



OPEN

## Quantitative gait analysis and prediction using artificial intelligence for patients with gait disorders

Nawel Ben Chaabane<sup>1,2,✉</sup>, Pierre-Henri Conze<sup>1,3</sup>, Mathieu Lempereur<sup>1,2,4</sup>, Gwénoél Quéllec<sup>1</sup>, Olivier Rémy-Néris<sup>1,2,4</sup>, Sylvain Brochard<sup>1,2,4</sup>, Béatrice Cochener<sup>1,2,4</sup> & Mathieu Lamard<sup>1,2</sup>

Quantitative Gait Analysis (QGA) is considered as an objective measure of gait performance. In this study, we aim at designing an artificial intelligence that can efficiently predict the progression of gait quality using kinematic data obtained from QGA. For this purpose, a gait database collected from 734 patients with gait disorders is used. As the patient walks, kinematic data is collected during the gait session. This data is processed to generate the Gait Profile Score (GPS) for each gait cycle. Tracking potential GPS variations enables detecting changes in gait quality. In this regard, our work is driven by predicting such future variations. Two approaches were considered: signal-based and image-based. The signal-based one uses raw gait cycles, while the image-based one employs a two-dimensional Fast Fourier Transform (2D FFT) representation of gait cycles. Several architectures were developed, and the obtained Area Under the Curve (AUC) was above 0.72 for both approaches. To the best of our knowledge, our study is the first to apply neural networks for gait prediction tasks.

Gait disorders are described as any deviation from normal walking or gait<sup>1</sup>. Their prevalence among adults rises with age. In the elderly population over the age of 70 years, they are present in approximately 35% of patients<sup>2,3</sup> and in 72% of patients over 80 years<sup>2</sup>. These statistics take into account whether such disorders result from neurological etiologies or not, which can be determined through laboratory work, clinical presentation, and diagnostic testing<sup>2</sup>. In fact, gait disorders etiologies include neurological conditions (e.g., sensory or motor impairments), orthopedic abnormalities (e.g., osteoarthritis and skeletal deformities), and medical conditions (e.g., heart failure, respiratory insufficiency, peripheral arterial occlusive disease, obesity)<sup>4,5</sup>. Cerebral palsy, as a group of neurological disorders, affects about 2 in every 1000 newborns. Its prevalence reaches 5–8% among newborns with very low birth weights or very pre-term deliveries. Gait disturbances have a tremendous impact on patients, especially on their quality of life<sup>1</sup>: they complain most often of pain, joint stiffness, numbness, or weakness<sup>6</sup>. Neurological gait disorders, in particular, are associated with lower cognitive function, depressed mood, and diminished quality of life<sup>7</sup>. To have insight into patients' conditions and therefore treat their gait disorders, clinicians historically used Observational Gait Analysis (OGA)<sup>8</sup>. OGA usually relies on a clinician's observation freeze-framed techniques and video slow-motion replay to record and analyze a patient's gait. It is subject to bias and has limited precision because it relies on the experience of the clinician. To overcome this limitation, Quantitative Gait Analysis (QGA) is considered. It uses instrumentation to quantify the gait cycle by recording temporal-spatial, kinematic, and kinetic data that is rarely gathered by observation. The challenge facing clinicians is to analyze a large amount of clinical data from QGA in order to determine the severity of the illness and select the most effective therapeutic strategy. It is a very tricky task because of the great disparity between patients (e.g., children and adults) and the diversity of their pathologies. In this context, our aim is to assist clinicians in analyzing this large amount of clinical data with an artificial intelligence applied to kinematics from QGA. The target objective is to go beyond objectively quantifying gait quality by predicting whether it will improve within the next visit. These predictions tend to help clinicians select the most effective treatment strategy. For this purpose, two approaches were considered: signal-based, which uses raw gait cycles, and image-based, which converts gait cycles into image-like representations, making them suitable for training image-based deep neural networks, especially pre-trained ones. In the signal-based approach, a Long Short Term Memory (LSTM) and a MultiLayer Perceptron (MLP) were designed from scratch. Their hyper-parameters were

<sup>1</sup>LaTIM UMR 1101 Laboratory, Inserm, Brest, France. <sup>2</sup>Western Brittany University, Brest, France. <sup>3</sup>IMT Atlantique, Brest, France. <sup>4</sup>University Hospital of Brest, Brest, France. ✉email: nawel.ben-chaabane@inserm.fr

tuned with KerasTuner<sup>9</sup>. The obtained results were compared to five state-of-the-art architectures<sup>10</sup>, including Fully Convolutional neural Network (FCN), Residual Network (ResNet), Encoder, Time Le-Net (t-LeNet), and Transformer. For the two tailored and state-of-the-art architectures, the influence of data augmentation was studied. In the image-based approach, the first step was to map the time representation of 1D gait cycles to a 2D frequency representation using the two-dimensional Fast Fourier Transform (2D FFT). Then, the obtained 2D FFT images were processed with four pre-trained Convolutional Neural Networks (CNN): VGG16, ResNet34, EfficientNet\_b0, and a Vision Transformer (ViT). The obtained results were compared to those of a tailored CNN with a much smaller number of parameters. The effectiveness of the proposed models was evaluated on a gait dataset collected from more than 700 patients.

## Materials and methods

### Data acquisition

This study was carried out in accordance with the tenets of the Declaration of Helsinki and with the approval of the Brest, France hospital's (CHRU's) Ethics Committee. Patients had also signed an informed consent. Our work was conducted between 2021 and 2022. Data collected between June 2006 and June 2021 from 734 patients (115 adults and 619 children) who had undergone clinical 3D gait analysis were used. Their identities were preserved by respecting medical secret and protecting patient confidentiality. All data were recorded using the same motion analysis system (Vicon MX, Oxford Metrics, UK) and four force platforms (Advanced Mechanical Technology, Inc., Watertown, MA, USA) in the same motion laboratory (CHU Brest) between 2006 and 2022. The data collected by the 15 infrared cameras (sampling rate of 100 or 120 Hz) were synchronized with the ground reaction forces recorded by the force platforms (1000 Hz or 1200 Hz). The 16 markers were placed according to the protocol by Kadaba et al.<sup>11</sup>. Marker trajectories and ground reaction forces were dual-pass filtered with a low-pass Butterworth filter at a cut-off frequency of 6 Hz. After an initial calibration in the standing position, all patients were asked to walk at a self-selected speed along a 10m walkway.

Gait kinematics were processed using the Vicon Plug-in Gait model. Kinematics were time-normalized to stride duration, from 0 to 100% from initial contact (IC) to the next IC of the ipsilateral foot. Nine gait joint angles (kinematic gait variables) were used: anteversion/retroversion of the pelvis, rotation of the pelvis, pelvic tilt, flexion/extension of the hip, abduction/adduction of the hip, internal/external rotation of the hip, flexion/extension of the knee, plantar/dorsiflexion of the ankle, and the foot's angle of progression. As a result, a gait cycle yielded  $101 \times 9$  measurements. Let  $E_{p,d}$  denote the gait session of patient  $p$  at datetime  $d$ . It can be written as follows:

$$E_{p,d} = \{C_{E_{p,d}}^1, C_{E_{p,d}}^2, \dots, C_{E_{p,d}}^K\} \quad (1)$$

where  $C_{E_{p,d}}^k$  is the  $k$ -th gait cycle of a gait session  $E_{p,d}$  and  $K$  the total number of gait cycles. Let  $c_{t,n}^{E_{p,d}^k}$  denote the gait cycle  $C_{E_{p,d}}^k$  value at time step  $t$  and joint angle  $n$ . To keep notations simple,  $c_{t,n}^{E_{p,d}^k}$  is referred to as  $c_{t,n}$  in what follows.  $C_{E_{p,d}}^k$  can simply be represented with a matrix of 101 lines and 9 columns, as follows:

$$C_{E_{p,d}}^k = \begin{bmatrix} c_{1,1} & c_{1,2} & \dots & c_{1,9} \\ c_{2,1} & c_{2,2} & \dots & c_{2,9} \\ \vdots & & & \\ c_{101,1} & c_{101,2} & \dots & c_{101,9} \end{bmatrix} \quad (2)$$

The Gait Profile Score (GPS), a “walking behavior score”, was computed for each gait cycle from the previously described joint angles<sup>12–14</sup>. The GPS is a single index measure that summarizes the overall deviation of kinematic gait data relative to normative data. It can be decomposed to provide Gait Variable Scores (GVS) for nine key component kinematic gait variables, which are presented as a Movement Analysis Profile (MAP). The GVS corresponding to the  $n$ -th kinematic variable,  $GVS_n$ , is given by<sup>15–17</sup>:

$$GVS_n = \sqrt{\frac{1}{T} \sum_{t=1}^T (c_{t,n} - c_{t,n}^{ref})^2} \quad (3)$$

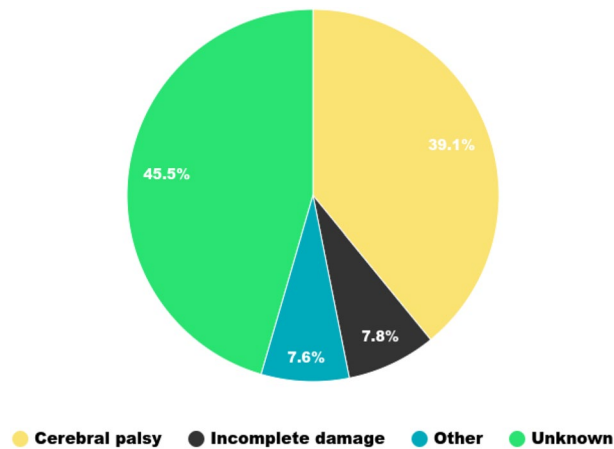
where  $t$  is a specific point in the gait cycle,  $T$  its total number of points (typically equal to 101<sup>18,19</sup>),  $c_{t,n}$  the value of the kinematic variable  $n$  at point  $t$ , and  $c_{t,n}^{ref}$  is its mean on the reference population (physiological normative). The GPS is obtained from the GVS scores<sup>15,17</sup> as follows:

$$GPS = \sqrt{\frac{1}{N} \sum_{n=1}^N GVS_n^2} \quad (4)$$

where  $N$  is the total number of kinematic variables (equal to 9 by definition).

### Gait database

We had a total of 1459 gait sessions from 734 patients (115 adults and 619 children). Each patient had an average of 1.988 gait sessions with a standard deviation of 1.515. 53,693 gait cycles were collected. Their average number per gait session is equal to 18 with a standard deviation of 6. Neurological conditions, notably cerebral palsy, are the most frequent etiologies, as we can see in Fig. 1.



**Figure 1.** Etiologies pie chart.

The average patient age within the first gait session is equal to 14 years, with a standard deviation of 16 years. The time delay between the first and last gait session (for patients with more than one gait session, i.e., 319) is equal to 3.92 years on average with a standard deviation of 3.24 years. Directly consecutive gait sessions are, on average, separated by approximately 740 days, with a standard deviation of 577 days. The shortest (resp. longest) time delay was equal to 4 (resp. 4438) days. We had 1384 pairs of directly consecutive gait sessions belonging to 319 patients (the remaining patients were removed since they had only one gait session). Involved gait conditions are various: without any equipment, with a cane, with a rollator, with an orthosis, with a prosthesis.. Only pairs of gait sessions without equipment were selected in order to be in the same condition (79% of all available pairs, i.e. 1152). The first gait sessions in these pairs were used for training. Models were fed the gait cycles of these first gait sessions (i.e., 21,167 gait cycles in total).

### GPS variation prediction

GPS variation prediction is similar enough to a Time Series Classification (TSC) issue that its proposed popular architectures should be adopted. Consecutive gait session pairs  $(E_{p,d}, E_{p,d+\Delta d})$  were considered. For each gait cycle  $C_{E_{p,d}}^k$  of the current gait session  $E_{p,d}$ , a GPS variation  $\Delta GPS$  was computed using:

$$\Delta GPS(C_{E_{p,d}}^k) = GPS_{avg}(E_{p,d+\Delta d}) - GPS(C_{E_{p,d}}^k) \quad (5)$$

where  $GPS_{avg}(E_{p,d+\Delta d})$  is the average GPS per cycle of  $E_{p,d+\Delta d}$  and  $GPS(C_{E_{p,d}}^k)$  the GPS of the current gait cycle  $C_{E_{p,d}}^k$ . The average GPS per cycle  $GPS_{average}(E_{p,d})$  of a gait session  $E_{p,d}$  is simply equal to:

$$GPS_{avg}(E_{p,d}) = \frac{\sum_{k=1}^K GPS(C_{E_{p,d}}^k)}{K} \quad (6)$$

$\Delta$

GPS was ranked in a binary fashion. Either it is negative, in which case the patient's gait improves (class 1), or it is positive, in which case the patient's gait worsens (class 0). The metric used is the Area Under the Curve (AUC).

The distribution of patients between training, validation, and test groups is provided in Table 1. Such a split put 73%, 12%, and 14% of total gait cycles within the training, validation, and test groups, respectively.

### Signal-based approach

To be exhaustive, one MLP, one recurrent neural network (LSTM), one hybrid architecture (Encoder), several CNN architectures (FCN, ResNet, t-LeNet), and a one-dimensional Transformer<sup>20</sup> were included. The MLP and LSTM were designed and developed from scratch. Their hyper-parameters were optimized manually. FCN, ResNet, Encoder, and t-LeNet are among the most effective end-to-end discriminative architectures regarding

Train	Validation	Test
224 patients	48 patients	47 patients
15509 cycles	2678 cycles	2980 cycles
844 gait session pairs	142 gait sessions pairs	166 gait session pairs
45.61 % of class 0	45.89 % of class 0	44.4% of class 0

**Table 1.** Data distribution for  $\Delta GPS$  prediction.

the TSC state-of-the-art<sup>10</sup>. These methods were also compared to the Transformer, a more recent and popular architecture. The Transformer does not suffer from long-range context dependency issues compared to LSTM<sup>21</sup>. In addition, it is notable for requiring less training. The Adam optimizer<sup>22</sup> and binary cross-entropy loss were employed<sup>23</sup>.

For MLP, gait cycles were flattened so that the input length was equal to 909 time steps. The number of neurons was the same across all the fully connected layers. Many values of this number were tested to find the best structure for our task. In the same way, the number of layers was optimized. The corresponding architecture is shown in Fig. 2.

LSTM layers were stacked, and a dropout was added before the last layer to avoid overfitting. The corresponding architecture is shown in Fig. 3.

For FCN, ResNet, Encoder and t-LeNet, the architectures proposed in Ref.<sup>10</sup> were considered. They are shown in Figs. 4, 5, 6 and 7, respectively. We followed an existing implementation<sup>24</sup> to set up the Transformer.

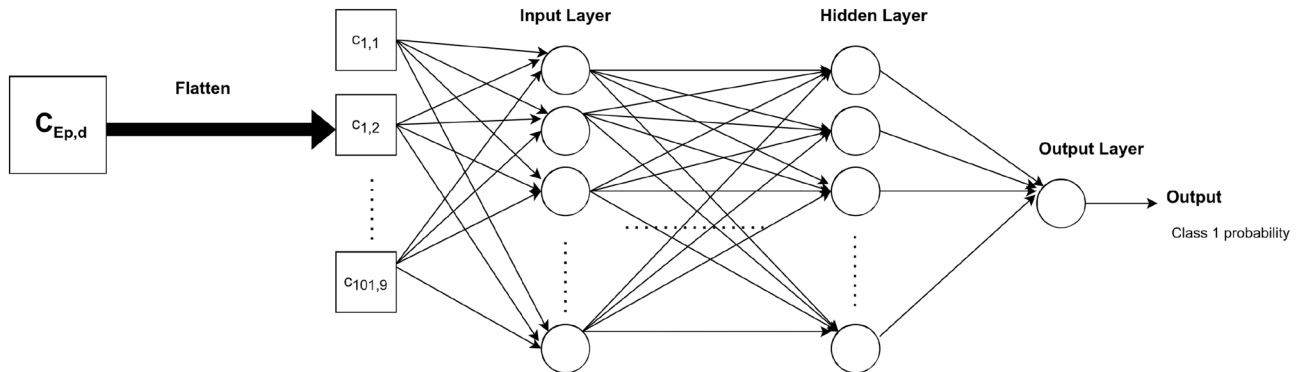


Figure 2. MLP architecture for prediction.

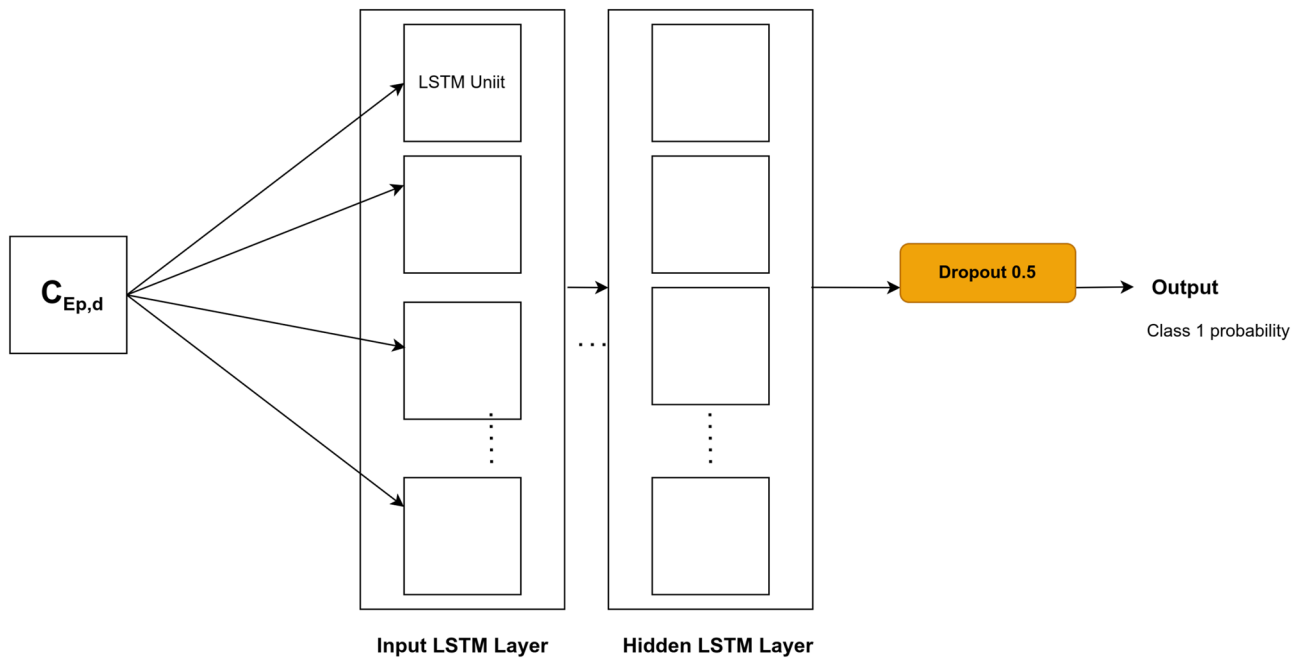


Figure 3. LSTM architecture for prediction.

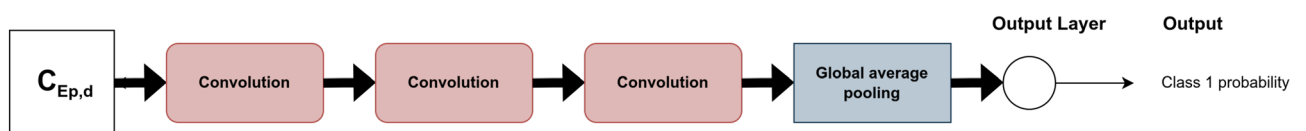


Figure 4. FCN architecture for prediction.

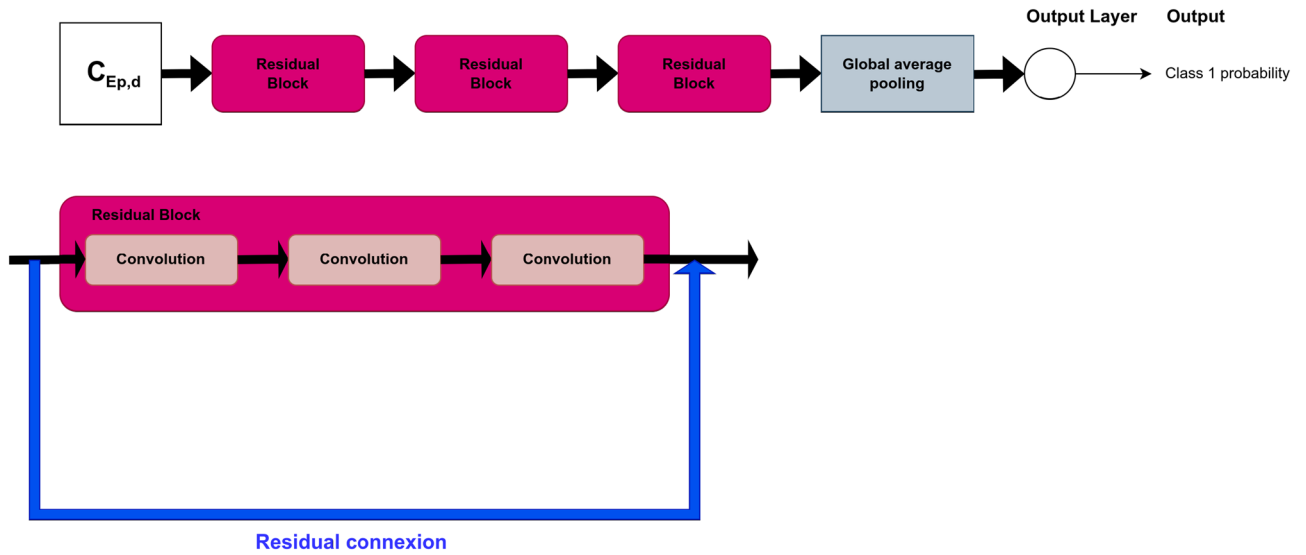


Figure 5. ResNet architecture for prediction.

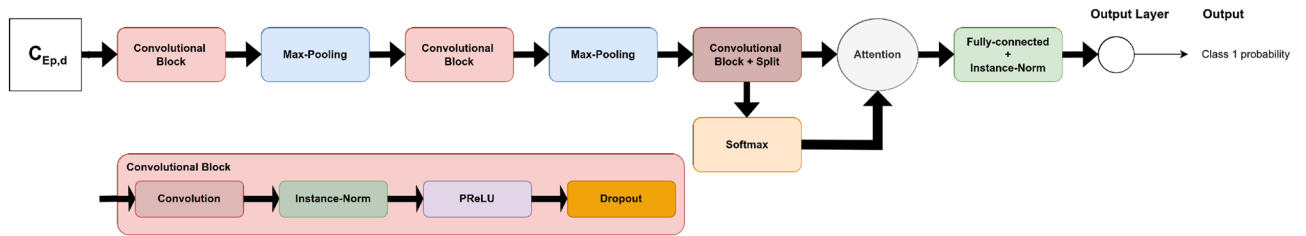


Figure 6. Encoder architecture for prediction.

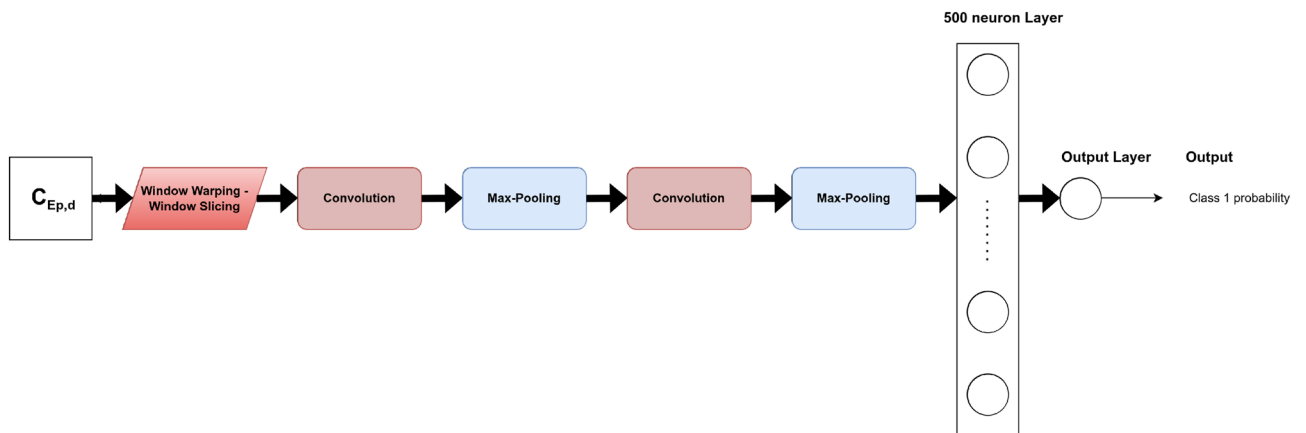


Figure 7. t-LeNet architecture for prediction.

**Data augmentation**

Different techniques of data augmentation were tested as a pre-processing step to avoid overfitting: jittering, scaling, window warping, permutation, and window slicing. Their hyperparameters were empirically optimized for each model. These are among the TSC literature’s most frequently utilized techniques, particularly when it comes from sensor data<sup>10</sup>.

*Image-based approach*

Image-based time series representation initiated a new branch of deep learning approaches that consider image transformation as an innovative pre-processing of feature engineering<sup>25</sup>. In an attempt to reveal features and patterns less visible in the one-dimensional sequence of the original time series, many transformation methods were developed to encode time series as input images.

In our study, sensor modalities are transformed to the visual domain using 2D FFT in order to utilize a set of pre-trained CNN models for transfer learning on the converted imagery data. The full workflow of our framework is represented in Fig. 8.

2D FFT is used to work in the frequency domain or Fourier domain because it efficiently extracts features based on the frequency of each time step in the time series. It can be defined as:

$$F(u, v) = \frac{1}{T \cdot N} \sum_{t=0}^T \sum_{n=0}^N c_{t,n} \exp \left( -j2\pi \left( \frac{ut}{T} + \frac{vn}{N} \right) \right) \quad (7)$$

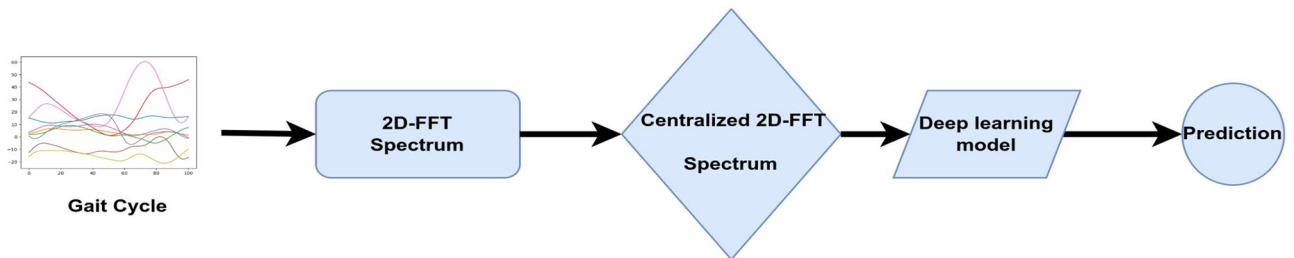
where  $F(u, v)$  is the direct Fourier transform of the gait cycle. It is a complex function that shows the phase and magnitude of the signal in the frequency domain.  $u$  and  $v$  are the frequency space coordinates. The magnitude of the 2D FFT  $|F(u, v)|$ , also known as the spectrum, is a two-dimensional signal that represents frequency information. Because the 2D FFT has translation and rotation attributes, the zero-frequency component can be moved to the center of  $|F(u, v)|$  without losing any information, making the spectrum image more visible. The centralized FFT spectrums were computed and fed to the proposed deep learning models. A centralized FFT spectrum for a given gait cycle is represented in Fig. 9.

## Proposed deep learning models

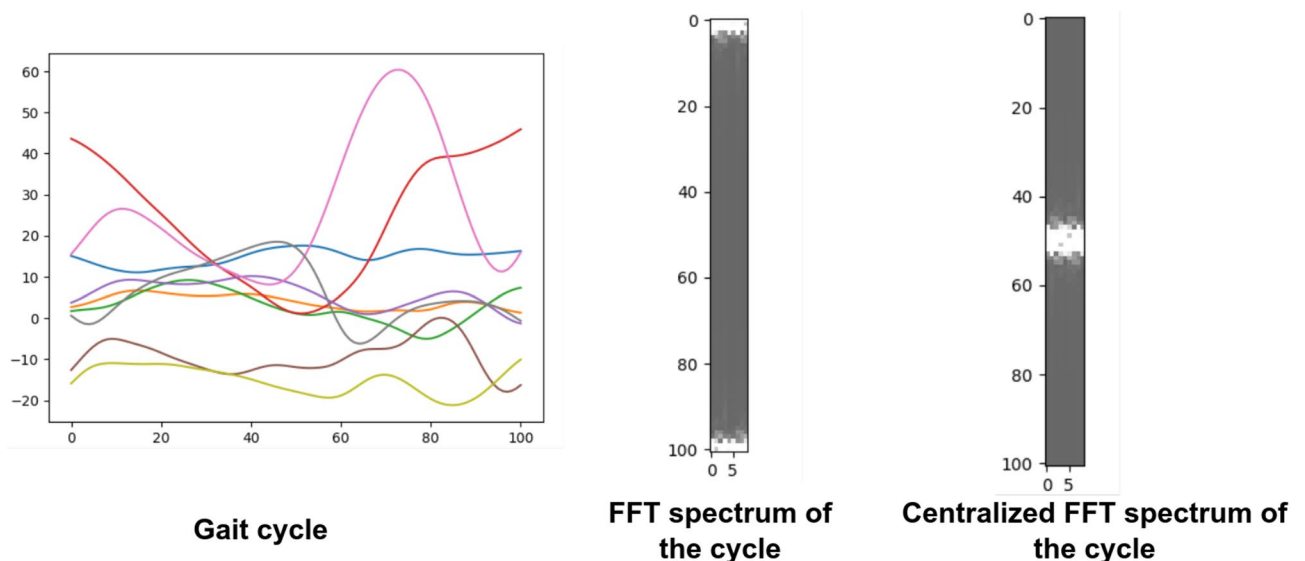
### Timm pre-trained models

The Timm library's<sup>26</sup> pre-trained VGG16, ResNet34, EfficientNet\_b0, and the Vision Transformer 'vit\_base\_patch16\_224' were investigated. They were pre-trained on a large collection of images, in a supervised fashion. For the Transformer, the pre-training was at a resolution of  $224 \times 224$  pixels. Its input images were considered as a sequence of fixed-size patches (resolution  $16 \times 16$ ), which were linearly embedded.

Converting our grayscale images to RGB images was not necessary because Timm's implementations support any number of input channels. The model's minimum input size for VGG16 is  $32 \times 32$ . The image's width dimension ( $N$ ) equals 9, which is less than 32. In order to fit the minimum needed size, 2D FFT images were repeated 4 times in this width dimension. Transfer learning with fine-tuning methods was employed. One neuron's final fully connected layer was used. In the same way that the top layers were trainable, all convolutional blocks were.



**Figure 8.** Proposed  $\Delta$ GPS prediction workflow for the image-based approach.



**Figure 9.** 2D FFT for a given gait cycle. (a) The gait cycle; (b) FFT spectrum of the gait cycle; (c) Centralized FFT spectrum of the gait cycle.

### Two-dimensional 2D CNN

The pre-trained Timm models are deep and sophisticated, with many layers. As a result, a CNN model with fewer parameters, designed from scratch, was conceived. The number of used two-dimensional convolutional layers was a hyper-parameter to optimize in a finite range of values {1, 2, 3, 4, 5}. After the convolutional block, a dropout function was applied. Following that, two-dimensional max-pooling (MaxPooling2D) and batch normalization were used. The flattened output of the batch normalization was then fed to a dense layer of a certain number of neurons to tune. In order to predict the  $\Delta GPS$ , our model had a dense output layer with a single neuron. The corresponding architecture is shown in Fig. 10.

The following are all of the architecture hyper-parameters to tune: the number of convolutional layers (num\_layers), the number of filters for each convolution layer (num\_filters), the kernel size of each convolution layer (kernel\_size), the dropout rate (dropout), the pooling size of the MaxPooling2D (pool\_size), the number of neurons in the dense layer (units), and the learning rate (lr). Five models with a varying number of convolutional layers (from 1 to 5) were tested. For each of them, the rest of the hyper-parameters were tuned using KerasTuner<sup>9</sup> to maximize the validation AUC.

## Results

In this section, prediction results are presented in terms of AUC.

### Signal-based approach

#### Without data augmentation

Results are given in Table 2. They are homogeneous on the validation set. LSTM and MLP perform equally well on the validation set. ResNet has the highest val AUC (0.709) for the state-of-the-art architectures. FCN achieves a comparable result to ResNet with a val AUC of 0.705. Encoder, t-LeNet and Transformer perform nearly equally well, with a val AUC above 0.63.

#### MLP

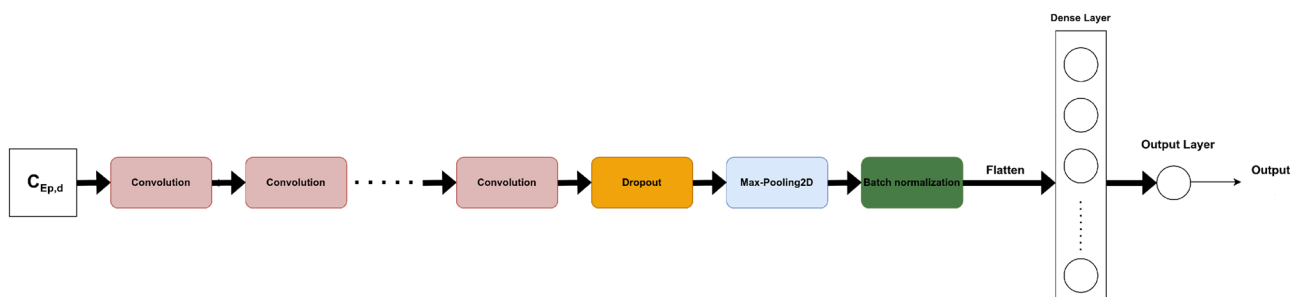
The best model has 4 layers of 200 neurons each. It is referred to as MLP\_4\_200. It gives a val AUC equal to 0.717.

#### LSTM

The best model has 4 LSTM layers of 500 units. It is referred to as LSTM\_4\_500. It gives a val AUC of 0.701.

#### Data augmentation

For all the architectures used, an overfitting behavior with very quick convergence was exhibited. To mitigate this, 5 data augmentation techniques already presented were combined. The order of their application was chosen randomly for each training batch. The best data augmentation parameters were found for each architecture. Results are presented in Table 3. Performances are slightly better after data augmentation except for the Transformer. In general, convergence is slower; it no longer appears in the first few epochs. FCN gives the best val AUC (0.723).



**Figure 10.** Tailored 2D CNN for prediction.

Model	Epoch of convergence	Val AUC	Test AUC	Test accuracy
MLP	55	<b>0.717</b>	<b>0.683</b>	<b>0.640</b>
LSTM	3	0.701	0.710	0.665
FCN <sup>10</sup>	30	0.705	0.689	0.641
ResNet <sup>10</sup>	8	0.709	0.699	0.663
Encoder <sup>10</sup>	25	0.631	0.691	0.648
t-LeNet <sup>10</sup>	5	0.653	0.689	0.64
Transformer <sup>20</sup>	851	0.640	0.709	0.626

**Table 2.** GPS variation prediction results. Best results in bold.



Model	Epoch of convergence	val AUC	test AUC	Test accuracy
MLP_4_200	2932	0.667	0.667	0.643
LSTM_4_500	2126	0.712	0.710	0.66
FCN <sup>10</sup>	1517	<b>0.723</b>	<b>0.717</b>	<b>0.598</b>
ResNet <sup>10</sup>	172	0.719	0.716	0.663
Encoder <sup>10</sup>	1942	0.640	0.632	0.604
t-LeNet <sup>10</sup>	61	0.625	0.665	0.618
Transformer <sup>30</sup>	322	0.546	0.561	0.467

**Table 3.** Prediction results after data augmentation. Significant values are in bold.

### Image-based approach

Table 4 presents the obtained results.

Despite the overfitting behavior of the pre-trained models, the test set's results are nearly identical to those from the validation set. The tested Timm models all produced results that were comparable, with a `val AUC` of greater than 0.63. The model with the highest efficiency, the CNN trained entirely from scratch, gives a `val AUC` of 0.726. It has two convolutional layers, and its hyper-parameter values are as follows: `num_filters = 4`, `kernel_size = 32`, `dropout = 0`, `pool_size = 8`, `units = 300` and `lr = 4.127 × 10-4`. Transfer learning is unlikely to have made a significant contribution because our images are visually insufficiently meaningful. Besides, there are not enough large datasets (from the same domain) available within the community to carry out such a transfer learning task.

### Discussion and conclusion

The goal of our study was to predict the  $\Delta GPS$  between two consecutive gait sessions in a binary fashion. If this variation is negative, gait quality gets better and vice versa. Globally, from scratch designed architectures gave slightly better results than state-of-the-art ones, which introduce too many parameters to optimize given the relatively small quantity of available data. As a result, a trade-off should be made between the amount of available training data, the complexity of the task, and performance. In the signal-based approach, in general, data augmentation techniques made some improvements in performance. Because of that, we suggest trying to find a way to improve their efficiency. In the image-based approach, developed from scratch CNN surpassed pre-trained Timm models. This can be explained by the fact that the source and target domains are so different. ROC (Receiver operating characteristic) curves for all models are presented in Fig. 11.

To have better insight into results, the ROC curves of the best models (i.e., FCN after data augmentation for the signal-based approach and CNN for the image-based one) were compared using the DeLong's test. This revealed a p-value of  $2.316 \times 10^{-4}$  for the two ROC curves at hand, which means that the AUCs of both models are significantly different. In other words, the FCN model after data augmentation, with a `val AUC` of 0.723 and a `test AUC` of 0.717, is meaningfully better than the CNN model. This outcome proves that knowledge extraction is more efficient on raw signals than synthetic images. In summary, for both approaches, the prediction results are encouraging despite the complexity of such a prediction task on so heterogeneous data. The `val AUC` and `test AUC` are above 0.7 for both approaches.

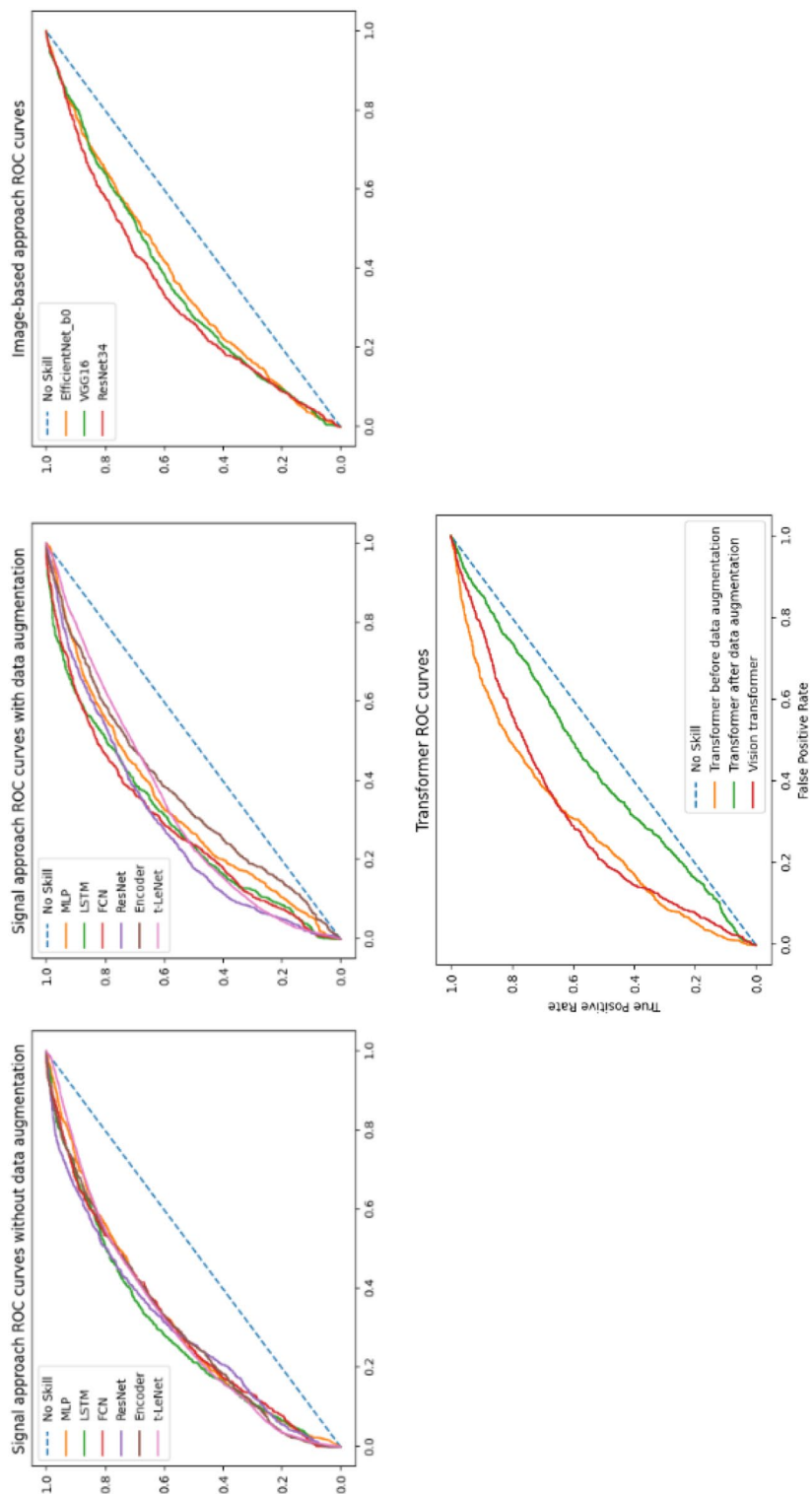
One limitation of this study is the fact that we were not able to validate our findings on external datasets because we did not have any other external data at our disposal. Actually, we were unable to find any publicly-available medical databases.

Our future work will focus on taking the different pathologies into account. Ways of having more data should be thought about as well.

Architecture	Number of parameters	val AUC	test AUC
VGG16	134,263,489	0.650	0.642
ResNet34	21,278,913	0.653	0.679
EfficientNet_b0	4,008,253	0.637	0.628
Vision Transformer	85,260,289	0.652	0.691
CNN	35,505	<b>0.726</b>	<b>0.693</b>

**Table 4.** Quantitative results. Best results in bold.





**Figure 11.** ROC curves for both approaches.

## Data availability

The dataset used and analysed during the current study is available from the corresponding author on reasonable request.

Received: 6 July 2023; Accepted: 13 December 2023

Published online: 28 December 2023

## References

1. Ataullah, A. H. M. & De Jesus, O. Gait Disturbances. in *StatPearls* (StatPearls Publishing, 2022).
2. Auvinet, B., Touzard, C., Montestruc, F., Delafond, A. & Goeb, V. Gait disorders in the elderly and dual task gait analysis: A new approach for identifying motor phenotypes. *J. Neuroeng. Rehabil.* **14**, 1–14 (2017).
3. Bishnoi, A. & Hernandez, M. E. Dual task walking costs in older adults with mild cognitive impairment: A systematic review and meta-analysis. *Aging Mental Health* **25**, 1618–1629 (2021).
4. Pirker, W. & Katzenschlager, R. Gait disorders in adults and the elderly. *Wiener Klinische Wochenschrift* **129**, 81–95. <https://doi.org/10.1007/s00508-016-1096-4> (2017).
5. Rodrigues, F., Domingos, C., Monteiro, D. & Morouço, P. A review on aging, sarcopenia, falls, and resistance training in community-dwelling older adults. *Int. J. Environ. Res. Public Health* **19**, 874 (2022).
6. Osoba, M. Y., Rao, A. K., Agrawal, S. K. & Lalwani, A. K. Balance and gait in the elderly: A contemporary review. *Laryngosc. Investig. Otolaryngol.* **4**, 143–153 (2019).
7. Rodriguez-Fernández, A., Lobo-Prat, J. & Font-Llagunes, J. M. Systematic review on wearable lower-limb exoskeletons for gait training in neuromuscular impairments. *J. Neuroeng. Rehabil.* **18**, 1–21 (2021).
8. Ridao-Fernández, C., Pinero-Pinto, E., Chamorro-Moriana, G. *et al.* Observational gait assessment scales in patients with walking disorders: Systematic review. *BioMed Res. Int.* **2019** (2019).
9. O'Malley, T. *et al.* Kerastuner. <https://github.com/keras-team/keras-tuner> (2019).
10. Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L. & Muller, P. A. Deep learning for time series classification: A review. *Data Mining Knowl. Discov.* **33**, 917–963 (2019).
11. Kadaba, M. P., Ramakrishnan, H. & Wootten, M. Measurement of lower extremity kinematics during level walking. *J. Orthop. Res.* **8**, 383–392 (1990).
12. Christian, J., Kröll, J. & Schwameder, H. Comparison of the classifier oriented gait score and the gait profile score based on imitated gait impairments. *Gait Posture* **55**, 49–54 (2017).
13. Holmes, S. J., Mudge, A. J., Wojciechowski, E. A., Axt, M. W. & Burns, J. Impact of multilevel joint contractures of the hips, knees and ankles on the gait profile score in children with cerebral palsy. *Clin. Biomech.* **59**, 8–14 (2018).
14. Robinson, L. *et al.* The relationship between the Edinburgh visual gait score, the gait profile score and gmfcs levels I–III. *Gait Posture* **41**, 741–743 (2015).
15. Baker, R. *et al.* The gait profile score and movement analysis profile. *Gait Posture* **30**, 265–269 (2009).
16. Barton, G. J., Hawken, M. B., Scott, M. A. & Schwartz, M. H. Movement deviation profile: A measure of distance from normality using a self-organizing neural network. *Hum. Movement Sci.* **31**, 284–294 (2012).
17. Jarvis, H. L. *et al.* The gait profile score characterises walking performance impairments in young stroke survivors. *Gait Posture* **91**, 229–234 (2022).
18. Winter, D. A. *Biomechanics and motor control of human gait: normal, elderly and pathological* (1991).
19. Zdero, R., Brzozowski, P., Schemitsch, E. H. *et al.* Experimental methods for studying the contact mechanics of joints. *BioMed Res. Int.* **2023** (2023).
20. Vaswani, A. *et al.* Attention is all you need. *Adv. Neural Inform. Process. Syst.* **30** (2017).
21. Dosovitskiy, A. *et al.* An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint. arXiv:2010.11929* (2020).
22. Kingma, D. P. & Ba, J. Adam: A method for stochastic optimization. *arXiv preprint. arXiv:1412.6980* (2014).
23. Ruby, U. & Yendapalli, V. Binary cross entropy with deep learning technique for image classification. *Int. J. Adv. Trends Comput. Sci. Eng.* **9** (2020).
24. GitHub-hsd1503/transformer1d: Pytorch implementation of transformer for 1D data—github.com. <https://github.com/hsd1503/transformer1d/tree/master>. [Accessed 29-11-2023].
25. Yang, C.-L., Chen, Z.-X. & Yang, C.-Y. Sensor classification using convolutional neural network by encoding multivariate time series as two-dimensional colored images. *Sensors* **20**, 168 (2019).
26. Wightman, R. Pytorch image models. <https://doi.org/10.5281/zenodo.4414861> (2019).

## Acknowledgements

This study received funding from the French government via the national research agency (Agence Nationale de la Recherche) as part of the investment programme for the future (Programme d'Investissement d'avenir), under the reference ANR-17-RHUS-0005 (FollowKnee Project).

## Author contributions

N.B. conceived the experiments and wrote the manuscript, with the help of M.L.A., P.H.C. and G.Q. She conducted the experiments. M.L.E. gave deep insights into the database. O.R., S.B., B.C. provided clinical inputs. All authors helped supervise the project, discussed the results, commented on the manuscript, and provided critical feedback.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to N.B.C.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024