loss using the Adam optimizer with an initial learning rate of 0.001 and batch size of 64. The accuracy of the validation set was used as a reference metric to control early stopping and learning rate reduction. For the learning rate reduction, if the validation set loss did not improve over one epoch, the learning rate was decreased by 50%. After ten epochs with no improvement in validation set loss, or after 50 epochs, early stopping was applied. Then the training stopped, and the model with the best validation accuracy was saved. The final prediction was the class with the highest probability estimation in the average probability estimation of the prediction of the five trained models. This was the rule for the InceptionTime, but we also applied it to ResNet and LSTM, because we found that it reduced the standard deviation of the results and mitigated local minima.

2.4. Statistics

We evaluated the models according to the test set predictions. We used six parameters to evaluate the performance of the DL models: accuracy, sensitivity/recall, specificity, F1 score, Area Under the Curve (AUC) score, and confusion matrix.

Accuracy quantifies how close the predictions of the DL model are to the true values. In other words, it gives the proportion of correct predictions in the total number of predictions (formula (7)) to determine if the prediction is balanced in each class and to provide an assessment of the prediction performance of the model. Specificity, equivalent to recall, and sensitivity are defined in formulas (9) and (10). The F1 score is defined in formula (11) and precision is defined in formula (8). The AUC value is the area under the ROC curve, which plots the curve having False Positive Rate (FPR) on the x-axis and True Positive Rate (TPR) on the y-axis. TPR is equivalent to recall, and FPR is defined in formula (12).

Table 3
Definition of the confusion matrix.

Number / Prediction value / True value	Positive	Negative
Positive	True positive	False negative
Negative	False positive	True negative

The Confusion Matrix is used to visualize the performance of an algorithm in machine learning. It consists of True Positive (TP), False Negative (FN), False Positive (FP), and True Negative (TN). The matrix is presented in Table 3.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{7}$$

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{9}$$

$$Specificity = \frac{TN}{TN + FP} \tag{10}$$

$$F_1 = \frac{2}{recall^{-1} + precision^{-1}} \tag{11}$$

$$FPR = \frac{FP}{TP + FN} \tag{12}$$

Table 4
Performance of the three networks to distinguish between pathological and healthy gait using the TDvsCPu dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.92±0.01	**0.93±0.02**	**0.93±0.01**	0.93±0.03	0.91±0.01	0.90±0.04
Sensitivity	**0.96±0.01**	**0.96±0.01**	0.95±0.01	0.95±0.02	0.96±0.01	0.96±0.01
Specificity	0.77±0.03	0.84±0.06	**0.84±0.03**	**0.87±0.06**	0.72±0.08	0.74±0.13
F1	0.95±0.01	0.95±0.01	**0.96±0.01**	**0.95±0.02**	0.95±0.01	0.93±0.02
AUC	0.87±0.01	0.90±0.03	**0.90±0.02**	**0.91±0.03**	0.84±0.04	0.85±0.06

Table 5
Performance of the three networks to distinguish between pathological and healthy gait using the TDvsNM dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.82±0.02	0.77±0.03	**0.85±0.01**	**0.84±0.02**	0.83±0.01	0.78±0.03
Sensitivity	0.88±0.04	0.83±0.06	**0.90±0.01**	**0.88±0.02**	0.90±0.01	0.87±0.03
Specificity	0.71±0.05	0.71±0.07	**0.76±0.03**	**0.79±0.03**	0.70±0.03	0.69±0.04
F1	0.87±0.01	0.78±0.02	**0.89±0.01**	**0.85±0.02**	0.88±0.01	0.80±0.02
AUC	0.79±0.01	0.77±0.02	**0.83±0.01**	**0.84±0.02**	0.80±0.02	0.78±0.03

Table 6
Performance of the three networks to distinguish between pathologies using the ToeWalkervsCPb dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.80±0.03	0.83±0.05	**0.81±0.01**	**0.84±0.02**	0.80±0.01	0.82±0.02
Sensitivity	**0.81±0.03**	0.79±0.06	0.79±0.01	**0.79±0.02**	0.79±0.02	0.77±0.03
Specificity	0.80±0.05	0.97±0.07	**0.87±0.01**	**1.00±0.00**	0.85±0.09	0.98±0.05
F1	0.86±0.02	0.88±0.04	**0.87±0.01**	**0.88±0.01**	0.86±0.01	0.87±0.02
AUC	0.80±0.03	0.88±0.05	**0.83±0.01**	**0.89±0.01**	0.82±0.04	0.88±0.02

Table 7
Performance of the three networks to distinguish between pathologies using the NMvsCPb dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.79±0.01	0.77±0.03	**0.83±0.02**	**0.83±0.02**	0.82±0.01	0.82±0.02
Sensitivity	0.70±0.04	0.65±0.04	**0.73±0.05**	**0.73±0.05**	0.74±0.02	0.70±0.04
Specificity	0.87±0.04	0.89±0.04	**0.90±0.02**	**0.93±0.02**	0.89±0.02	0.93±0.03
F1	0.74±0.02	0.74±0.03	**0.79±0.03**	**0.81±0.03**	0.78±0.02	0.79±0.03
AUC	0.78±0.02	0.77±0.02	**0.82±0.02**	**0.83±0.02**	0.81±0.01	0.82±0.02

Table 8
Performance of the three networks to distinguish between pathologies using the TDvsNMvsCPb dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.75±0.02	0.74±0.03	**0.77±0.02**	**0.77±0.03**	0.76±0.02	0.74±0.02

To mitigate the effect of randomness in training, the values are the mean and standard deviation of the results of the ten models trained for ten different random initializations. The DL models classify the joint angle time series cycles directly. However, predictions for individual children are more useful to clinicians. Since the dataset contains multiple cycles for each child, we used the majority vote of the predictions of cycles for a given child as that child's diagnosis result.

3. Results

The details of the confusion matrix and data set prediction by DL methods can be found in Tables 4, 5, 6, 7, 8 and, 9 and in Appendix A.3. For readability purposes, we report here only the diagnostic accuracy of the "cycle" and "subject" analyses for the best network architecture. The codes for our experiments are available at https://github.com/LANZhengyang/Gait_DeepLearning_Diagnostic_Tool

3.1. Q1 - Distinguishing between pathological and healthy gait

LSTM best distinguished between pathological and healthy gait (TDvsCPu and TDvsNM datasets) (Tables 4 and 5). For the TDvsCPu dataset, LSTM accuracy was 0.93 ± 0.01 and F1 score was 0.96 ± 0.01 for cycles. For subjects, accuracy was 0.93 ± 0.03 and F1 score was 0.95 ± 0.02. For the TDvsNM dataset, LSTM accuracy was 0.85 ± 0.01

Table 9
Performance of the three networks to estimate onset time for a known pathology using the StrokeBeforeAfter dataset. Best performances are in bold.

Performance / Model / Indicator	ResNet		LSTM		InceptionTime	
	Cycles	Subjects	Cycles	Subjects	Cycles	Subjects
Accuracy	0.88±0.11	0.94±0.07	**0.98±0.03**	**0.99±0.04**	0.95±0.04	0.99±0.04
Sensitivity	0.75±0.32	0.70±0.46	**1.00±0.00**	**1.00±0.00**	1.00±0.00	1.00±0.00
Specificity	0.93±0.05	0.98±0.05	**0.97±0.05**	**0.98±0.05**	0.93±0.05	0.98±0.05
F1	0.74±0.26	0.67±0.45	**0.97±0.05**	**0.97±0.10**	0.93±0.06	0.97±0.10
AUC	0.84±0.17	0.84±0.23	**0.99±0.02**	**0.99±0.02**	0.97±0.03	0.99±0.02

and F1 score was 0.89 ± 0.01 for cycles. For subjects, accuracy was 0.84 ± 0.02 and F1 score was 0.85 ± 0.02.

3.2. Q2 - Distinguishing between pathologies

LSTM best distinguished between pathologies (ToeWalkervsCPb, NMvsCPb, and TDvsNMvsCPb datasets) (Tables 6–8). For the ToeWalkervsCPb dataset, accuracy was 0.81 ± 0.01 and F1 score was 0.87 ± 0.01 for cycles. For subjects, accuracy was 0.84 ± 0.02 and F1 score was 0.88 ± 0.01. For the NMvsCPb dataset, accuracy was 0.83 ± 0.01 and F1 score was 0.79 ± 0.03 for cycles. For subjects, accuracy was 0.82 ± 0.02 and F1 score was 0.81 ± 0.03. For the TDvsNMvsCPb dataset, accuracy was 0.77 ± 0.02 for cycles and 0.77 ± 0.02 for subjects.

3.3. Q3 - Estimation of onset time for a known pathology

LSTM best estimated the onset time for a known pathology (StrokeBeforeAfter dataset) (Table 9). Accuracy was 0.98 ± 0.03 and F1 score was 0.97 ± 0.05 for cycles. For subjects, accuracy was 0.99 ± 0.04 and F1 score was 0.97 ± 0.10.

4. Discussion

This study showed that DL combined with 3DGK data from children with pathological gait diagnosed the etiology of the gait disorder (or onset time of the pathology) with a high accuracy. The level of accuracy was similar when using only one gait cycle or the whole set of gait cycle for one given child. The accuracy (0.93) of the LSTM for the TDvsCPu dataset was similar to the accuracy (0.98) of AbnormNet (a neural network with convolutional and fully connected layers) [25], a tool for identifying knee abnormalities during gait in children with cerebral palsy that uses the automated, comprehensive gait assessment score (A-GAS) [47] as input. These results provide new perspectives for gait analysis, enabling 3DGK+DL to be used for the diagnosis of gait disorder etiology.

4.1. Clinical implications

Currently, the diagnosis of a pathology that causes a gait disorder is based on information that is not gait related, such as the child's developmental history, brain MRI, or more invasive exams such as intramuscular EMG or muscle biopsy. The results showed that 3DGK+DL successfully differentiated between pathological and healthy gait, and between the gait patterns of different pathologies in both binary and multiply classification diagnosis. The capacity to diagnose different pathologies using only gait characteristics could be a "game changer". First, this means that the information contained within the gait kinematics (i.e., gait patterns, the "way a person walks"), is specific to the diagnosis, and that there is no need to consider spatiotemporal parameters (speed, step width, etc.) or muscle activation (EMG). Clinicians often formulate hypotheses about the diagnosis from the gait patterns, but those hypotheses are based on experience, and the observations are not objective. The next step for research will be to identify the features of the 3D kinematics used by the algorithm to distinguish between pathologies and to compare them to the features used by clinicians. Second, given the good to excellent results of 3DGK+DL, this tool can be used as part of the diagnostic decision tool e.g., to help to diagnose CP (i.e., differentiate from TD), or to differentiate between CP and other NM diseases in the early stages, without using invasive complementary exams. Our preliminary results pave the way towards such clinical implementation, but further studies are needed at this stage. Thirdly, being able to determine the date of stroke (before or after 2 years of age) using only gait pattern data after the age of two years is a real challenge. If these results are confirmed in a larger dataset, it would mean that the gait kinematics themselves, without spatiotemporal or muscle activation data, contain specific information related to the child's age at the time of the "clinical event". Such information can be crucial in some clinical cases in which the dating of the stroke may help to understand a complex diagnosis that has been made.

4.2. Deep learning - A promising approach

The results show that the three DL models designed to analyze time series signals successfully performed different diagnosis tasks. The LSTM-based model performed better generally, with a smaller standard deviation for the different diagnostic modes and sample sizes. The main difference between DL and conventional ML is that feature extraction is performed automatically by the model in DL. Because the multiple layers enable the progressive extraction of higher-level features from the raw input, there is no need for the manual selection of relevant features, which may require expert knowledge. This architecture has several advantages for the combination of DL with 3DGA: (1) automatic feature engineering reduces data exploration time and the need for domain expertise, (2) DL can handle multi-dimensional and multi-variety data in dynamic or uncertain environments, such as real-world clinical environments, and (3) DL can solve complex problems, such as multi-class disease classification. The high learning capacity and good generalization of DL means it can theoretically be used to fit any function or learn any features. Provided with some time series signals labeled with the corresponding pathology type, the DL model can automatically extract features and learn them without any clinical expertise. This ability is not limited to the dataset and pathologies evaluated in this study; the method could be used to extract specific information related to other pathologies from 3D gait signals, and even non-gait time series signals, and to serve as a new diagnostic tool. An additional advantage of the diagnostic tool proposed in this study is that all inputs in the DL model are captured using an accessible marker-based motion capture system based on a specific standardized protocol, thus ensuring higher accuracy and repeatability than traditional gait analysis measurements. In other words, the diagnostic results of our model are more objective than traditional semi-subjective diagnostics.

4.3. Limitations

This study has four main limitations. First, the datasets were unbalanced. The CP sample in the TDvsCPu dataset was three times larger than the TD sample. In the StrokeBeforeAfter dataset, the Before sample was four times larger than the After sample and the number

of toe walkers was three times smaller than the number of children with bilateral cerebral palsy. Because of this unbalanced distribution of the dataset, DL methods tend to overfit to the majority classes, which causes the model to predict the cycle to the majority class and then drop into local minima, particularly ResNet. This leads to biased and high variance results. However, by averaging the prediction probability estimation over five models, the variance of the results was reduced.

Second, in the test set, some subjects were consistently misdiagnosed by all three models. Furthermore, the three models performed significantly less accurately with the ToeWalkervsCPb dataset than the other datasets. This may be related to limitations in the feature extraction of the chosen DL models or to the fact that a larger training set is required for the model to converge fully.

Third, although the DL models yielded excellent diagnostic results, some interpretation is missing for use in a clinical setting and to improve the clinician's understanding of gait signals.

Finally, the recurrent neural network LSTM combined with 3D gait kinematic data had good to excellent diagnostic accuracy for the 6 datasets. The analysis of DL performance is limited to existing diagnosis types, and it cannot predict the diagnosis for a child with an unknown diagnosis. The accuracy of a multi-classification problem generally decreases as the number of classes increases. To validate a larger multi-classification system a larger database would be needed, for example by combining databases from different 3D CGA laboratories. Our encouraging results on single database justify the need for the development of a larger, multicenter database and DL algorithms. Another approach to determining the etiology of a child with unknown pathology would be to identify similarities between gait signals, assuming that children with the same pathology have the same gait patterns. This diagnostic tool is an adjunct to early diagnosis, based primarily on a combination of clinical history, a standardized neuromotor assessment and magnetic resonance imaging findings. DL combined with 3DGK remains a statistical method.

4.4. Perspectives

4.4.1. Investigation of the misdiagnosed subjects

To solve the problem of consistent misdiagnosis of some subjects, the model could be scaled to reduce the number of parameters and facilitate convergence. Misdiagnosis may also because by the data. To assess the impact of the data on the misdiagnosis, the characteristics of the misdiagnosed cycle could be visualized and analyzed, and the magnitude of the impact of the data could be understood through discussion with clinicians. In addition, although the accuracy of all three DL models was high, it would be pertinent to design a DL architecture specifically for clinical gait analysis that more efficiently extracts pathology-related features, instead of using off-the-shelf and general DL models for TSC.

4.4.2. Interpretation of deep learning model

Further work is required so that 3DGK+DL can be used in clinical settings and clinicians can interpret the results. For example, it is important to understand which features are extracted by DL models for TSC. This would both improve the credibility of the classification results for clinical applications as well as improve understanding of the gait signal. This could be done using Explainable Artificial Intelligence (XAI) methods [48–51].

4.4.3. Other deep learning based approaches for diagnosis

We proposed an approach for the diagnosis of the etiology of gait disorders using DL combined with 3DGA data; however, other approaches could also successfully perform this task. For example, artificial intelligence could be combined with other input sources such as video-based gait analysis or electromyography (EMG), for a high diagnostic accuracy. Video-based gait analysis methods are simpler to use and lower in cost than 3DGA systems [52–54]. DL methods, such as OpenPose [55] and AlphaPose [56] are already used to estimate the human joint position features from the video. ML or DL methods could then make a diagnosis based on the extracted features. The diagnostic accuracy of such methods ranges from 80% to 96% depending on the pathology [57–60]. Gait disorders caused by Parkinson's disease have been the most extensively studied, and the diagnostic accuracy is 96% in a binary classification [58].

EMG records electrical signals produced by skeletal muscles. It plays a significant role in the diagnosis of neuromuscular disorders [61]. Studies have shown that signal processing methods (such as Wavelet Transform [62–64], Hilbert–Huang transform [65]) and machine learning techniques (such as KNN, decision trees, SVM [66,67], Etc.) can diagnose the etiology of gait disorders with up to 99% accuracy [65,66].

5. Conclusion

This study demonstrated that the three DL methods, ResNet, LSTM, and InceptionTime, combined with 3DGK data can accurately perform four types of diagnostic task: distinguishing between pathological and normal gait, detection of pathology, diagnosis of the time of onset of pathology and discrimination between different pathologies. The use of our diagnostic tool requires no human intervention and no expert knowledge. In addition, our work demonstrated that kinematic gait data contain specific information related to the child's etiology. In the clinical setting, where data collection is standardized, the trained model can easily be used to directly assist decision-making. The proposed methods can be easily extended to other diagnoses.

CRediT authorship contribution statement

Zhengyang Lan: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology. **Mathieu Lempereur:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Data curation. **Gwenael Gueret:** Methodology, Data curation. **Laetitia Houx:** Writing – review & editing, Methodology, Data curation. **Marine Cacioppo:** Methodology, Data curation. **Christelle Pons:** Methodology, Data curation. **Johanne Mensah:** Methodology, Data curation. **Olivier Rémy-Néris:** Methodology, Data curation. **Abdeldjalil Aïssa-El-Bey:** Writing – review & editing, Writing – original draft, Supervision, Methodology. **François Rousseau:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Funding acquisition. **Sylvain Brochard:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Data curation.

Declaration of competing interest

None of the authors have any financial or personal conflict of interest with regard to this work.

Acknowledgments

We thank Johanna Robertson, PT, PhD, translator and medical writer for language editing. The research leading to these results was supported by the ANR AI4CHILD Project, Grant ANR-19-CHIA-0015 funded by the French National Research Agency and Brest Metropole Océane.

Appendix

A.1. Data split

See Fig. 5.

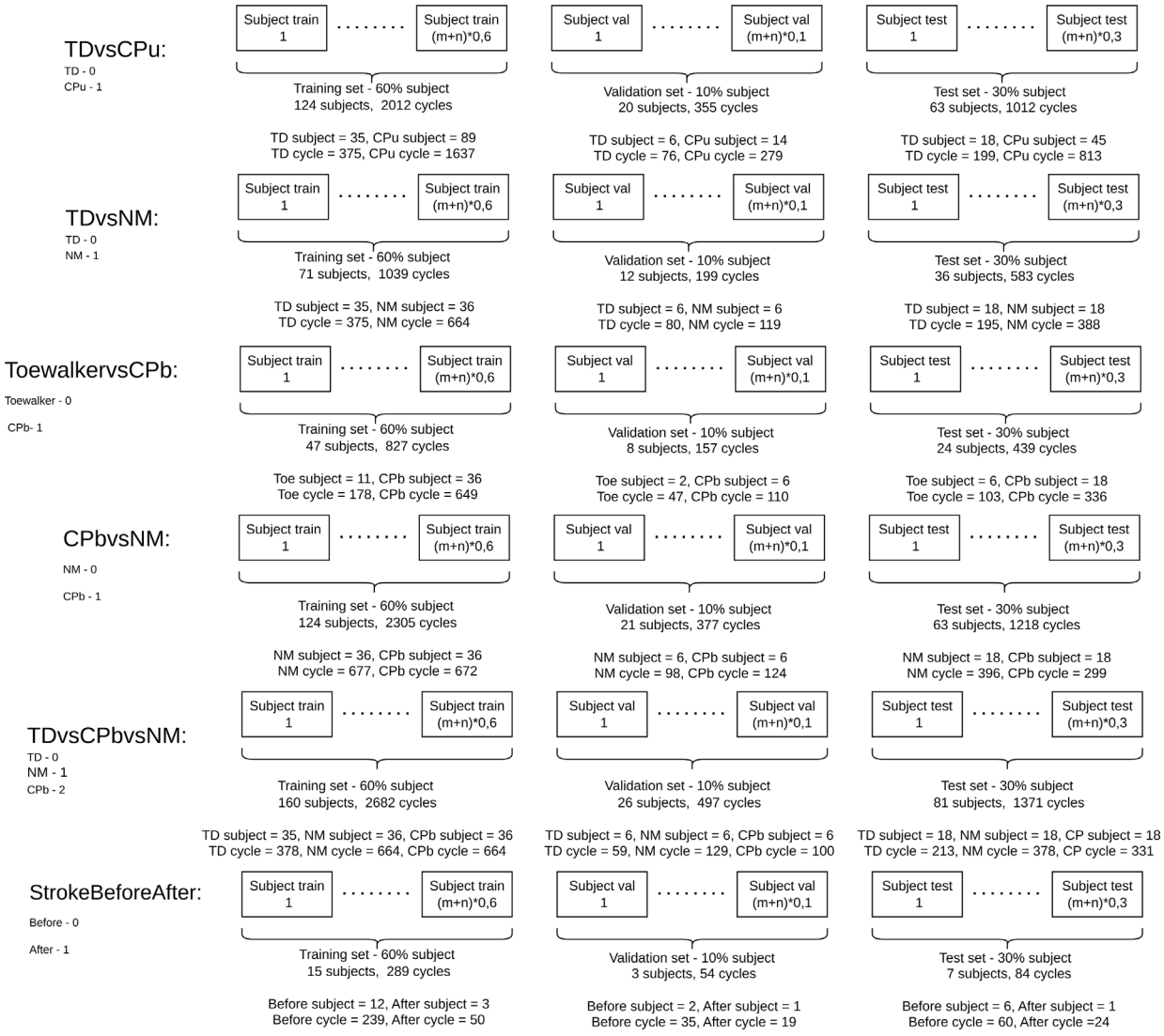


Fig. 5. Data split strategy for the six datasets used in this study.

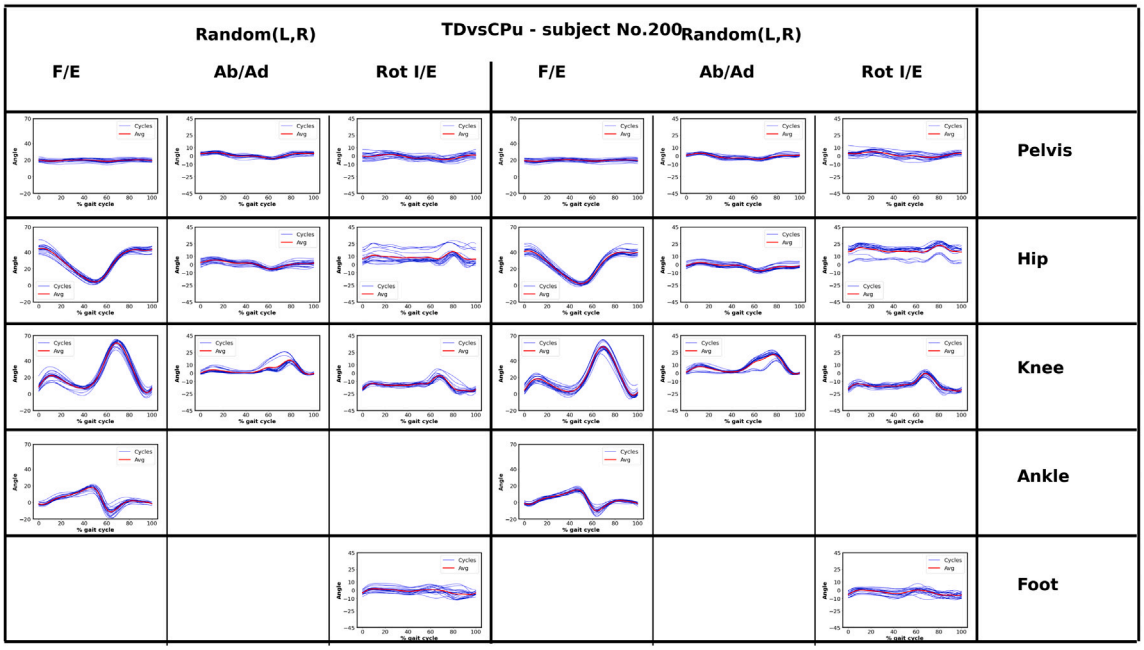


Fig. 6. Joint angles throughout the gait cycle of a TD subject.

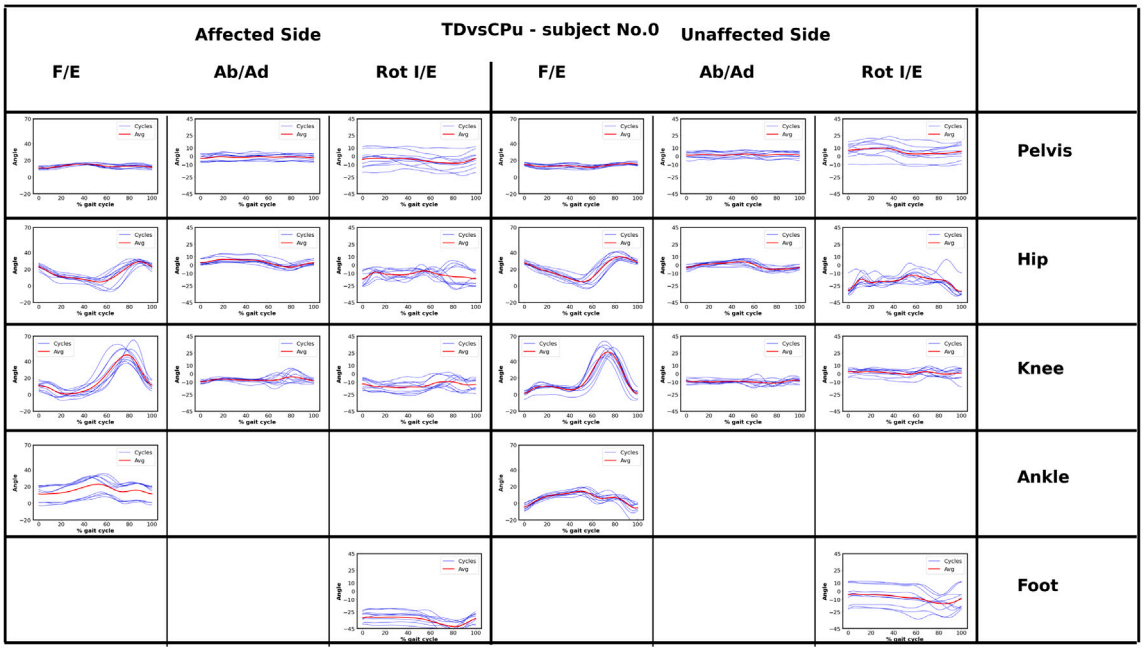


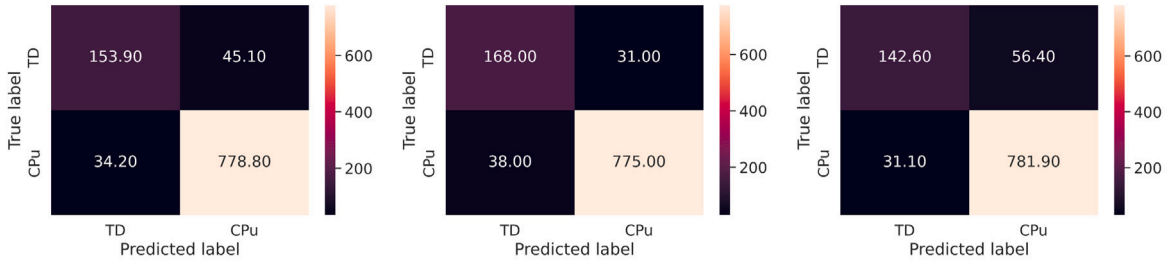
Fig. 7. Joint angles throughout the gait cycle of a CPu subject.

A.2. Data visualization

See Figs. 6 and 7.

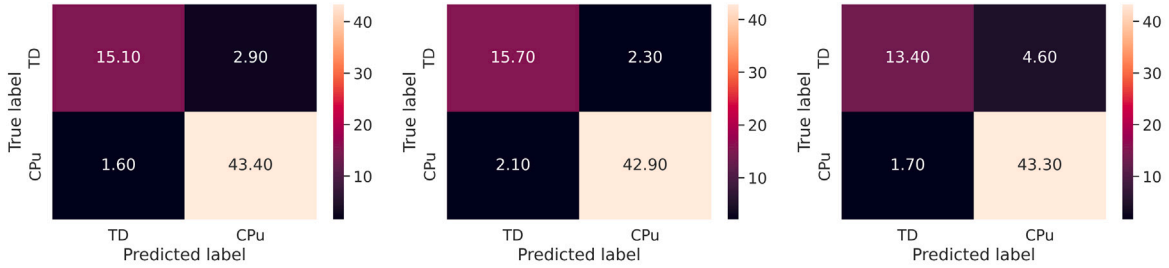
A.3. Confusion matrix

See Figs. 8–19.



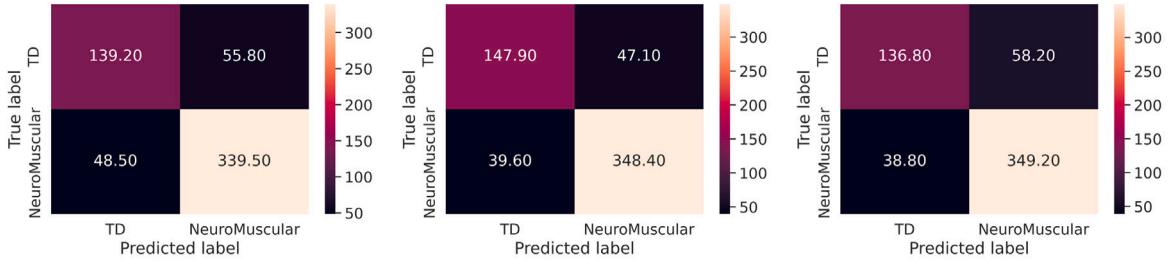
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 8. Confusion matrix of the prediction cycles of the three DL model for TDvsCPu.



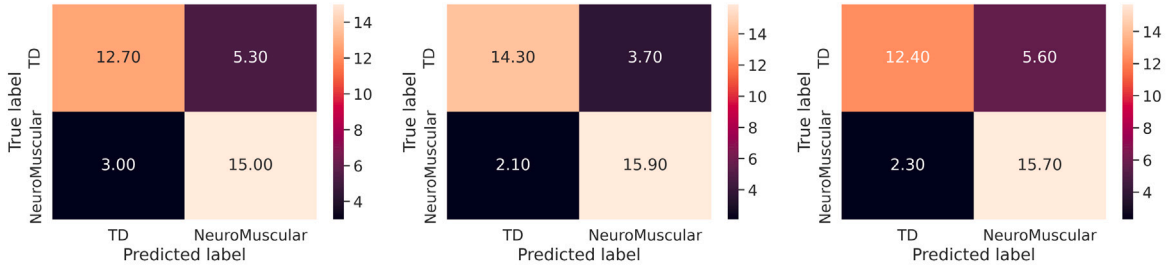
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 9. Confusion matrix of the prediction subjects of the three DL model for TDvsCPu.



(a) ResNet (b) LSTM (c) InceptionTime

Fig. 10. Confusion matrix of the prediction cycles of the three DL model for TDvsNM.

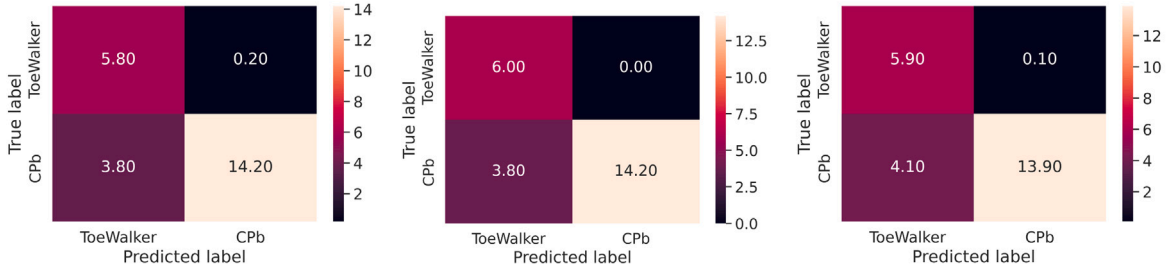


(a) ResNet (b) LSTM (c) InceptionTime

Fig. 11. Confusion matrix of the prediction subjects of the three DL model for TDvsNM.

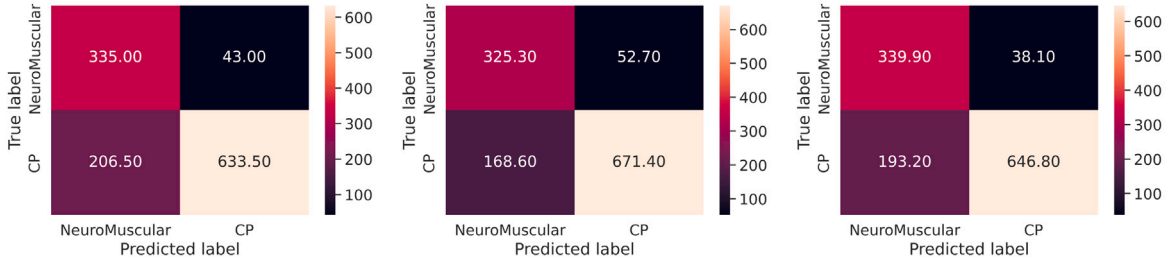
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 12. Confusion matrix of the prediction cycles of the three DL model for ToeWalkervsCPb.



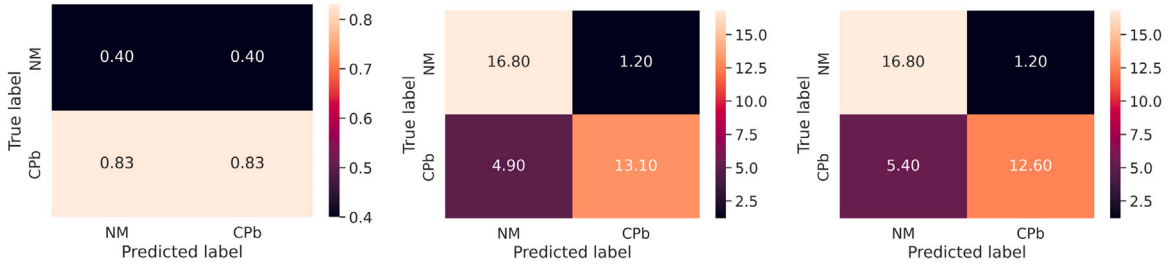
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 13. Confusion matrix of the prediction subjects of the three DL model for ToeWalkervsCPb.



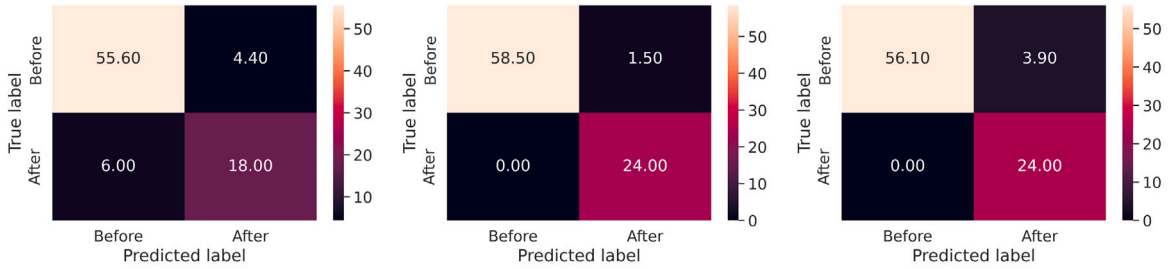
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 14. Confusion matrix of the prediction cycles of the three DL model for CPbvsNM.



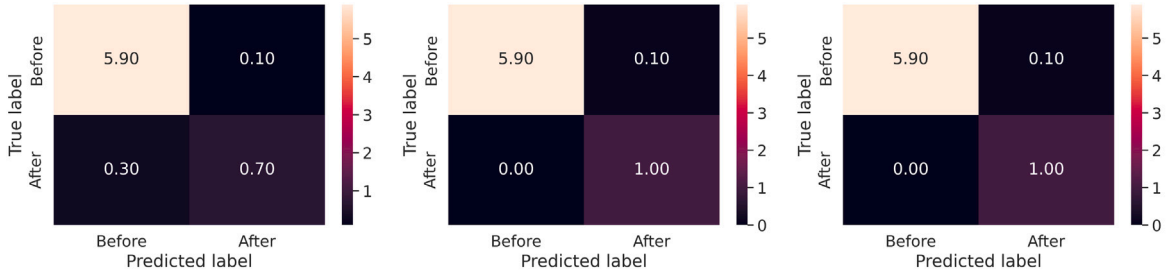
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 15. Confusion matrix of the prediction subjects of the three DL model for CPbvsNM.



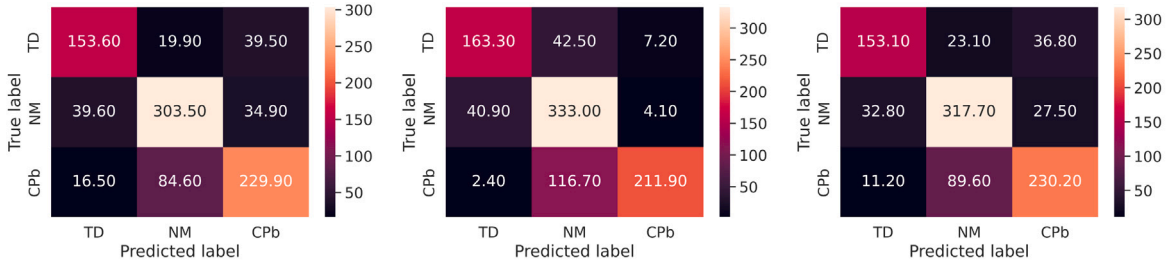
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 16. Confusion matrix of the prediction cycles of the three DL model for StrokeBeforeAfter.



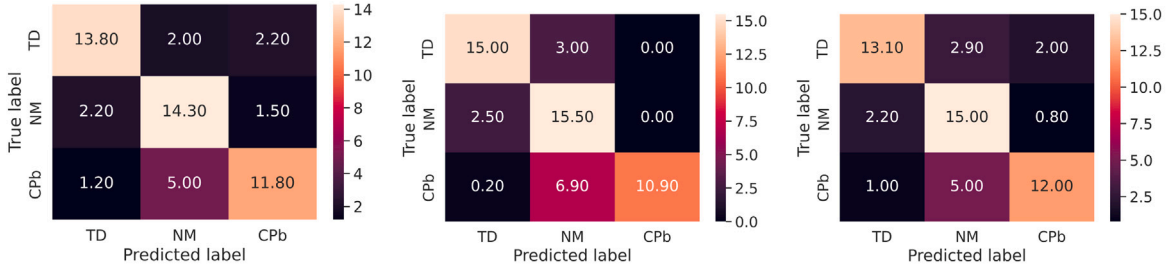
(a) ResNet (b) LSTM (c) InceptionTime

Fig. 17. Confusion matrix of the prediction subjects of the three DL model for StrokeBeforeAfter.



(a) ResNet (b) LSTM (c) InceptionTime

Fig. 18. Confusion matrix of the prediction cycles of the three DL model for TDvsCPbvsNM.



(a) ResNet (b) LSTM (c) InceptionTime

Fig. 19. Confusion matrix of the prediction subjects of the three DL model for TDvsCPbvsNM.

References

[1] P. Rosenbaum, N. Paneth, A. Leviton, M. Goldstein, M. Bax, A report: The definition and classification of cerebral palsy - april 2006, Dev. Med. Child Neurol. 49 (2007) 8–14.

[2] A. Michael-Asalu, G. Taylor, H. Campbell, L.-L. Lelea, R.S. Kirby, Cerebral palsy: Diagnosis, epidemiology, genetics, and clinical update, Adv. Pediatr. 66 (2019) 189–208.

[3] H. Stolze, S. Klebe, C. Baecker, C. Zechlin, L. Friege, S. Pohle, G. Deuschl, Prevalence of gait disorders in hospitalized neurological patients, Mov. Disorders 20 (1) (2005) 89–94.

[4] S. Lord, B. Galna, L. Rochester, Moving forward on gait measurement: toward a more refined approach, Mov. Disorders 28 (11) (2013) 1534–1543.

[5] R.Z.U. Rehman, S. Del Din, Y. Guan, A.J. Yarnall, J.Q. Shi, L. Rochester, Selecting clinically relevant gait characteristics for classification of early Parkinson's disease: A comprehensive machine learning approach, Sci. Rep. 9 (1) (2019) 17269.

[6] M. Yang, H. Zheng, H. Wang, S. McClean, J. Hall, N. Harris, A machine learning approach to assessing gait patterns for complex regional pain syndrome, Med. Eng. Phys. 34 (6) (2012) 740–746.

[7] A. Habersack, S.F. Fischerauer, T. Kraus, H.-P. Holzer, M. Svehlik, Kinematic and kinetic gait parameters can distinguish between idiopathic and neurologic toe-walking, Int. J. Environ. Res. Public Health 19 (2) (2022) 804.

[8] M.P. Kadaba, H.K. Ramakrishnan, M.E. Wootten, Measurement of lower extremity kinematics during level walking, J. Orthop. Res. Off. Publ. Orthop. Res. Soc. 8 (3) (1990) 383–392.

[9] I. Bautmans, B. Jansen, B. Van Keymolen, T. Mets, Reliability and clinical correlates of 3D-accelerometry based gait analysis outcomes according to age and fall-risk, Gait Posture 33 (3) (2011) 366–372.

[10] P. Putz, S. Durstberger, C. Kaufmann, M. Klinger, K. Plessl, J. Rejtö, K. Widhalm, C. Male, I. Pabinger, 3D gait analysis, haemophilia joint health score, leg muscle laterality and biomarkers of joint damage: A cross-sectional comparative assessment of haemophilic arthropathy, Haemophilia 26 (6) (2020) e323–e333.

[11] R. Rucco, V. Agosti, F. Jacini, P. Sorrentino, P. Varriale, M. De Stefano, G. Milan, P. Montella, G. Sorrentino, Spatio-temporal and kinematic gait analysis in patients with frontotemporal dementia and Alzheimer's disease through 3D motion capture, Gait Posture 52 (2017) 312–317.

[12] A. Romano, M. Favetta, T. Schirinzi, S. Summa, S. Minosse, A. D'Amico, M. Catteruccia, M. Petrarca, E. Castelli, E. Bertini, G. Vasco, Evaluation of gait in duchenne muscular dystrophy: Relation of 3D gait analysis to clinical assessment, Neuromuscular Disord. 29 (12) (2019) 920–929.

[13] M. Galli, V. Cimolin, V. Crugnola, L. Priano, F. Menegoni, C. Trotti, E. Milano, A. Mauro, Gait pattern in myotonic dystrophy (steinert disease): A kinematic, kinetic and EMG evaluation using 3D gait analysis, J. Neurol. Sci. 314 (1) (2012) 83–87.

[14] M. Liparoti, M. Della Corte, R. Rucco, P. Sorrentino, M. Sparaco, R. Capuano, R. Minino, L. Lavorgna, V. Agosti, G. Sorrentino, S. Bonavita, Gait abnormalities in minimally disabled people with multiple sclerosis: A 3D-motion analysis study, Multiple Scler. Relat. Disord. 29 (2019) 100–107.

[15] A. Fouasson-Chailloux, Y. Maugars, C. Vinatier, M. Trossaert, P. Menu, F. Rannou, J. Guicheux, M. Dauty, Clinical relevance of 3D gait analysis in patients with haemophilia, Haemophilia 24 (5) (2018) 703–710.

[16] E. Molteni, E. Beretta, D. Altomonte, F. Formica, S. Strazzer, Combined robotic-aided gait training and 3D gait analysis provide objective treatment and assessment of gait in children and adolescents with acquired hemiplegia, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBC, 2015, pp. 4566–4569, http://dx.doi.org/10.1109/EMBC.2015.7319410.

[17] P. De Blasiis, M.F. Siani, A. Fullin, M. Sansone, M.A.B. Melone, S. Sampaolo, E. Signoriello, G. Lus, Short and long term effects of nabiximols on balance and walking assessed by 3D-gait analysis in people with multiple sclerosis and spasticity., Multiple Scler. Relat. Disord. 51 (2021) 102805.

[18] B. Noehren, J. Scholz, I. Davis, The effect of real-time gait retraining on hip kinematics, pain and function in subjects with patellofemoral pain syndrome, Br. J. Sports Med. 45 (9) (2011) 691–696.

[19] E. Broström, S. Hagelberg, Y. Haglund-Åkerlind, Effect of joint injections in children with juvenile idiopathic arthritis: Evaluation by 3D-gait analysis, Acta Paediatr. 93 (7) (2004) 906–910.

[20] B.M. Eskofier, P. Federolf, P.F. Kugler, B.M. Nigg, Marker-based classification of Young–Elderly gait pattern differences via direct PCA feature extraction and SVMs, Comput. Methods Biomech. Biomed. Eng. 16 (4) (2013) 435–442.

[21] R. Begg, J. Kamruzzaman, A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data, J. Biomech. 38 (3) (2005) 401–408.

[22] J. Wu, J. Wang, L. Liu, Feature extraction via KPCA for classification of gait patterns, Hum. Mov. Sci. 26 (3) (2007) 393–411.

[23] Y. Zhou, R. Romijnders, C. Hansen, J. van Campen, W. Maetzler, T. Hortobágyi, C.J.C. Lamoth, The detection of age groups by dynamic gait outcomes using machine learning approaches, Sci. Rep. 10 (1) (2020) 4426.

[24] D. Laroche, A. Tolambiya, C. Morisset, J.F. Maillefert, R.M. French, P. Ornetti, E. Thomas, A classification study of kinematic gait trajectories in hip osteoarthritis, Comput. Biol. Med. 55 (2014) 42–48.

[25] R. Bajpai, A. Tiwari, D. Joshi, R. Khatavkar, Abnormnet: A neural network based suggestive tool for identifying gait abnormalities in cerebral palsy children, in: 2022 International Conference for Advancement in Technology, ICONAT, 2022, pp. 1–5, http://dx.doi.org/10.1109/ICONAT53423.2022.9725832.

[26] S. Aich, P.M. Pradhan, J. Park, H.C. Kim, A machine learning approach to distinguish Parkinson's disease (PD) patient's with shuffling gait from older adults based on gait signals using 3D motion analysis, Int. J. Eng. Technol. 7 (3.29) (2018) 153–156.

[27] A.S. Alharthi, S.U. Yunas, K.B. Ozanyan, Deep learning for monitoring of human gait: A review, IEEE Sens. J. 19 (21) (2019) 9575–9591.

[28] L. Medeiros, H. Almeida, L. Dias, M. Perkusich, R. Fischer, A gait analysis approach to track parkinson's disease evolution using principal component analysis, in: 2016 IEEE 29th International Symposium on Computer-Based Medical Systems, CBMS, 2016, pp. 48–53, http://dx.doi.org/10.1109/CBMS.2016.14.

[29] K. Chia, I. Fischer, P. Thomason, H.K. Graham, M. Sangeux, A decision support system to facilitate identification of musculoskeletal impairments and propose recommendations using gait analysis in children with cerebral palsy, Front. Bioeng. Biotechnol. 8 (2020).

[30] A.S. Alharthi, A.J. Casson, K.B. Ozanyan, Gait spatiotemporal signal analysis for Parkinson's disease detection and severity rating, IEEE Sens. J. 21 (2) (2021) 1838–1848.

[31] G. Park, K.M. Lee, S. Koo, Uniqueness of gait kinematics in a cohort study, Sci. Rep. 11 (2021) 15248.

[32] R. Begg, J. Kamruzzaman, A comparison of neural networks and support vector machines for recognizing Young-old gait patterns, in: Conference on Convergent Technologies for Asia-Pacific Region TENCON 2003, Vol. 1, 2003, pp. 354–358, http://dx.doi.org/10.1109/TENCON.2003.1273344.

[33] D.A. Winter, Biomechanics and Motor Control of Human Gait, second ed., Univ. of Waterloo Press, Waterloo, Ontario, 1991.

[34] R. Baker, J.L. McGinley, M.H. Schwartz, S. Beynon, A. Rozumalski, H.K. Graham, O. Tirosh, The gait profile score and movement analysis profile, Gait Posture 30 (3) (2009) 265–269.

[35] C. Rat, B. Tudrej, S. Kinouani, C. Guineberteau, P. Bertrand, V. Renard, O. Saint-Lary, Regulatory framework for research in general practice, Exerc. Revue Francophone Med. Gen. (135) (2017) 327–334.

[36] H.I. Fawaz, G. Forestier, J. Weber, L. Idoumghar, P.-A. Muller, Deep learning for time series classification: A review, Data Min. Knowl. Discov. 33 (4) (2019) 917–963, arXiv:1809.04356.

[37] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2015, http://dx.doi.org/10.48550/arXiv.1512.03385, arXiv:1512.03385.

[38] Z. Wang, W. Yan, T. Oates, Time series classification from scratch with deep neural networks: A strong baseline, 2016, http://dx.doi.org/10.48550/arXiv.1611.06455, arXiv:arXiv:1611.06455.

[39] S. Hochreiter, J. Schmidhuber, Long short-term memory, Neural Comput. 9 (8) (1997) 1735–1780.

[40] H.I. Fawaz, B. Lucas, G. Forestier, C. Pelletier, D.F. Schmidt, J. Weber, G.I. Webb, L. Idoumghar, P.-A. Muller, F. Petitjean, InceptionTime: Finding AlexNet for time series classification, Data Min. Knowl. Discov. 34 (6) (2020) 1936–1962, arXiv:1909.04939.

[41] S. Saadatnejad, M. Oveisi, M. Hashemi, LSTM-based ECG classification for continuous monitoring on personal wearable devices, IEEE J. Biomed. Health Inf. 24 (2) (2020) 515–523.

[42] P. Nagabushanam, S. Thomas George, S. Radha, EEG signal classification using LSTM and improved neural network algorithms, Soft Comput. 24 (13) (2020) 9981–10003.

[43] G. Klarenbeek, R.I.A. Harmanny, L. Cifola, Multi-target human gait classification using LSTM recurrent neural networks applied to micro-Doppler, in: 2017 European Radar Conference, EURAD, 2017, pp. 167–170, http://dx.doi.org/10.23919/EURAD.2017.8249173.

[44] M. Khokhlova, C. Migniot, A. Morozov, O. Sushkova, A. Dipanda, Normal and pathological gait classification LSTM model, Artif. Intell. Med. 94 (2019) 54–66.

[45] A. Rivadulla, X. Chen, G. Weir, D. Cazzola, G. Trewartha, J. Hamill, E. Preatoni, Development and validation of FootNet; a new kinematic algorithm to improve foot-strike and toe-off detection in treadmill running, PLoS One 16 (8) (2021) e0248608.

[46] W. Falcon, The PyTorch Lightning team, PyTorch Lightning, 2019, http://dx.doi.org/10.5281/zenodo.3828935, URL: https://github.com/Lightning-AI/lightning.

[47] R. Bajpai, D. Joshi, A-GAS: a probabilistic approach for generating automated gait assessment score for cerebral palsy children, IEEE Trans. Neural Syst. Rehabil. Eng. 29 (2021) 2530–2539.

[48] M.T. Ribeiro, S. Singh, C. Guestrin, "Why should I trust you?": Explaining the predictions of any classifier, 2016, http://dx.doi.org/10.48550/arXiv.1602.04938, arXiv:arXiv:1602.04938.

[49] S. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, 2017, http://dx.doi.org/10.48550/arXiv.1705.07874, arXiv:arXiv:1705.07874.

[50] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: visual explanations from deep networks via gradient-based localization, in: 2017 IEEE International Conference on Computer Vision, ICCV, 2017, pp. 618–626, http://dx.doi.org/10.1109/ICCV.2017.74.

[51] S. Bach, A. Binder, G. Montavon, F. Klauschen, K.-R. Müller, W. Samek, On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation, PLoS One 10 (7) (2015) e0130140.

[52] P. Soda, A. Carta, D. Formica, E. Guglielmelli, A low-cost video-based tool for clinical gait analysis, in: 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2009, pp. 3979–3982, http://dx.doi.org/10.1109/IEMBS.2009.5333623.

[53] J.P. Ferreira, T. Liu, P. Coimbra, P. Coimbra, Parameter analysis and selection for human gait characterization using a low cost vision system, in: 2017 4th International Conference on Systems and Informatics, ICSAI, 2017, pp. 198–203, http://dx.doi.org/10.1109/ICSAI.2017.8248289.

[54] A. Zanela, T. Schirinzi, N.B. Mercuri, A. Stefani, C. Romagnoli, G. Annino, V. Bonaiuto, R. Cerroni, Using a video device and a deep learning-based pose estimator to assess gait impairment in neurodegenerative related disorders: A pilot study, Appl. Sci. 12 (9) (2022) 4642.

[55] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, Y. Sheikh, OpenPose: Realtime multi-person 2D pose estimation using part affinity fields, 2019, http://dx.doi.org/10.48550/arXiv.1812.08008, arXiv:arXiv:1812.08008.

[56] H.-S. Fang, J. Li, H. Tang, C. Xu, H. Zhu, Y. Xiu, Y.-L. Li, C. Lu, AlphaPose: Whole-body regional multi-person pose estimation and tracking in real-time, IEEE Trans. Pattern Anal. Mach. Intell. (2022) 1–17.

[57] Y.-m. Tang, Y.-h. Wang, X.-y. Feng, Q.-s. Zou, Q. Wang, J. Ding, R.C.-j. Shi, X. Wang, Diagnostic value of a vision-based intelligent gait analyzer in screening for gait abnormalities, Gait Posture 91 (2022) 205–211.

[58] T. Connie, T.B. Aderinola, T.S. Ong, M.K.O. Goh, B. Erfianto, B. Purnama, Pose-based gait analysis for diagnosis of Parkinson's disease, Algorithms 15 (12) (2022) 474.

[59] W. Li, X. Chen, J. Zhang, J. Lu, C. Zhang, H. Bai, J. Liang, J. Wang, H. Du, G. Xue, Y. Ling, K. Ren, W. Zou, C. Chen, M. Li, Z. Chen, H. Zou, Recognition of freezing of gait in Parkinson's disease based on machine vision, Front. Aging Neurosci. 14 (2022) 921081.

[60] H. Zhang, E.S.L. Ho, X. Zhang, H.P.H. Shum, Pose-based tremor classification for Parkinson's disease diagnosis from video, in: L. Wang, Q. Dou, P.T. Fletcher, S. Speidel, S. Li (Eds.), Medical Image Computing and Computer Assisted Intervention, MICCAI 2022, in: Lecture Notes in Computer Science, Springer Nature Switzerland, Cham, 2022, pp. 489–499, http://dx.doi.org/10.1007/978-3-031-16440-8_47.

[61] T. Ahmed, M.K. Islam, EMG signal classification for detecting neuromuscular disorders, J. Phys. Conf. Ser. 1921 (1) (2021) 012043.

[62] A. Subasi, E. Yaman, EMG signal classification using discrete wavelet transform and rotation forest, in: A. Badnjevic, R. Škrbić, L. Gurbeta Pokvić (Eds.), CMBEBIH 2019, in: IFMBE Proceedings, Springer International Publishing, Cham, 2020, pp. 29–35, http://dx.doi.org/10.1007/978-3-030-17971-7_5.

[63] A. Belkhou, A. Achmamad, A. Jbari, Classification and diagnosis of myopathy EMG signals using the continuous wavelet transform, in: 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science, EBBT, 2019, pp. 1–4, http://dx.doi.org/10.1109/EBBT.2019.8742051.

[64] V. Kehri, A.R. N, EMG signal analysis for diagnosis of muscular dystrophy using wavelet transform, SVM and ANN, Biomed. Pharmacol. J. 11 (3) (2018) 1583–1591.

[65] J.R. Torres-Castillo, C.O. López-López, M.A. Padilla-Castañeda, Neuromuscular disord. detection through time-frequency analysis and classification of multi-muscular EMG signals using Hilbert-Huang transform, Biomed. Signal Process. Control 71 (2022) 103037.

[66] R. Dubey, M. Kumar, A. Upadhyay, R.B. Pachori, Automated diagnosis of muscle diseases from EMG signals using empirical mode decomposition based method, Biomed. Signal Process. Control 71 (2022) 103098.

[67] D.C. Toledo-Pérez, J. Rodríguez-Reséndiz, R.A. Gómez-Loenzo, J.C. Jauregui-Correa, Support vector machine-based EMG signal classification techniques: a review, Appl. Sci. 9 (20) (2019) 4402.