

Article

# Analysis and Classification of Motor Dysfunctions in Arm Swing in Parkinson's Disease

Tobias Steinmetzer <sup>1,2,†</sup> , Michele Maasch <sup>2,†</sup>, Ingrid Bönninger <sup>2,†</sup> and Carlos M. Travieso <sup>1,\*,†</sup> 

<sup>1</sup> Institute for Technological Development and Innovation in Communications, University of Las Palmas de Gran Canaria, Campus Universitario de Tafira, 35017 Las Palmas de Gran Canaria, Spain; tobias.steinmetzer@b-tu.de

<sup>2</sup> Institute of Medical Technology, Brandenburg University of Technology Cottbus, 01968 Senftenberg, Germany; michele.maasch#1@b-tu.de (M.M.); ingrid.boenninger@b-tu.de (I.B.)

\* Correspondence: ctravieso@dsc.ulpgc.es

† These authors contributed equally to this work.

Received: 29 October 2019; Accepted: 24 November 2019; Published: 3 December 2019



**Abstract:** Due to increasing life expectancy, the number of age-related diseases with motor dysfunctions (MD) such as Parkinson's disease (PD) is also increasing. The assessment of MD is visual and therefore subjective. For this reason, many researchers are working on an objective evaluation. Most of the research on gait analysis deals with the analysis of leg movement. The analysis of arm movement is also important for the assessment of gait disorders. This work deals with the analysis of the arm swing by using wearable inertial sensors. A total of 250 records of 39 different subjects were used for this task. Fifteen subjects of this group had motor dysfunctions (MD). The subjects had to perform the standardized Timed Up and Go (TUG) test to ensure that the recordings were comparable. The data were classified by using the wavelet transformation, a convolutional neural network (CNN), and weight voting. During the classification, single signals, as well as signal combinations were observed. We were able to detect MD with an accuracy of 93.4% by using the wavelet transformation and a three-layer CNN architecture.

**Keywords:** wavelet transformation; gait analysis; inertial sensors; Parkinson's disease; machine learning; wearable sensors

## 1. Introduction

The life expectancy of humankind is increasing worldwide. Life expectancy is projected to increase in the 35 industrialised countries with a probability of at least 65% for women and 85% for men. There is a 90% probability that life expectancy at birth among South Korean women in 2030 will be higher than 86.7 years, the same as the highest worldwide life expectancy in 2012, and a 57% probability that it will be higher than 90 years [1]. Due to the increasing life expectancy, the number of old-age diseases is also increasing. One of them is PD. At present, there are 10 million people affected by this disease, and the trend is increasing [2]. Parkinson's disease is a neurodegenerative disease and is currently incurable. However, the progression of the disease can be delayed by medication. For this reason, an exact diagnosis is very important so that the medication can be adjusted as well as possible to the particular person. There are different rating scales for the uniform assessment, e.g., the Unified Parkinson's Disease Rating Scale (UPDRS) [3]. With the help of this rating scale, for example, cognitive and motor performance are assessed. One of the motor tests is the Timed Up and Go (TUG). The assessment is visual and therefore subjective. For this reason, many researchers are working on the objective evaluation of this test.

Most of the research on gait analysis deals with the analysis of leg motion [4–10]. However, the analysis of the arm movement is also important for the assessment of a gait disorder. Stationary systems that use cameras or ultrasound [11–19] and mobile systems with inertial sensors [20–22] are used to measure the arm swing.

In [11], the arm swings of Parkinson's patients and healthy persons with the help of a Kinect camera were compared. Significant differences in amplitude and speed were observed. The arm movements of Parkinson's patients also often showed asymmetry. The PD group showed significant reductions in arm swing magnitude (left,  $p = 0.002$ ; right,  $p = 0.006$ ) and arm swing speed (left,  $p = 0.002$ ; right,  $p = 0.004$ ) and significantly greater arm swing asymmetry (ASA) ( $p < 0.001$ ). An accuracy of more than 90% in distinguishing healthy people from persons with PD was also achieved using a Kinect camera in [12]. Classification between healthy and non-healthy subjects is performed based on the five most relevant features and the two new obtained features from LDA, using four different classifiers, support vector machine (SVM), multilayer perceptron (MLP), the radial basis (RB) neural network, and k nearest neighbor (KNN). Using the motion capture system Motek CAREN in [13], it was detected that Parkinson's patients have a different jerk and arm swing length compared to healthy people. The fact that Parkinson's patients in the early stages have a larger ASA could be confirmed in [14] with the Vicon and the Baton Rouge motion lab system. The  $p$ -value for distinguishing healthy individuals from individuals with Parkinson's disease was 0.003. A Kinect system was used in [16] to detect the differences in speed, amplitude, and symmetry in arm movement between healthy people and people in the early stages of Parkinson's disease. In [17], it was investigated which model method provided the best results when using a Kinect to detect Parkinson's disease stages. The best results with an accuracy of 93.4% were obtained with a special Bayesian network classifier using 10-fold cross-validation. The relevant features were related to left shin angles, left humerus angles, frontal and lateral bends, left forearm angles, and the number of steps during a spin. For the recordings in [18], a Kinect system was used in combination with an e-Motion capture program. The proposed system classifies PD into three different stages related to freezing of gait (FoG). An accuracy of 93.4% was reached using the features of the movement and position of the left arm, the trunk position for slightly displaced walking sequences, and left shin angle for straight walking sequences. However, they obtained a better accuracy of 96.23% for a classifier that only used features extracted from slightly displaced walking steps and spin walking steps.

In [15], an automatic method for the treatment of levodopa-induced dyskinesia (LID) was developed. Gyroscopes were used on the abdomen and chest and the abdomen, chest, wrists, and ankles. In general, an average detection rate of 90% for Parkinson's patients was achieved, and the average detection rate and the precision of the individual classes (LID, Parkinson, healthy) were 80% and 77%, respectively. Several classification techniques have been used for LID assessment, including the naive Bayes classifier, KNN, fuzzy lattice reasoning (FLR), decision trees, random forests (RF), and neural networks using a multilayer perceptron (MLP).

The method used in [19] consisted of guiding patients with early Parkinson's on a treadmill and measuring their movements with an ultrasound device on each side. The results were a reduced arm swing amplitude in the patients and a longer stride length compared to healthy people.

In [20], a sensor unit was used on each forearm. This sensor unit consisted of two triaxial G-Link accelerometers that were attached to an aluminum bar. Arm swing asymmetry (ASA), maximal cross-correlation (MXC), and instantaneous relative phase (IRP) of bilateral arm swing were compared between PD and controls. PD subjects demonstrated significantly higher ASA ( $p = 0.002$ ) and lower MXC ( $p < 0.001$ ) than controls.

An accelerometer was placed on the upper arm, as well as a magnetic angular rate and gravity (MARG) device on the shoulder in [21]. The Denavit–Hartenberg model was used, and the algorithm was based on the pseudoinverse of the Jacobian by the acceleration of the upper arm. The accuracy of this method was demonstrated by the use of an optoelectronic system for control purposes.

A similar system was used in [22] with nearly the same sensors and sensor position. An eigenvector method was suggested to compare the axes of the left and right hand. The results showed a difference between people with Parkinson's disease and healthy people.

In our approach, we want to propose a medical wearable system that:

- (a) classifies between subjects with motor dysfunctions and a control group based exclusively on arm motions
- (b) uses 3D data from the accelerometer, gyroscope, and magnetometer
- (c) includes new parameters
- (d) is small and easy to use
- (e) is not bound to a location
- (f) requires a small number of sensors
- (g) is low cost

According to the previously mentioned classification, this paper is organized as follows. Section 2 describes our materials. The section is divided into the medical experiment protocol, the hardware used, and the dataset. Then, in Section 3, a description of our methods and how we apply the methods to our data are described. Section 4 include the results. Finally, a discussion and comparison is found in Section 5.

## 2. Materials

### 2.1. Protocol

We decided to use the TUG test as a suitable test for recording gait data. Among other things, it is used to evaluate the motor performance of the UPDRS. For the test, only a chair with a backrest and armrests was needed. At first, the test person was sitting on a chair. Upon a command from the test leader, the test person stood up and walked straight ahead for ten meters at an appropriate speed to a mark. At the mark, the test person turned around and walked ten meters straight ahead, back to the chair. The test person sat down in the chair. The test and the recording were then finished. We divided the TUG into two different parts for later analysis of the data. Part (A) contained all data of the TUG including standing up and sitting down in the chair. Part (B) included going straight to the mark, turning around, and going straight back to the chair. Parts (A) and (B) are shown in Figure 1. The aim of this splitting was to extract the gait data from the complete recording.

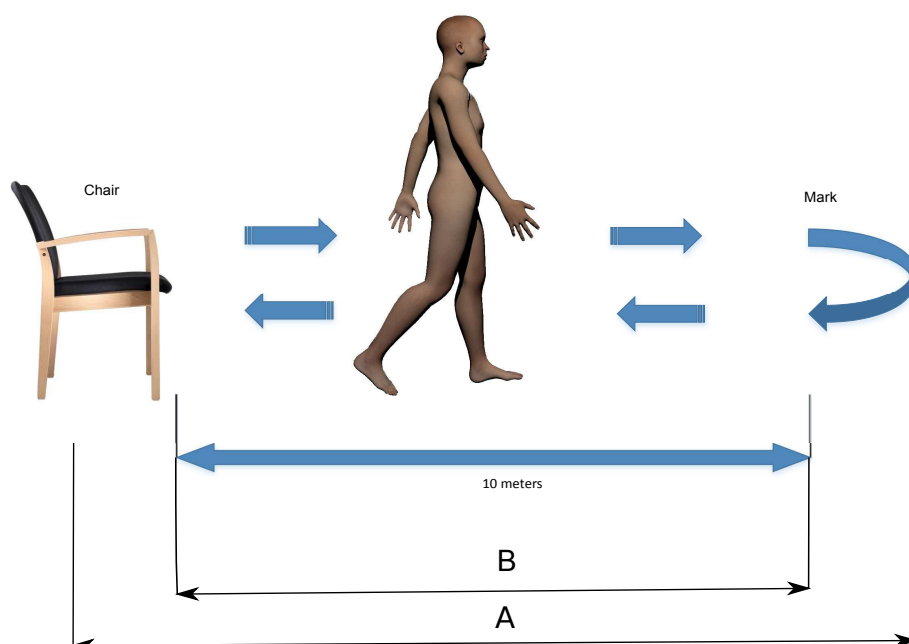
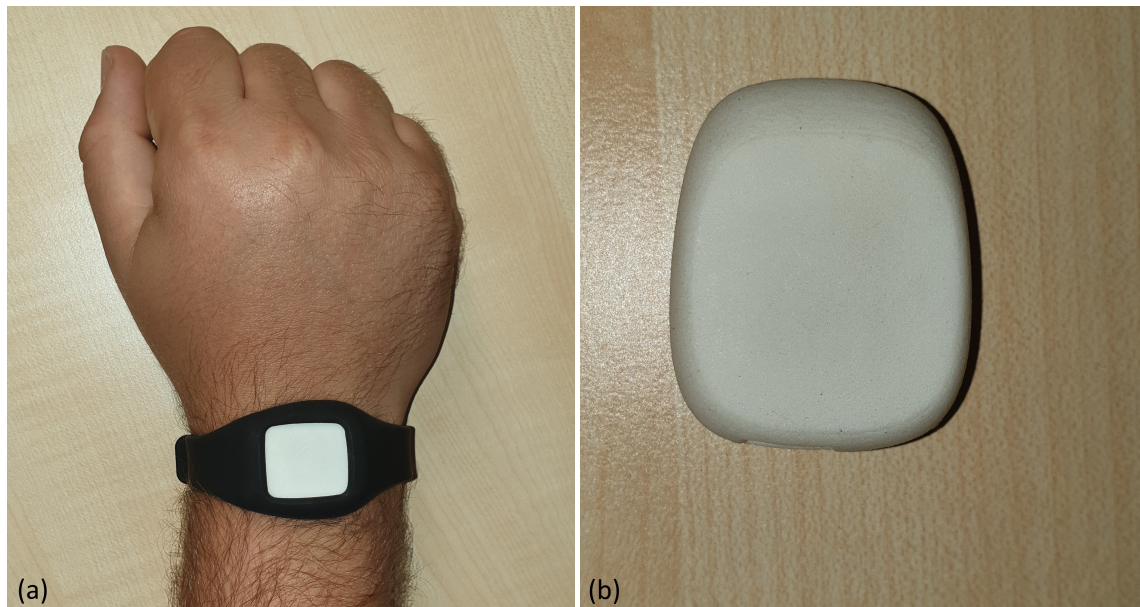


Figure 1. Process of the Timed Up and Go (TUG) test.

## 2.2. Hardware

For data recording, we used two wristbands with the Meta Motion Rectangle wearable sensors from Mbiientlab; see Figure 2 [23]. This is an inertial measurement unit (IMU) sensor. It consists of a BMI 160 with a 3-axis gyroscope and a 3-axis accelerometer and a BMM 150 with a 3-axis magnetometer. By using the Bosch sensor fusion algorithm, the Euler angle and linear acceleration can be obtained [24]. The x-axis corresponds to the gait direction.



**Figure 2.** (a) Wristband with the Meta Motion Rectangle sensor. (b) Position of the sensor during the measurement.

## 2.3. Data

### 2.3.1. Dataset

To create a dataset for later analysis, we worked together with the Niederlausitz Clinic in the study “Development of a digital Parkinson Disease Assessment” (ethics request granted in December 2018 by Ethics Committee Brandenburg). All persons were evaluated by the physicians. A total of 39 different persons with 250 recordings were available for the dataset. Of these, there were 15 motor dysfunction patients with 80 recordings and 24 persons with 170 recordings as the control group. Table 1 summarizes the data.

**Table 1.** Amount of persons and records from the Parkinson’s and control groups.

Label	Persons	Records
Motor dysfunction	15	80
Control	24	170

### 2.3.2. Sensor Data

While the subjects performed the TUG test, 3D Euler angles and 3D linear acceleration of the arms were captured. The signals for the Euler angles and the linear acceleration were the result of the sensor fusion algorithm from Bosch. Both signals were recorded at a frequency of 100 Hz. The algorithm for the sensor fusion used the data from the accelerometer, gyroscope, and magnetometer. Figure 3 shows at the top the 3D Euler angles and at the bottom the 3D linear acceleration signals. In Figure 3,

the complete signal of one wristband during the TUG test is shown. Furthermore, Part (A) contains all recorded data and Part (B) the data between the black dotted lines, the active walking parts.

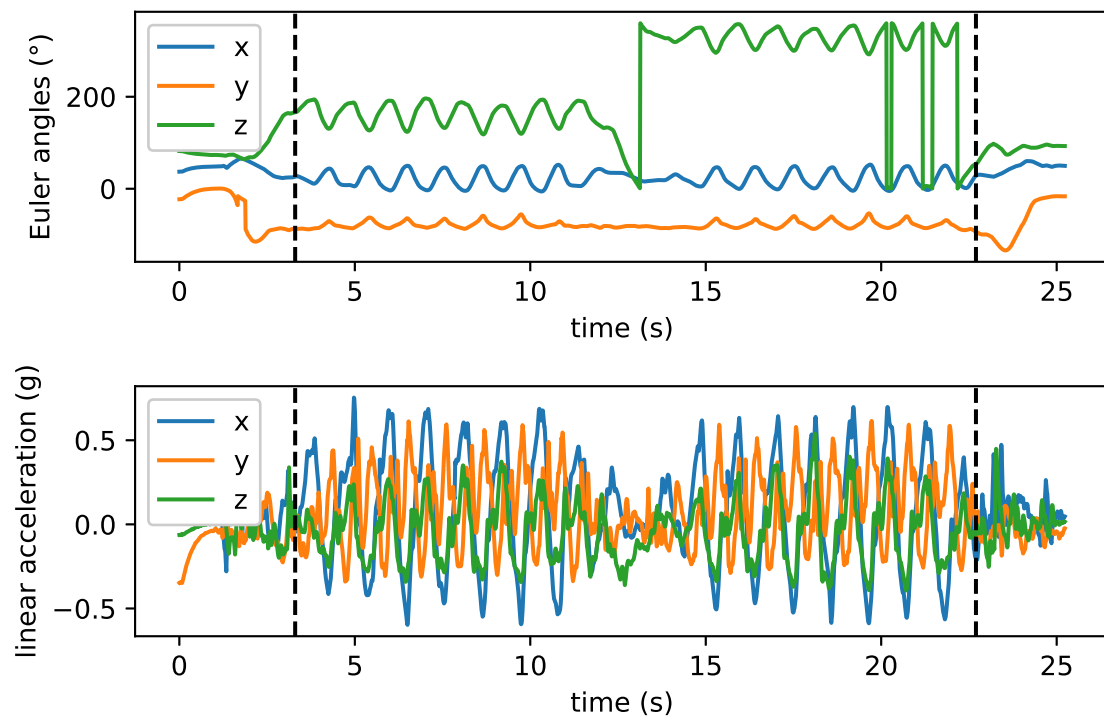


Figure 3. Euler angles and linear acceleration of one wristband for the TUG test.

### 3. Methods

#### 3.1. Removing Jumps

Figure 3 shows that some jumps existed in the signal of the z-axis of the Euler angle. This was because the value range of the sensor was between  $0^\circ$  and  $360^\circ$ . This made the signal unstable. To correct this, we removed all jumps that were greater than a threshold of  $300^\circ$ . In Equation (1), our procedure is shown. If the absolute value of the difference of two successive sensor values  $|x_i - x_{i+1}| > 300$ , a correction of the signal was performed, where  $i \in \{1, \dots, N\}$ .  $N$  indicates the length of the signal. The result of the cleanup are given in Figure 4.

$$x_{i+1} = \begin{cases} x_{i+1} - 360 & \text{for } x_i < x_{i+1} \\ x_{i+1} + 360 & \text{for } x_i > x_{i+1} \end{cases} \quad (1)$$

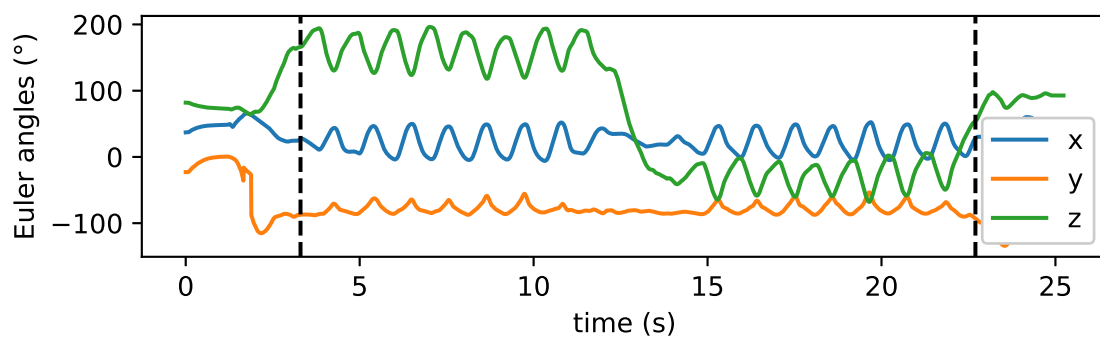


Figure 4. Euler angles without jumps.

### 3.2. Derivation

It was not possible to create a classifier that could classify the subjects with motor dysfunctions (MD) and no MD by using the Euler angles, because the Euler angles were measured in absolute values. This means that the angles were not calibrated to a starting value at the beginning of the recording. For this reason, we calculated the derivative of each axis of the Euler angles. For this purpose, we calculated the difference between two successive measured values. The equation of the first order discrete derivative can be seen in (2), where  $N$  is the length of the signal,  $x_i$  is the signal at index  $i$ , and  $x'_i$  is the value for the difference at  $i$ . The result of the derivation can be seen in Figure 5. The derivation makes the signals more comparable for different recordings. This is because the relative angle is used by the derivation.

$$X'_i = x_{i+1} - x_i, i \in \{1, 2, \dots, (N - 1)\} \quad (2)$$

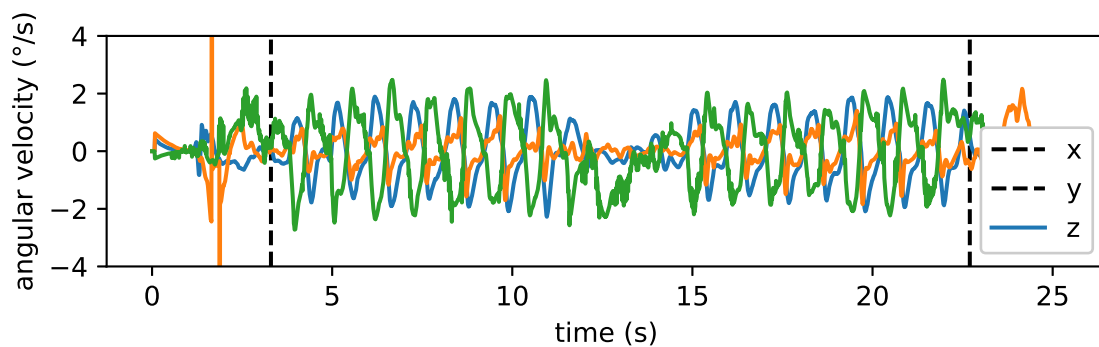


Figure 5. Derivation of the Euler angles.

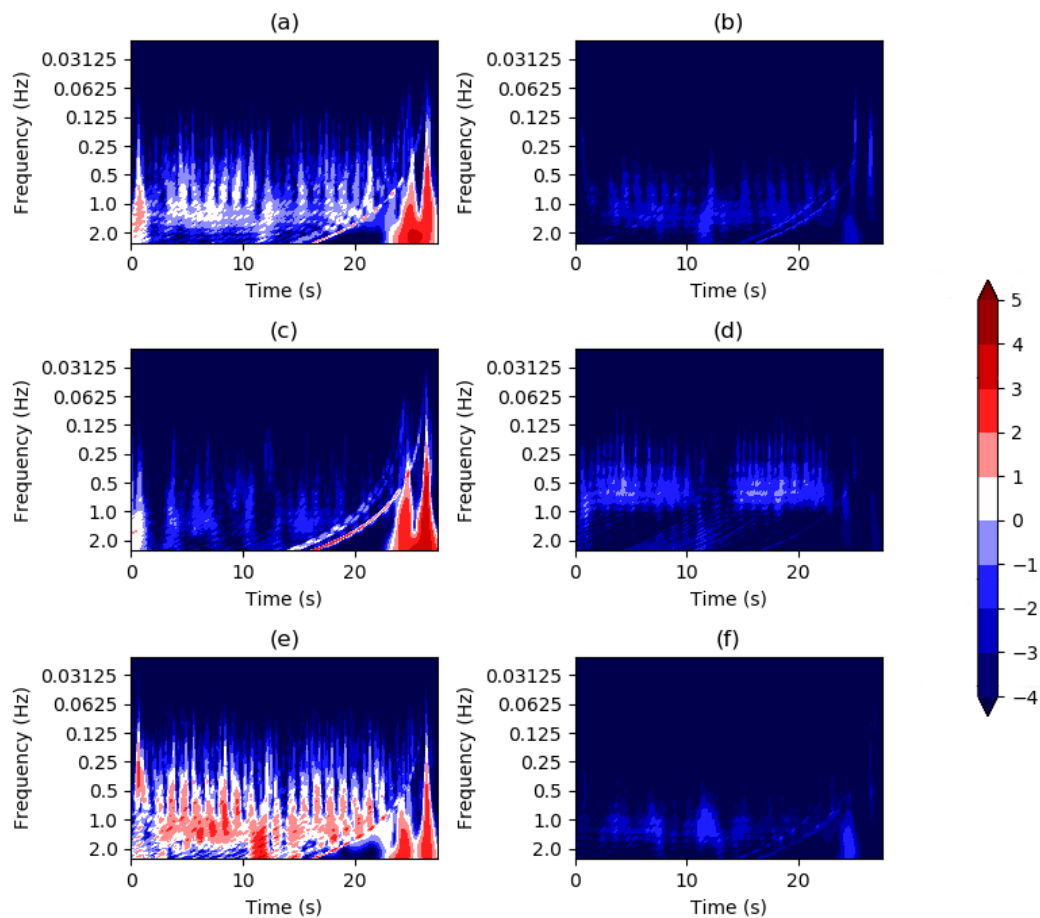
### 3.3. Resampling

Before CNN can interpret the data, the signal must have a uniform length. To do this, we resampled the data to a length of 512 values. For resampling, we used the Python library SciPy [25].

### 3.4. Wavelet Transformation

When considering static signals, the Fourier transformation is very well suited. Unfortunately, there are hardly any static signals in the real world. Every signal changes its frequency dynamically in time. This also applies to the human gait. The gait is a dynamic process. For this reason, it does not make sense to use Fourier analysis.

The origin of the data was a temporal series; therefore, we preferred the use of the wavelet transform in order to increase the information, by decomposition of the time frequency. After the experiments, the accuracy showed a useful feature extracted from this transform. For the wavelet transformation, a signal was convoluted with a wavelet template. By selecting the kernel, we ensured that the ranges around 1.2 Hz (frequency of the arm swing [26]) had a high amplitude. With this template, we calculated the wavelet transformation over the complete signal. In our case, these were the x-, y-, and z-axes of the derived Euler angle and the x-, y-, and z-axes of the linear acceleration of both wristbands. Figure 6 shows the scalograms of the individual signals of one wristband. On the y-axis, the frequencies are shown in Hertz and on the x-axis the time in seconds. For the calculation of the wavelet transformations, we used the Python library PyWavelets [27].



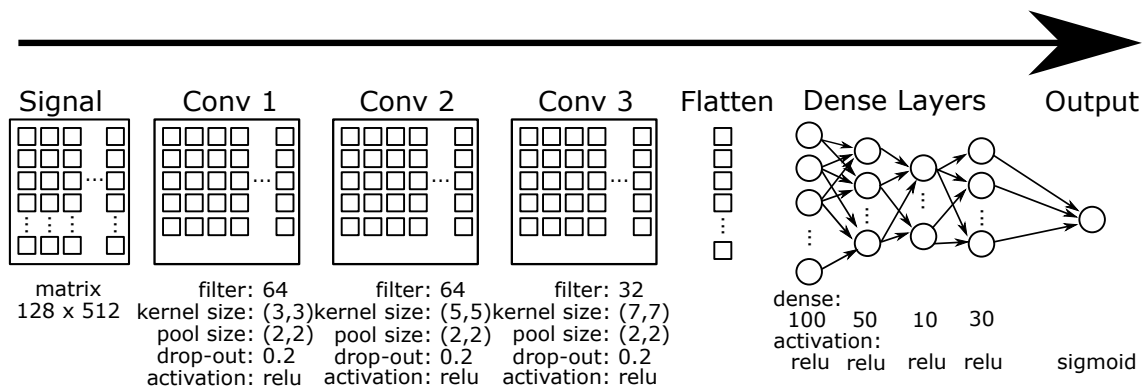
**Figure 6.** (a) x-axis of the derived Euler angle. (b) y-axis of the derived Euler angle. (c) z-axis of the derived Euler angle. (d) x-axis of linear acceleration. (e) y-axis of linear acceleration. (f) z-axis of linear acceleration.

Figure 6a,c,e corresponds to the x-, y-, and z-axes of the derived Euler angles. We calculated for each signal the continuous wavelet transformation with the Morlet wavelet. It can be seen that there was a high amplitude from 0.25 Hz. In the lower frequency data  $< 0.25$ , the individual arm swings can be seen.

Figure 6b,d,f reflects the x-, y-, and z-axes of linear acceleration. We calculated for each signal the continuous wavelet transformation with the Morlet wavelet. With these data, it can be seen that the largest amplitude was in the range of 1 Hz. This corresponds to the natural arm swing since this corresponds to a frequency of approximately 1.2 Hz [26].

### 3.5. CNN

In image classification, as well as other signals, the application of CNNs has been very successful. The difference from common NNs is that a CNN searches for a local pattern in the input signal. When using multiple CNN layers, one after the other, larger patterns can be detected [28,29]. Thus, a CNN often provides better classification results than NN. In our case, we achieved the best results with the use of three convolution layers. Then, we applied one NN with three encoders and one decoder. Our used CNN with the configuration is shown in Figure 7. We used Python and the Keras library to create the CNN [30]. We obtained the architecture for our CNN by systematically testing. We wanted to keep the number of CNN layers as small as possible. However, with less than three layers, no useful results were available.



**Figure 7.** Construction of a single signal CNN for classification.

In order to have a useful input for the CNN, we resampled the signal to a uniform length of 512 values; see Section 3.3. We then applied a wavelet transformation to the signal; see Section 3.4. This gave us a  $128 \times 512$  matrix for the signal. We used this matrix as input for the CNN. As the activation function, we used the ReLU function for all convolution layers. We also used the ReLU function in the hidden layers of the encoder and decoder. The equation of the ReLU function can be seen in Equation (3). The characteristic of the ReLU function is that the weight of the output is not negative. In the output layer, we used the sigmoid function; see Equation (4). After each convolution layer, we performed a two-dimensional max-pooling with a pool size of  $2 \times 2$  and a drop out with a probability of 0.2.

$$f(x) = \max(0, x) \quad (3)$$

$$f(x) = \frac{1}{1 + e^{-x}} \quad (4)$$

The first convolutional layer searched for the smallest pattern from the signal. For convolution, we used a  $3 \times 3$  matrix. In total, we created 64 different filters in the first convolutional layer. In the second convolutional layer, we increased our kernel size to  $5 \times 5$  and created 64 filters again. The third convolutional layer had a kernel size of  $7 \times 7$ , and the filters created were reduced to 32 pieces. After the convolutional layers, we used a flatten layer so that the signal could be interpreted by the dense layers. In the dense layers, we started with three encoder layers with 100, 50, and 10 neurons, followed by a decoder layer with 30 neurons. Finally, we obtained our prediction in the output layer. Since we had a binary problem, a single neuron was used. For the training of the models, we used a batch size of 50 and 50 epochs. For training, we used an Intel Core i7-6700HQ with 2.6 GHz with four cores. Furthermore, the system used 16 GB RAM. The computer required approximately 45 min to train a model.

### 3.6. Multi-Channel CNN

In the last section, we presented our architecture for a single signal. To achieve better and more robust results, we wanted to use multiple channels  $x$ ,  $y$  Euler angles, and  $x$  of linear acceleration for classification. For this reason, we created an  $m$ -dimensional input. For the third dimension, we used the number of  $m$  different signals used. Figure 8 shows the construction. Another difference was that the first convolutional layer created 128 filters. The model was similar to the one in Figure 7. The computer required approximately 2 h to train a model.



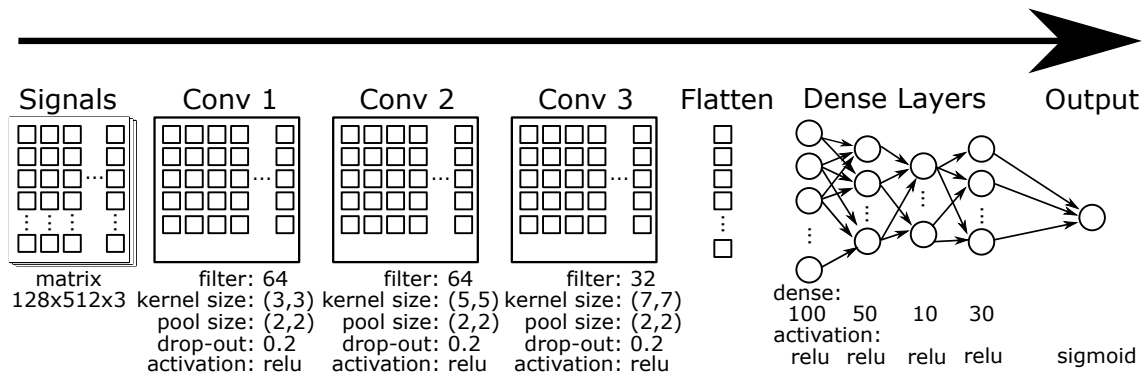


Figure 8. Construction of a 3-channel CNN to use three different signals for classification.

### 3.7. Weight Voting

The multi-channel CNN was trained with 3 signals at the same time. The difference in voting was that for each signal, a separate model was trained, which was independent of the other models. In our case, we had a binary problem, so the calculation for the voting was easy. We used the predicted classes and calculated the average of all predictions; see Equation (5), where  $m_i$  is the prediction of a model from a classifier and  $M$  is the number of classifiers.

$$v = \frac{1}{M} \sum_{i=1}^M m_i, i \in \{1, 2, \dots, M\} \tag{5}$$

If  $v \geq 0.5$ , then the predicted class is MD and in all other cases, no MD; see Equation (6).

$$prediction = \begin{cases} MD, & v \geq 0.5 \\ noMD, & v < 0.5 \end{cases} \tag{6}$$

### 3.8. Evaluation

We decided to use 3-fold cross-validation for the classification to make the results of our applied methods reasonable. We used 66.6% of the data for training and 33.3% for testing. For each measurement, we calculated the sensitivity, specificity (precision), recall, F1-score, and accuracy. For this, we used the confusion matrix in Table 2.

Table 2. Binary confusion matrix.

	Classes	
	Positive	Negative
predicted positive	TP true positive	FP false positive
predicted negative	FN false negative	TN true negative

Sensitivity (recall) is a widespread measurement in medicine. It indicated the ratio of predicted MD to all MD inside our test data; see Equation (7). The specificity described how well our system can distinguish MD from the control group (no MD). It was the ratio of predicted non-MD persons in all test data where healthy persons were present; see Equation (8). Precision was the proportion of correctly predicted MD to all MD; see Equation (9). Accuracy was the ratio of all correctly recognized

MD and no MD to all test data; see Equation (10). The F1-score (F1) was the harmonious average between precision and recall; see Equation (11).

$$\text{recall} = \text{sensitivity} = \frac{TP}{TP + FN} \quad (7)$$

$$\text{specificity} = \frac{TN}{FP + TN} \quad (8)$$

$$\text{precision} = \frac{TP}{TP + FP} \quad (9)$$

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN} \quad (10)$$

$$F1 = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \quad (11)$$

### 3.9. Methodology

After we have presented our material and methods, we will now discuss in this section how we applied these methods. In the presentation of the dataset, we already said that we divided our recording into two different parts. First, we classified Parts (A) and (B), which comprised the complete recording of the TUG test. The other scenario was that we only used Part (B). In Part (B), only the gait was used. Figure 9 shows the complete algorithm of the classification. In principle, we distinguished between the signals of the Euler angles and the linear acceleration. First, we removed the jumps within a signal of the Euler angles and then calculated the derivation of the signal. This made the signal more comparable. These steps were not necessary for linear acceleration. Then, we set the signals to a uniform length. This was necessary so that the signals could be interpreted by CNN later during classification. After resampling, we calculated the wavelet transformation for each individual signal. We used the resulting scalograms for the classification. In the classifications, we analyzed three different cases. At first, we classified each signal individually by CNN. This allowed us to show which axis of the sensors was very important. In the second case of classification, we used the three best signals for a multi-channel CNN. The third case was that we used the three best signals for classification by voting.

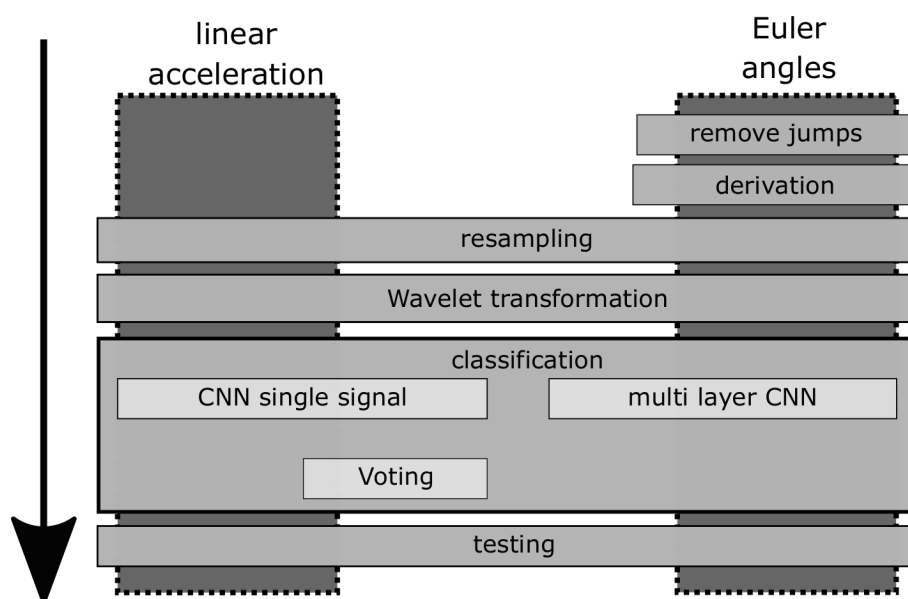


Figure 9. Classification process for detecting motor dysfunctions in arm swinging.

## 4. Results

### 4.1. Parts (A) and (B) of TUG

#### 4.1.1. Single Layer

To find out which sensor data were particularly useful for classification, we first separated all signals from each other. The results are shown in Table 3. In the table, we applied three-fold cross-validation to the sensor data. Furthermore, we optically separated the results from the Euler angles and the linear acceleration with a double line. For each signal, we calculated the precision, specificity, recall, F1-score, and accuracy. In every cell, we show the mean  $\bar{x} = \frac{1}{N} \sum_{i=1}^N x_i, i \in \{1, 2, \dots, N\}$  plus or minus the standard deviation  $s = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2}, i \in \{1, 2, \dots, N\}$ , where  $N$  is the length of the signal. The columns with the best results are highlighted with bold. It can be seen that the x-axis of the Euler angle and the x-axis of the linear acceleration produced the best results. Furthermore, it can be seen that the z-axis of the Euler angle and linear acceleration provided the lowest results.

**Table 3.** Results of a single signal by CNN classification. Parts (A) and (B) of the TUG test are used.

Signal	Sensitivity	Specificity	Recall	F1-Score	Accuracy
x Euler angles	<b>0.918 ± 0.071</b>	<b>0.939 ± 0.043</b>	<b>0.887 ± 0.085</b>	<b>0.898 ± 0.017</b>	<b>0.928 ± 0.009</b>
y Euler angles	0.891 ± 0.014	0.874 ± 0.016	0.775 ± 0.072	0.829 ± 0.047	0.882 ± 0.009
z Euler angles	0.57 ± 0.505	0.844 ± 0.186	0.606 ± 0.527	0.587 ± 0.514	0.821 ± 0.173
x linear acceleration	<b>0.907 ± 0.101</b>	<b>0.901 ± 0.048</b>	<b>0.846 ± 0.036</b>	<b>0.873 ± 0.046</b>	<b>0.908 ± 0.015</b>
y linear acceleration	0.857 ± 0.031	0.888 ± 0.056	0.841 ± 0.066	0.848 ± 0.032	0.877 ± 0.027
z linear acceleration	0.795 ± 0.118	0.863 ± 0.044	0.74 ± 0.043	0.761 ± 0.037	0.841 ± 0.009

#### 4.1.2. Signal Combination

To get better results in the classification, we decided to combine the individual layers. For the combination, there were several possibilities. On the one hand, it was possible to use an ensemble classifier like voting. On the other hand, we could use a multi-channel CNN. In Table 3, the x-axis of the Euler angles and the linear acceleration produced the best results. The third was the Euler angles of the y-axis. In this section, we used these three signals to improve our results. The results are shown in Table 4. We again used three-fold cross-validation for our results. Each cell represented the result as  $\bar{x} \pm s$ , as introduced in Section 4.1.1.

**Table 4.** Classification results by combining the x- and y-axis of Euler angles and the x-axis of linear acceleration. Parts (A) and (B) of the TUG test are used.

Layer	Sensitivity	Specificity	Recall	F1-Score	Accuracy
3 channel CNN	0.934 ± 0.047	0.932 ± 0.013	0.899 ± 0.026	0.928 ± 0.043	0.933 ± 0.024
3 signal voting	0.915 ± 0.078	0.9 ± 0.02	0.821 ± 0.052	0.862 ± 0.026	0.902 ± 0.018

Table 4 shows the results of the signal combination classification. The three channel CNN achieved better results than the three signal voting. The three channel CNN was also better than any signal in Table 3.

### 4.2. Part (B) of TUG

#### 4.2.1. Single Layer

In this section, we present our results if only Part (B) of the TUG test was used for classification. In Table 5, you can see the results for a CNN classification for each axis of the sensors. As in Section 4.1.1, we used three-fold cross-validation and calculated the average  $\bar{x}$  plus or minus the standard deviation  $s$ . The best results for each sensor and each column are marked with bold. Like the analysis of

the complete TUG test, the x-axis provided the best results for Euler angles and linear acceleration. However, the results were not as accurate as in Section 4.1.1.

**Table 5.** Results of a single signal by CNN classification. Only Part (B) of the TUG test is used.

Signal	Sensitivity	Specificity	Recall	F1-Score	Accuracy
x Euler angles	$0.873 \pm 0.027$	$0.899 \pm 0.043$	$0.822 \pm 0.088$	$0.844 \pm 0.039$	$0.887 \pm 0.018$
y Euler angles	$0.793 \pm 0.037$	$0.855 \pm 0.016$	$0.756 \pm 0.046$	$0.772 \pm 0.006$	$0.831 \pm 0.015$
z Euler angles	$0.763 \pm 0.202$	$0.943 \pm 0.049$	$0.904 \pm 0.088$	$0.809 \pm 0.099$	$0.821 \pm 0.138$
x linear acceleration	$0.909 \pm 0.012$	$0.9 \pm 0.044$	$0.822 \pm 0.088$	$0.862 \pm 0.053$	$0.903 \pm 0.032$
y linear acceleration	$0.804 \pm 0.041$	$0.832 \pm 0.043$	$0.705 \pm 0.078$	$0.748 \pm 0.033$	$0.821 \pm 0.024$
z linear acceleration	$0.563 \pm 0.496$	$0.794 \pm 0.147$	$0.508 \pm 0.468$	$0.52 \pm 0.453$	$0.774 \pm 0.111$

#### 4.2.2. Signal Combination

Table 6 shows the results of the signal combination of Part (B) of the TUG test. For the results, three-fold cross-validation was applied and for each cell, and the average  $\bar{x}$  plus or minus the standard deviation  $s$  was calculated. The three signal voting performed best. However, the results were marginally better than the single signal CNN classification in Table 5. Furthermore, the results were not as good as if the complete TUG test was used for the classification.

**Table 6.** Classification results by combining the x- and y-axis of Euler angles and the x-axis of linear acceleration. Only Part (B) of the TUG test is used.

Layer	Sensitivity	Specificity	Recall	F1-Score	Accuracy
3 layer CNN	$0.888 \pm 0.045$	$0.847 \pm 0.027$	$0.677 \pm 0.065$	$0.766 \pm 0.042$	$0.856 \pm 0.024$
3 signal voting	$0.914 \pm 0.03$	$0.901 \pm 0.043$	$0.822 \pm 0.088$	$0.863 \pm 0.04$	$0.903 \pm 0.024$

## 5. Discussion

In Tables 3 and 5, the x-axis always shows the best results. The x-axis corresponds to the movement in the sagittal plane. According to the literature, the most important characteristics of human gait are also present in this plane [31,32]. For this reason, it is a logical conclusion that the features with the highest significance are present on this axis.

We presented our results in the previous section. We compared the results when the complete TUG test, Parts (A) and (B), was used for the classification, as well as if we only used the gait, Part (B), for the classification. The results showed that for the classification of motor dysfunctions, the gait alone gave quite good results with an accuracy of 90.3%, but when looking at the complete test, we obtained even better results with an accuracy of 93.3%. From this, we concluded that the complete TUG test was necessary for the analysis of motor dysfunctions.

Furthermore, we classified each signal separately. During the classification, we found out that the x-axis of the Euler angle and linear acceleration gave the best results, independent of whether Parts (A) and (B), as well as only Part (B) were used for the classification. From this, we concluded that the x-axis was the most relevant.

The conclusion was that we obtained better results through the combination of the signals compared to single signals. In the classification of Parts (A) and (B), the three-channel CNN proved to be the best solution. When classifying with only Part (B), voting was the best choice.

Table 7 shows our classification results compared to the corresponding state-of-the-art works. Our results were comparable to the results from large, expensive, and stationary video based systems.

**Table 7.** Comparison of classification results with other works.

Reference	Description	Accuracy
Our System	IMU sensors	93.3%
[12]	Kinect camera	90%
[17]	Kinect, Bayesian network	93.4%
[18]	Kinect and e-Motion capture program	96.23%
[15]	Gyroscope	90%

Our system delivered better results than the wearable system that also classified the data [15]. We could not make a comparison with the other works because they focused on a statistical evaluation of the data. CNN in combination with wavelet transformations was a powerful technique for arm swing analysis.

**Author Contributions:** Conceptualization, C.M.T.; methodology, T.S.; software, T.S. and M.M.; validation, T.S. and M.M.; formal analysis, T.S. and M.M.; investigation, T.S. and I.B.; resources, T.S. and I.B.; data curation, T.S.; writing, original draft preparation, T.S., I.B., and M.M.; writing, review and editing, T.S., I.B., and C.M.T.; visualization, T.S.; supervision, C.M.T.; project administration, C.M.T.

**Funding:** This research received no external funding.

**Acknowledgments:** Many thanks go to the physicians Dorela Erk, Markus Christoph Reckhardt, and Fritjof Reinhardt of the Niederlausitz Clinic for their assessment of the symptoms of the subjects. Thanks also go to the Brandenburg University of Technology Cottbus-Senftenberg for financing the equipment.

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

## References

- Kontis, V.; Bennett, J.E.; Mathers, C.D.; Li, G.; Foreman, K.; Ezzati, M. Future life expectancy in 35 industrialised countries: Projections with a Bayesian model ensemble. *Lancet* **2017**, *389*, 1323–1335. [CrossRef]
- Parkinson's Foundation. Available online: <https://www.parkinson.org/Understanding-Parkinsons/Statistics> (accessed on 1 October 2019).
- Goetz, C.G.; Tilley, B.C.; Shaftman, S.R.; Stebbins, G.T.; Fahn, S.; Martinez-Martin, P.; Poewe, W.; Sampaio, C.; Stern, M.B.; Dodel, R.; et al. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov. Disord.* **2008**, *23*, 2129–2170. [CrossRef] [PubMed]
- Mazumder, O. Development and Control of Active Lower Limb Exoskeleton for Mobility Regeneration and Enhancement. Ph.D. Thesis, Indian Institute of Engineering Science and Technology, Shibpur, India, 2018.
- Jasni, F.; Hamzaid, N.A.; Al-Nusairi, T.Y.; Yusof, N.H.M.; Shasmin, H.N.; Cheok, N.S. Feasibility of A Gait Phase Identification Tool for Transfemoral Amputees using Piezoelectric-Based In-Socket Sensory System. *IEEE Sens. J.* **2019**, *19*, 6437–6444. [CrossRef]
- Stoelben, K.J.V.; Pappas, E.; Mota, C.B. Lower extremity joint moments throughout gait at two speeds more than 4 years after ACL reconstruction. *Gait Posture* **2019**, *70*, 347–354. [CrossRef] [PubMed]
- Balzer, J.; Marsico, P.; Mitteregger, E.; van der Linden, M.L.; Mercer, T.H.; van Hedel, H.J. Influence of trunk control and lower extremity impairments on gait capacity in children with cerebral palsy. *Disabil. Rehabil.* **2018**, *40*, 3164–3170. [CrossRef] [PubMed]
- Prakash, C.; Sujil, A.; Kumar, R.; Mittal, N. Linear Prediction Model for Joint Movement of Lower Extremity. In *Recent Findings in Intelligent Computing Techniques*; Springer: Singapore, 2019; pp. 235–243.
- Steinmetzer, T.; Bonninger, I.; Priwitzer, B.; Reinhardt, F.; Reckhardt, M.C.; Erk, D.; Travieso, C.M. Clustering of Human Gait with Parkinson's Disease by Using Dynamic Time Warping. In Proceedings of the 2018 IEEE International Work Conference on Bioinspired Intelligence (IWOB), San Carlos, Costa Rica, 18–20 July 2018; pp. 1–6.
- Steinmetzer, T.; Bönninger, I.; Reckhardt, M.; Reinhardt, F.; Erk, D.; Travieso, C.M. Comparison of algorithms and classifiers for stride detection using wearables. *Neural Comput. Appl.* **2019**, 1–12. [CrossRef]

11. Ospina, B.M.; Chaparro, J.A.V.; Paredes, J.D.A.; Pino, Y.J.C.; Navarro, A.; Orozco, J.L. Objective arm swing analysis in early-stage Parkinson's disease using an RGB-D camera (Kinect). *J. Parkinson's Dis.* **2018**, *8*, 563–570. [CrossRef] [PubMed]
12. Spasojević, S.; Santos-Victor, J.; Ilić, T.; Milanović, S.; Potkonjak, V.; Rodić, A. A vision-based system for movement analysis in medical applications: The example of Parkinson disease. In *International Conference on Computer Vision Systems*; Springer: Cham, Switzerland, 2015; pp. 424–434.
13. Baron, E.I.; Koop, M.M.; Streicher, M.C.; Rosenfeldt, A.B.; Alberts, J.L. Altered kinematics of arm swing in Parkinson's disease patients indicates declines in gait under dual-task conditions. *Parkinsonism Relat. Disord.* **2018**, *48*, 61–67. [CrossRef] [PubMed]
14. Lewek, M.D.; Poole, R.; Johnson, J.; Halawa, O.; Huang, X. Arm swing magnitude and asymmetry during gait in the early stages of Parkinson's disease. *Gait Posture* **2010**, *31*, 256–260. [CrossRef] [PubMed]
15. Tsiouras, M.G.; Tzallas, A.T.; Rigas, G.; Tsouli, S.; Fotiadis, D.I.; Konitsiotis, S. An automated methodology for levodopa-induced dyskinesia: Assessment based on gyroscope and accelerometer signals. *Artif. Intell. Med.* **2012**, *55*, 127–135. [CrossRef] [PubMed]
16. Castaño, Y.; Navarro, A.; Arango, J.; Muñoz, B.; Orozco, J.L.; Valderrama, J. Gait and Arm Swing Analysis Measurements for Patients Diagnosed with Parkinson's Disease, using Digital Signal Processing and Kinect. In Proceedings of the SSN2018, Valdivia, Chile, 29–31 October 2018; pp. 71–74.
17. Dranca, L.; de Mendarozketa, L.D.A.R.; Goñi, A.; Illarramendi, A.; Gomez, I.N.; Alvarado, M.D.; Rodríguez-Oroz, M.C. Using Kinect to classify Parkinson's disease stages related to severity of gait impairment. *BMC Bioinform.* **2018**, *19*, 471. [CrossRef] [PubMed]
18. Castaño-Pino, Y.J.; Navarro, A.; Muñoz, B.; Orozco, J.L. Using Wavelets for Gait and Arm Swing Analysis. In *Wavelet Transform and Complexity*; IntechOpen: London, UK, 2019.
19. Roggendorf, J.; Chen, S.; Baudrexel, S.; Van De Loo, S.; Seifried, C.; Hilker, R. Arm swing asymmetry in Parkinson's disease measured with ultrasound based motion analysis during treadmill gait. *Gait Posture* **2012**, *35*, 116–120. [CrossRef] [PubMed]
20. Huang, X.; Mahoney, J.M.; Lewis, M.M.; Du, G.; Piazza, S.J.; Cusumano, J.P. Both coordination and symmetry of arm swing are reduced in Parkinson's disease. *Gait Posture* **2012**, *35*, 373–377. [CrossRef] [PubMed]
21. Bertomeu-Motos, A.; Lledó, L.; Díez, J.; Catalan, J.; Ezquerro, S.; Badesa, F.; Garcia-Aracil, N. Estimation of human arm joints using two wireless sensors in robotic rehabilitation tasks. *Sensors* **2015**, *15*, 30571–30583. [CrossRef] [PubMed]
22. Viteckova, S.; Kutilek, P.; Lenartova, J.; Kopecka, J.; Mullerova, D.; Krupicka, R. Evaluation of movement of patients with Parkinson's disease using accelerometers and method based on eigenvectors. In Proceedings of the 2016 17th International Conference on Mechatronics-Mechatronika (ME), Prague, Czech Republic, 7–9 December 2016; pp. 1–5.
23. MbiLab. MetaWear RG/RPro. 2016. Available online: <https://mbientlab.com/docs/MetaWearRPROPSv0.8.pdf> (accessed on 4 January 2016).
24. BOSCH Sensortec, Data Sheet BMI160. 2015. Available online: [https://ae-bst.resource.bosch.com/media/\\_tech/media/datasheets/BST-BMI160-DS000-07.pdf](https://ae-bst.resource.bosch.com/media/_tech/media/datasheets/BST-BMI160-DS000-07.pdf) (accessed on 2 December 2019).
25. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. SciPy: Open Source Scientific Tools For Python. 2019. Available online: <http://www.scipy.org/> (accessed on 2 December 2019).
26. Hausdorff, J.M.; Cudkowicz, M.E.; Firtion, R.; Wei, J.Y.; Goldberger, A.L. Gait variability and basal ganglia disorders: Stride-to-stride variations of gait cycle timing in Parkinson's disease and Huntington's disease. *Mov. Disord.* **1998**, *13*, 428–437. [CrossRef] [PubMed]
27. Lee, G.R.; Gommers, R.; Wasilewski, F.; Wohlfahrt, K.; O'Leary, A. PyWavelets: A Python package for wavelet analysis. *J. Open Source Softw.* **2019**, *4*, 1237. [CrossRef]
28. Abadi, M.; Abadi, M.; Barham, P.; Chen, J.; Chen, Z.; Davis, A.; Dean, J.; Devin, M.; Ghemawat, S.; Irving, G.; et al. Tensorflow: A system for large-scale machine learning. In Proceedings of the 12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16), Savannah, GA, USA, 2–4 November 2016.
29. Sze, V.; Chen, Y.H.; Yang, T.J.; Emer, J.S. Efficient processing of deep neural networks: A tutorial and survey. *Proc. IEEE* **2017**, *105*, 2295–2329. [CrossRef]
30. Chollet, F. Keras. 2015. Available online: <https://keras.io> (accessed on 11 October 2019).

31. Tafazzoli, F.; Safabakhsh, R. Model-based human gait recognition using leg and arm movements. *Eng. Appl. Artif. Intell.* **2010**, *23*, 1237–1246. [[CrossRef](#)]
32. Zhang, R.; Vogler, C.; Metaxas, D. Human gait recognition at sagittal plane. *Image Vis. Comput.* **2007**, *25*, 321–330. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).