*electronics*

MDPI

*Article*

# Improved Deep Learning for Parkinson's Diagnosis Based on Wearable Sensors

**Jintao Yu [1,*], Ke Meng [1], Tingwei Liang [2], He Liu [1] and Xiaowen Wang [1]**

[1] Department of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China; mengke@s.hrbcu.edu.cn (K.M.); liuhe@s.hrbcu.edu.cn (H.L.); wangxiaowen@s.hrbcu.edu.cn (X.W.)
[2] Department of Aerospace, Harbin Institute of Technology, Harbin 150001, China; liangtingwei@hit.edu.cn
* Correspondence: 101129@hrbcu.edu.cn

**Abstract:** Parkinson's disease is a neurodegenerative disease that seriously affects the quality of life of patients. In this study, we propose a new Parkinson's diagnosis method using deep learning techniques. The method takes multi-channel sensor signals as inputs, and the full convolutional and LSTM blocks of the model perceive the same time-series inputs from two different views, and connect the extracted spatial features with temporal features. In order to improve the detection performance, a channel attention mechanism was incorporated into the model, and a data augmentation approach was used to eliminate the effect of unbalanced datasets on model training. The pd vs. hc and pd vs. dd classification tasks were performed, which improved accuracy by 4.25% and 8.03%, respectively, compared to the previous best results. Both improvements were higher than the previous methods using machine learning combined with feature extraction. To utilize the available data resources more effectively, this study conducted the pd vs. hc vs. dd triple classification task for the first time, which improved the model's ability to identify disease features. In that task, the accuracy rate reached 78.23%. The experimental results fully demonstrated the effectiveness of the proposed deep learning method for Parkinson's diagnosis.

**Keywords:** Parkinson's disease; deep learning; diagnosis; long short-term memory; attention mechanism



**Citation:** Yu, J.; Meng, K.; Liang, T.; Liu, H.; Wang, X. Improved Deep Learning for Parkinson's Diagnosis Based on Wearable Sensors. *Electronics* **2024**, *13*, 4638. https://doi.org/10.3390/electronics13234638

Academic Editor: José Carlos Castillo

Received: 18 October 2024
Revised: 21 November 2024
Accepted: 21 November 2024
Published: 25 November 2024



## 1. Introduction

In recent years, the incidence of age-related cognitive impairment has increased with the aging of the population [1]. Parkinson's disease (PD) is a widespread neurodegenerative condition that impacts the central nervous system. Characterized by its progressive nature, PD leads to the degeneration of neurons, particularly those producing dopamine, which plays a pivotal role in regulating movement. It is marked by a range of motor impairments, including rigid muscles, tremors in the hands, and a slowdown in movement known as bradykinesia. As the disease progresses, it severely affects patients' quality of life. While no therapies currently exist that can halt the neurodegenerative process or promote regeneration in Parkinson's disease, timely diagnosis and intervention are crucial. They play a significant role in mitigating the disease's impact and associated costs [2].

Indeed, the medical community faces significant challenges in identifying Parkinson's disease in its early stages. Even as their health may decline, individuals can enhance their quality of life significantly with prompt medical attention and appropriate interventions [3]. This is challenging because PD symptoms coincide with those of other diseases, putting PD at risk of being unrecognized or, worse, misdiagnosed. A score is then given based on the criteria of the Unified Parkinson's Disease Rating Scale (EADS) [4] and the Simpson–Angus Scale (SAS) [5].

In the last decade or so, as various sensors and electronic devices have become smaller in size, higher in performance, and lower in cost, these electronic components have become more widely used in everyday life [6]. Compared to the drawbacks of deploying external

devices to recognize the human body's activity status, an approach that is costly and poorly portable, wearable sensors can be conveniently used with integrated sensors to collect various behavioral data from the human body [7]. Wearable accelerometers or Inertial Measurement Units (IMUs) have been widely used in medical research [8,9], especially in situations where motion may be affected. Existing studies have shown that monitoring localized discharge signals can be effective in providing early warning of disease, and this approach has promising applications [10,11]. A multitude of data modalities have been meticulously analyzed in the context of Parkinson's disease (PD), encompassing a range of physiological and behavioral aspects. These include the assessment of hand movements, gait and balance, eye movements, and speech capture [12,13]. The majority of studies utilizing these modalities have reported high accuracy rates in differentiating individuals with PD from healthy controls (HC), a capability that is pivotal for the clinical application of such diagnostic tools [14,15]. This high accuracy in discrimination is a significant advancement, as it not only aids in the early identification of PD but also holds the potential to enhance treatment strategies and improve the quality of life for those affected by the disease [16]. Furthermore, the integration of multiple data sources, such as gait data combined with clinical and metabolomics information, has been shown to enhance the robustness of predictive models and improve diagnostic accuracy [17]. This comprehensive approach to diagnostics reflects the multifaceted nature of PD and underscores the value of multimodal data analysis in capturing the complexity of the disease. Current research has shown that analytics using speech capture or action-based tasks have an accuracy rate of over 90% [18,19].

Previous research has harnessed the power of smartphones to monitor daily movement patterns and to document short self-administered tasks like finger tapping or spiral drawing [20,21]. These studies have leveraged the ubiquity and convenience of smartphones to gather valuable health data in a non-invasive and user-friendly manner. By utilizing smartphones, researchers can collect a wealth of data on individuals' movement habits and motor functions over extended periods, providing insights into their health status and potentially aiding in the early detection of conditions like Parkinson's disease [22]. The ability to track these subtle motor symptoms over time could be particularly beneficial for monitoring disease progression and evaluating the effectiveness of treatments.

While the classification model excels at differentiating PD from healthy subjects, it may struggle to accurately classify other movement and neurological disorders. This limitation arises because the model captures general movement anomalies rather than the distinctive characteristics that define PD. In clinical reality, this is very important because neurologists may not know in advance whether a patient has PD or a similar movement disorder. Furthermore, while numerous past studies have relied on passive monitoring methods, incorporating interactive assessments could unveil the advantages of capturing subtle, often unnoticed phenomena. For instance, Parkinson's-disease-related tremors may emerge in conjunction with cognitive tasks or resurface when individuals perform actions such as lifting and sustaining the position of their arms [23–25].

The assessment and diagnosis of Parkinson's disease through sensor-collected feature data relies on manual analysis, and specialized and complex data analysis is not clinician-friendly; thus, there is a need to find an objective and accurate technique for automated data analysis. Machine learning methods are gradually being applied in the field of Parkinson's disease diagnosis due to their excellent performance in data analysis.

Lyons and Pahwa [26] utilized a wearable device to collect data, subsequently developing a sophisticated classification system designed to differentiate between essential tremor and tremors associated with Parkinson's disease. Kostikis et al. [27] introduced a practical smartphone-based tool designed to accurately evaluate upper limb tremors in individuals with PD. When the phone was held or securely attached to the subject's hand, it captured motion data that was subsequently used to compute a comprehensive set of metrics. Sajal et al. [28] used a smartphone with a built-in triaxial accelerometer sensor to capture resting tremor data and speech data from patients diagnosed with PD and healthy

individuals. The methodology involved in this study utilized resting tremor and vowel articulation data from patients who had been previously diagnosed and labeled using the Unified Parkinson's Disease Rating Scale (UPDRS), as well as data from healthy individuals. This comprehensive dataset was employed to train and fine-tune machine learning (ML) models, aiming to enhance their ability to accurately detect Parkinson's disease.

Although sensor- and machine-learning-based research has yielded some results, there are currently some drawbacks. Traditional machine learning algorithms rely heavily on feature engineering, and complex feature engineering has limitations for clinical dissemination [29]. Second, machine learning algorithms also have their own shortcomings in specific applications. For example, the KNN algorithm is weak in processing nonlinear data and has a high computational cost when analyzing complex data; the decision tree algorithm is more sensitive to sample imbalance, and is prone to favor the majority class of samples in the training process, leading to overfitting; the SVM algorithm is sensitive to parameter training and kernel function, making it difficult to cope with the task of a larger number of samples [30].

The advent of deep learning effectively addressed the limitations of traditional machine learning, which often relied on intricate feature engineering. Deep learning models could process data in a more straightforward and efficient manner, eliminating the need for manual feature extraction. Currently, deep learning models had been extensively employed in the evaluation and diagnosis of Parkinson's disease. The primary areas of research in this context included early screening, the assessment of movement disorders, and pathological analysis of the disease [31]. For instance, Qin et al. [32] harnessed the power of surface electromyography (sEMG) to quantify the severity of tremors in Parkinson's disease patients. They utilized a Convolutional Neural Network (CNN) to learn the patterns and similarities within sEMG signals associated with tremors. In their study, they also compared several traditional machine learning models. This demonstrated the potential of deep learning to provide more accurate and nuanced assessments of tremor severity in Parkinson's disease, offering valuable insights for clinical diagnosis and treatment planning.

The Long Short-Term Memory Fully Convolutional Network (LSTM_FCN) was a model for time-series classification proposed by Karim et al. [33]. The model combined the deep learning model structures of LSTM and FCN and could enhance the FCN model to significantly improve its performance by a nominal increase in the number of parameters. There have been studies that could be used for ECG data classification and have achieved better results, relative to other time-series classification models. Therefore, the model given above was used in this study to extract features for Parkinson's disease detection for classification study.

## 2. Materials and Methods

### *2.1. Dataset*

In this study, the Parkinson's disease smartwatch (PADS)dataset (https://physionet.org/content/parkinsons-disease-smartwatch/1.0.0/ (accessed on 12 June 2024)) was used, and the data pertained to 469 individual cases [34]. The data presented in Table 1 categorize participants into three distinct groups: (1) Parkinson's disease (PD) patients, who encompass a broad spectrum of ages and diagnostic stages, (2) differential diagnosis (DD) subjects, which include those with primary tremor, atypical Parkinsonian disorders, and (3) healthy controls (HC). To ensure a balanced comparison, the healthy controls were carefully matched in age to the PD group. The dataset comprised a total of 276 PD patients, 79 healthy individuals, and 114 patients with differential diagnoses. This stratification allowed for a comprehensive comparison of the different groups, enabling researchers to analyze the distinct characteristics of each population and assess the effectiveness of the diagnostic models in distinguishing between them. By including these diverse groups, the study aimed to provide a more nuanced understanding of Parkinson's disease and its related movement disorders.

**Table 1.** Participant samples including Parkinson's disease (PD), healthy controls (HC), and differential diagnosis (DD).

| Class | Participant Group | Sample Size (Male, Female) | Age Years (Std) |
|---|---|---|---|
| PD | Parkinson's disease | 276 (195, 81) | 65.4 (9.6) |
| HC | healthy controls | 79 (29, 50) | 62.9 (12.5) |
| DD | differential diagnosis | 114 (57, 57) | 62.4 (11.5) |

Data for the experiment were collected through sensorimotor recordings during an interactive assessment that encompassed 11 distinct neuromotor steps. The assessment steps were designed by movement disorder specialists with the main aim of creating easy-to-understand checks and capturing the most relevant motor features, and each assessment took approximately 15 min. During the data collection phase, participants were seated comfortably in an armchair and equipped with two smartwatches (Apple Watch Series 4 (Apple Inc., Cupertino, CA, USA)), with one on each wrist. Throughout the assessment, these smartwatches simultaneously captured acceleration and rotation data along the three spatial axes ($x$, $y$, $z$) at a frequency of 100 Hz. The interactive evaluation consisted of eleven steps, with three steps lasting for 20 s each and the remaining eight steps lasting for 10 s each, as detailed in Table 2.

**Table 2.** Smartwatch-based assessment steps.

| Steps | Durations | Description | Task Category |
|---|---|---|---|
| 1a | 20 | Resting with closed eyes while sitting. | Resting |
| 1b | 20 | Resting while the patient is calculating sevens. | Resting |
| 2 | 10 | Lift and extend arms. | Postural |
| 3 | 10 | Arms remain lifted. | Postural |
| 4 | 10 | Hold one-kilogram weight in each hand for 5 s. | Postural |
| 5 | 10 | Point index finger to the examiner's lifted hand. | Kinetic |
| 6 | 10 | Drink from glass. | Kinetic |
| 7 | 10 | Cross and extend both arms. | Kinetic |
| 8 | 10 | Bring both index fingers to each other. | |
| 9 | 10 | Tap own nose with index finger. | Kinetic |
| 10 | 20 | Entrainment. The examiner stomps on the ground, setting the pace. | Postural |

Steps: lists the different activities in the experiment. Durations: the duration of each step, which may be in seconds. Description: a description of the specific actions or requirements for each step. Task category: "Resting", "Postural", and "Kinetic". These steps are designed to assess a subject's resting state, postural control, and motor function through a series of movements, which can be useful in diagnosing diseases such as Parkinson's disease.

To maintain consistency in the duration of each segment, the 20 s recording was divided into two parts, with each participant having a total of 14 10 s time series. With this setup, each participant acquired 168 channels of time-series data (14 recordings $\times$ 2 arms $\times$ 2 sensors $\times$ 3 axes = 168). Following a thorough analysis of the dataset by Julian Varghese et al. [34], it was determined that three specific evaluation steps did not contribute significantly to the classification process and were therefore eliminated: step 3, which involved "lifting and holding" the arms; step 5, which required "pointing a finger"; and step 8, which entailed "touching the index finger". This decision was made to streamline the dataset and focus on the steps that provided the most relevant information for distinguishing between Parkinson's disease patients and other groups [35]. After removing the redundant evaluation steps, the time-series data corresponding to each participant became 132 channels.

Since the experimental equipment collects both acceleration and gyroscope data, there was an opportunity to explore the impact of different sensor types on the evaluation task. Specifically, three different experimental configurations were evaluated in this study:

1. A setup that relied solely on acceleration data (acc);
2. A setup that utilized only rotation data (rot);

3. A comprehensive setup that integrated data from both sensors to provide a more nuanced interpretation of the data (both).

### *2.2. Preprocessing*

In data processing, the time series of the first 50 waiting sampling points that did not have contrasting characteristics were cropped. Initially, the acceleration data underwent preprocessing to reduce the impact of the Earth's gravitational field. A technique akin to segmented linear smoothing, known as l1-trend filtering, was employed to adjust for variations in the direction of gravity [36].

Since the dataset used in this study was unbalanced, with much more data collected from PD patients than from HC and DD populations, data augmentation techniques were applied during training. Data augmentation is a technique that enriches the dataset by creating additional, varied instances from the original data, preserving the labels. This method helped to uncover unseen input patterns, curb overfitting, balance class distributions within the dataset, and bolster the model's ability to handle diverse real-world inputs, thereby enhancing its overall robustness and generalization capabilities [37].

Data enhancement techniques were prevalent in domains like image recognition, where minor variations such as dithering, scaling, cropping, distortion, and rotation naturally occur during observation. These methods, which do not alter the original image's category labels, are integral to image data augmentation, as referenced in [38]. Extending this concept to time-series data, most enhancement strategies rely on random transformations of the training dataset. This includes the introduction of random noise [39], random cropping [40], random scaling [41], and both random warping and frequency warping along the time dimension [42], all aimed at enriching the dataset without changing its underlying labels. Similar to Um et al. [43], random time bending and axis rotation were performed, and the enhancement process is shown in Figure 1. The waveforms of the data before and after enhancement are shown in Figure 2. To maintain the authenticity of the PD samples in the experiment, no data augmentation was applied to them. To achieve a relatively balanced state in the sample data, the HC was expanded by a factor of two, and the DD was expanded by a factor of one. Therefore, there are now 276 PD, 237 HC, and 228 DD samples.

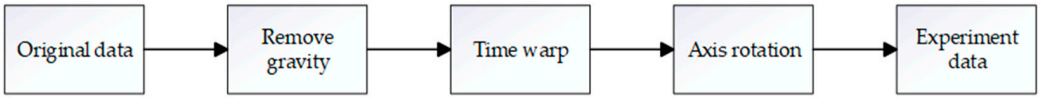



**Figure 1.** Data augmentation pipeline.

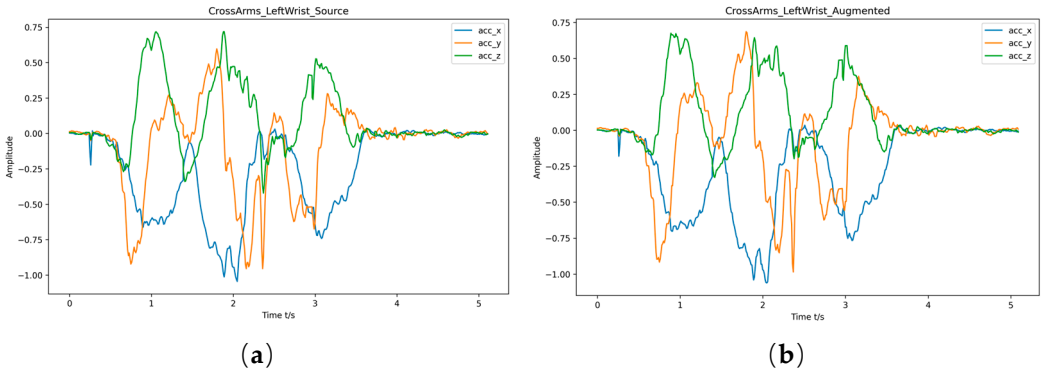



**Figure 2.** Enhancement before and after comparison chart. (**a**) Raw data; (**b**) enhanced data.

### *2.3. Model*

In order to detect Parkinson's disease patients from the participants, a novel deep learning model for sensor data was proposed. The model combined LSTM_FCN and a channel attention mechanism. Temporal convolution, a technique that involved applying

convolutional operations along the time axis, had proven to be an effective learning model for time-series classification problems [44]. Temporal convolution was an effective method for time-series classification. It used convolution along the time axis. Fully convolutional networks with temporal convolution served as feature extractors [45]. The fully connected (FC) layer flattened the convolutional feature maps into a one-dimensional sequence, which aided in connecting to activation functions for category prediction. All the FC layers in FCN were replaced by the Global Average Pooling (GAP), which made the input size of the network unrestricted and strengthened the network's generalization ability. In addition, the structure of FCN made it more competent in time-series classification tasks. Since time-series data differed greatly in dimension and length from image inputs, FCNs could efficiently retain and extract complex information from time series to prevent the loss of key features [46]. The structure of FCN is shown in Figure 3.

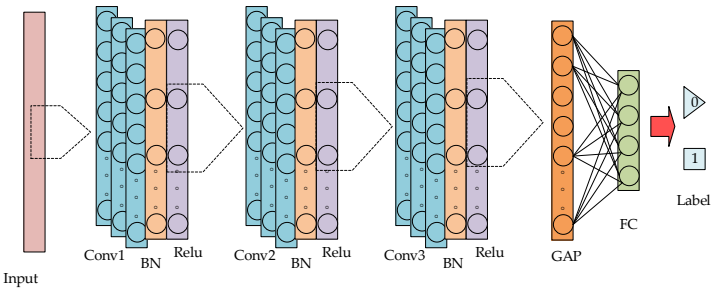



**Figure 3.** The structure of FCN. The output from the GAP layer is fed into the Softmax function. All convolutional layers in the network utilize one-dimensional convolution kernels. In the figure, 'Conv' denotes spatial convolution.

The abovementioned convolutional neural networks focused on extracting spatial features, while temporal features for time-series classification also had important information, of which LSTM was the classical network for extracting temporal features [47]. LSTM evolved from Recurrent Neural Networks (RNN), and both were built on the core idea of preserving the historical information of the network [48]. As the name suggested, RNN was characterized by a cyclic recursive structure that allowed it to preserve the information of all past nodes, and for the current node, it could obtain historical data from both the previous nodes and the sibling nodes, which ensured the reliability of its updates. However, it was obvious that since all the data were saved, when the amount of information and the network gradually became more and more complex, the model overload would inevitably produce problems such as gradient explosion or gradient disappearance. In addition, the redundant historical data on the node's updating choice was also a big test. In order to solve the abovementioned problems, LSTM was proposed as a structure for data selection, i.e., filtering and saving valid data in the redundant historical information; the core of processing data selection was the forgetting gate, which could decide what information the model discarded in the process of memorization.

The Channel Attention Mechanism (CAM) was a feature relabeling technique used in deep learning models, especially in Convolutional Neural Networks (CNN) [49]. This mechanism aimed to enhance the model's ability to capture critical information by learning the importance weights of different feature channels. In the structural diagram of Figure 4, first, a global average pooling operation was applied to the feature map of each channel to compress the spatial dimensions (height and width) of each channel into a single value. This operation could be seen as an aggregation of spatial information for each channel to obtain a global representation. These global representations were then nonlinearly transformed using a fully connected layer (also called a dense layer). Typically, this fully connected layer was first downscaled and then upscaled back to the original number of channels. This was undertaken in order to learn the weights of each channel, i.e., the importance of each channel to the final task. After the fully connected layer, a Sigmoid activation function was

used to ensure that the weights of each channel were between 0 and 1. In this way, the model learned the relative importance of each channel.

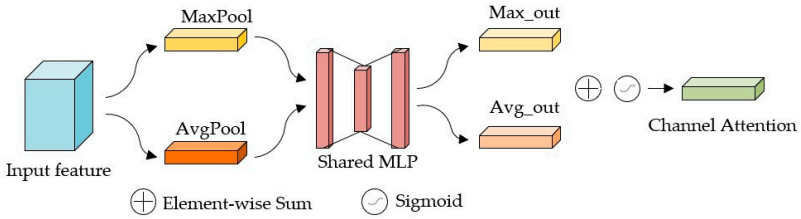



**Figure 4.** Structural diagram of channel attention.

As depicted in Figure 5, the full convolutional block and the LSTM block processed the same time-series input but from two distinct perspectives. The full convolutional block was composed of three sequentially arranged temporal convolutional blocks, each equipped with filters with 128, 256, and 128 layers, respectively. Within each block, a temporal convolutional layer was paired with batch normalization [50]. Feature extraction was accomplished by passing the data through a third convolutional block. The feature-extracted data were multiplied with the channel attention weights to incorporate the channel attention into the global features. At the same time, the time-series input was transmitted to the dimensional blending layer. Following the dimensional blending transformation of the time series, the data were fed into the LSTM block. This particular block in the model architecture consisted of a standard LSTM layer, which was adept at capturing temporal dependencies within the data. It was succeeded by a dropout layer, a regularization technique that was crucial for enhancing the model's generalization capabilities. This random deactivation forced the network to develop redundant representations and prevented it from becoming too reliant on any particular set of neurons. The outputs from both the FCN and the LSTM blocks were then concatenated. This combined data stream was subsequently forwarded to the SoftMax classification layer, which assigned a probability distribution over the predicted output classes, facilitating the final classification decision.

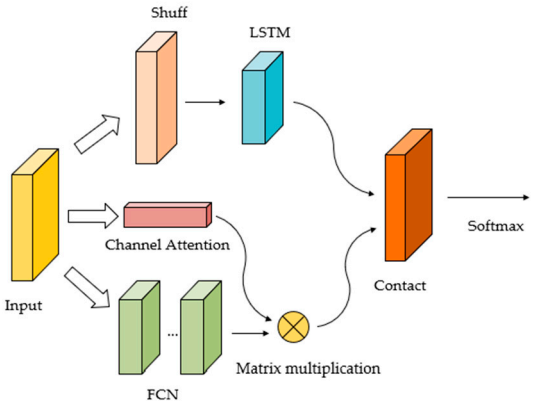



**Figure 5.** Architecture of CA_LSTM_FCN.

In this study, the proposed model was designed to process sensor data with time-series characteristics and was able to capture the temporal dynamics and spatial features of the data. Features were extracted from the time-series data using LSTM, which captured the dynamic characteristics of the data over time. With FCN, the model was able to extract spatial features from sensor data, including local patterns and global structural information. The time-series features extracted by LSTM and the spatial features extracted by FCN were fused to obtain a more comprehensive data analysis. This fusion not only enhanced the model's ability to capture temporal dynamics but also improved its ability to recognize spatial features, enabling the model to consider both the temporal and spatial

dimensions of time-series data. In order to allow the model to adaptively focus on key parts of the sequence, an attention mechanism was introduced. This mechanism allowed the model to dynamically adjust its attention to different time steps, thus improving the ability to recognize critical information. Through the abovementioned feature extraction and processing steps, the model was able to extract rich temporal dynamics and spatial features from the sensor data with time-series characteristics, providing strong support for subsequent analysis and prediction tasks.

### *2.4. Evaluate*

In addressing the binary classification problem at hand, this study opted for the Binary Cross Entropy Loss function, specifically the BCEWithLogitsLoss. This loss function was particularly efficient as it amalgamated the sigmoid activation function with the binary cross-entropy loss into a single operation. This combination allowed for a streamlined optimization of the model's parameters by minimizing the difference between the predicted probabilities and the actual target values.

The model parameters were refined through the use of the Adam optimizer, which was set with a learning rate (lr) of 0.001 [51]. The Adam optimizer stood out for its ability to automatically adjust the learning rate based on the estimated first and second moments of the gradients, thereby combining the benefits of the Momentum and RMSprop (Root Mean Square Propagation) algorithms. This adaptive feature made Adam highly versatile, as it could adjust to the varying demands of different stages of training. By doing so, Adam helped the model to converge more effectively, enhancing its ability to learn from the data and make accurate predictions.

To evaluate the effectiveness of the classifier, several key metrics were used: accuracy, precision, recall, and F1 scores [52]. These metrics were crucial for understanding how well the model performed in distinguishing between Parkinson's disease samples (considered positive) and other samples (considered negative).

Accuracy was a metric that indicated the percentage of all samples—both positive and negative—that were correctly identified by the model. It provided a broad view of how well the model was performing across the board, capturing its ability to make correct predictions for the majority of the data points it was evaluated against.

Precision was a measure of the accuracy of the positive predictions. It indicated the proportion of samples predicted as positive that were actually positive, thus providing insight into the model's ability to correctly identify true PD cases without a high rate of false positives.

Recall, or sensitivity, was the metric that showed how well a model could find all the positive instances in the data. A high recall meant the model did a good job of identifying actual positive cases without missing too many.

The F1 score was a blend of precision and recall, giving a balanced measure that considered both types of errors. It was very handy when there was an uneven number of instances in different classes, giving a fuller picture of the model's performance than either precision or recall by themselves.

By calculating these metrics, it was possible to obtain a full picture of the strengths and weaknesses of the model and make an informed decision about its suitability for practical application in the diagnosis of Parkinson's disease. Higher values across all these metrics indicated a more effective and reliable model.

$$\text{accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{1}$$

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{2}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{3}$$

$$\text{F1 score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (4)$$

In summary, the data from the dataset were organized by category. Estimated low-frequency gravity trends were subtracted from the signal data to remove offsets in the physiological data. The experimental samples contained enhanced data and labels. Temporal dynamics and spatial features were extracted using the proposed CA_LSTM_FCN model. The PyTorch (2.2.2) framework was used for deep learning model development and Sklearn (1.3.2) for model evaluation. A 5-fold cross-validation was performed to ensure a balanced proportion of categories in each fold. During training, losses were recorded, and the model performance was evaluated after each epoch. The model was evaluated on the validation set, and the accuracy, precision, recall, and F1 score were calculated. An early stop mechanism was implemented to stop training early if the performance was not improved in multiple consecutive epochs.

## 3. Results

To confirm the effectiveness of the LSTM_FCN network enhanced with the channel attention mechanism in this study, experiments comparing Parkinson's disease vs. healthy controls (pd vs. hc) and Parkinson's disease vs. differential diagnosis (pd vs. dd) on accelerated (acc), rotated (rot), and combined datasets (both) were performed. The outcomes of these experiments for the aforementioned tasks are documented in Table 3.

**Table 3.** Performance of the CA_LSTM_FCN model on the dataset.

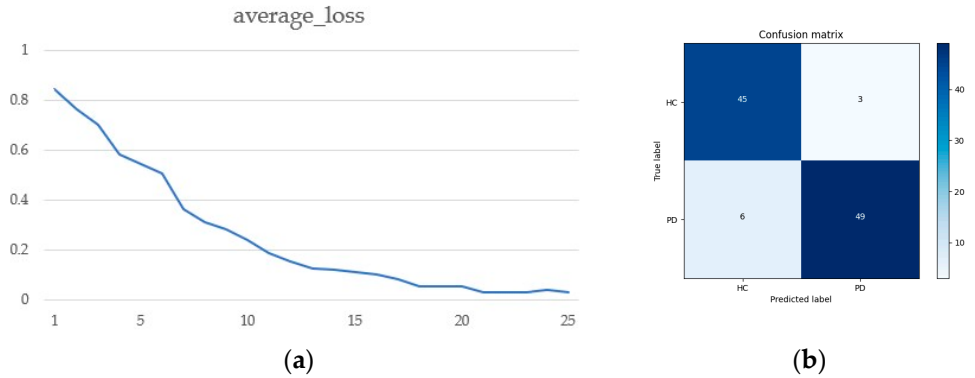
| Task | Dataset | Accuracy | Precision | Recall | F1 |
|---|---|---|---|---|---|
| pd vs. hc | acc | 81.55% | 83.33% | 81.82% | 83.25% |
| | rot | 88.35% | 97.78% | 80.13% | 88.63% |
| | acc + rot | 92.15% | 95.08% | 89.73% | 91.24% |
| pd vs. dd | acc | 72.28% | 72.13% | 80.00% | 76.58% |
| | rot | 76.24% | 75.41% | 83.64% | 79.11% |
| | acc + rot | 85.38% | 88.75% | 83.42% | 85.51% |

The classification task of pd vs. hc is a set of comparison experiments chosen by many researchers to distinguish people with Parkinson's disease from the normal population. The comparative analysis of the experimental results showed that our proposed model achieved a performance of more than 80% across all datasets, especially on both datasets, with an accuracy of 92.15%. Although the precision was reduced by 2.7% compared to the rot set performed, the recall was improved from 80% to nearly 90%, improving the performance of predicting Parkinson's disease samples as positive, which was the best among the three data subsets.

Compared to the pd vs. hc classification task, pd vs. dd may be a little harder, but it is not as easy to distinguish between PD and HC populations with large behavioral differences. Upon analysis, the data presented in Table 3 indicate that the model's performance in this particular task did not match the efficacy observed in pd vs. hc. Nonetheless, the most promising experimental outcomes were achieved when evaluating both datasets, yielding an accuracy of 85.38%, a precision of 88.75%, a recall of 83.42%, and an F1 score of 85.51%. These figures underscore the model's capability, albeit noting room for improvement in certain classification scenarios. Figures 6 and 7 show the validation loss function and confusion matrix for the two classification tasks on the acc and rot sets, respectively.

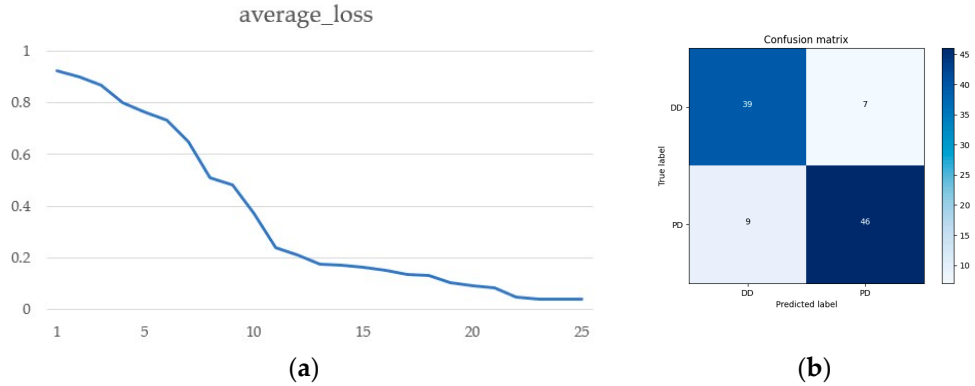
Overall, the model showed the best performance in detecting Parkinson's disease samples on both sets. Compared to using a single sensor dataset, the model proposed in this paper could capture body movements and vibrations more comprehensively on data combining both sensors, and this multidimensional data fusion could lead to a more comprehensive understanding of the characteristics of tremor in Parkinson's patients, thus improving the accuracy of symptom detection. Despite achieving better performance

across all metrics, the use of both sets led to an increase in data processing time while improving metrics.




(**a**) (**b**)

**Figure 6.** Pd vs. hc: (**a**) loss function; (**b**) confusion matrix.




(**a**) (**b**)

**Figure 7.** Pd vs. dd: (**a**) loss function; (**b**) confusion matrix.

The performance of our proposed model was benchmarked against previous outcomes on the same dataset. Julian Varghese et al. [34] employed a variety of machine learning techniques, including support vector machines (SVM), CatBoost, and neural networks (NN) for this dataset. They utilized both manually extracted features and the BOSS (Bag-of-SFA-Symbols) algorithm for automated feature extraction. Our model's results were then compared to the outcomes achieved by these methodologies to evaluate its effectiveness and potential advantages. This comparison helped to situate our approach within the context of existing research and underscored any improvements in accuracy, efficiency, or scalability that our model may offer.

As shown in Figure 8, for the pd vs. hc task, the network model effect of our proposed CA_LSTM_FCN was higher than that presented by Julian Varghese et al., who combined a machine learning approach with feature extraction in terms of accuracy and precision; the results were only slightly lower than the other methods in terms of recall and F1 value. From Figure 9, it can be noticed that the performance of the proposed method in this paper exceeded the previous results in both cases, with an improvement of nearly 20% in recall and more than 10% in other metrics.

However, the previous two dichotomous tasks, which limited the comparisons to two-by-two, amounted to classification experiments under a priori conditions. Therefore, a new experiment was added to this, namely pd vs. hc vs. dd. This task puts together samples from all the participants in the dataset in a three-classification experiment. It allows for a better diagnosis from an unqualified population to see if they are likely to have Parkinson's disease, in line with the intended effect of this project.

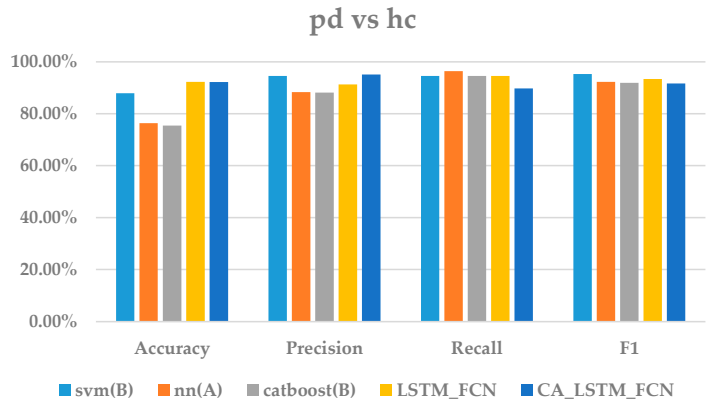



**Figure 8.** Comparison chart of several methods for the pd vs. hc task. A represents a human-defined feature extraction method, and B represents feature extraction using the BOSS algorithm.

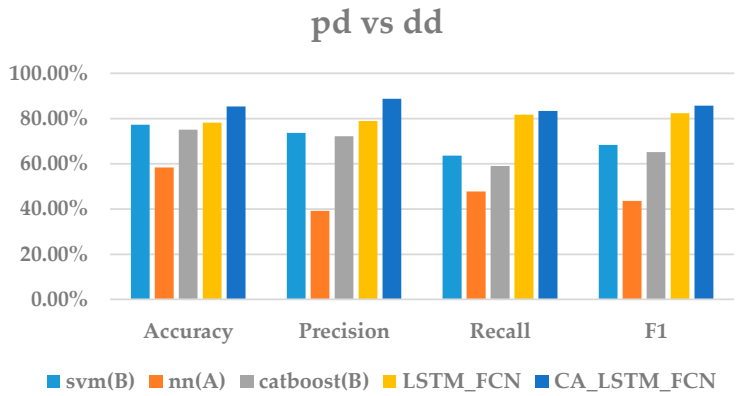



**Figure 9.** Comparison chart of several methods for the pd vs. dd task.

In this experiment, five-fold cross-validation was incorporated into the hyperparameter optimization process to help select hyperparameters that perform well in different subsets of the data, thereby improving the robustness of the model. Hyperparameter optimization plays a pivotal role in the performance of neural networks, but it can be a complex and time-consuming process. Traditionally, experts rely on their expertise and trial-and-error methods to find the best hyperparameter settings. However, an automated approach to searching for optimal neural network architectures offers a more systematic alternative to this hit-or-miss method.

In this study, the hyperparameter optimization task was simplified by implementing a regularized evolutionary algorithm [50]. This approach is adept at thoroughly searching the hyperparameter space, identifying configurations that significantly enhance the model's performance. During the learning process, Adam's algorithm was chosen, which is known for its fast convergence and high accuracy in all situations.

There are a few key hyperparameters that require particular attention throughout the tuning process, given as follows:

- The number of hidden layers within the network, which influences the model's ability to grasp intricate patterns;
- The count of RNN layers, essential for capturing the temporal dynamics within sequential data;
- The dropout rates for both the RNN layers ('rnn_dropout') and the fully connected layers ('fc_dropout'), which play a key role in mitigating overfitting by preventing the network from becoming too reliant on any specific weights during training.

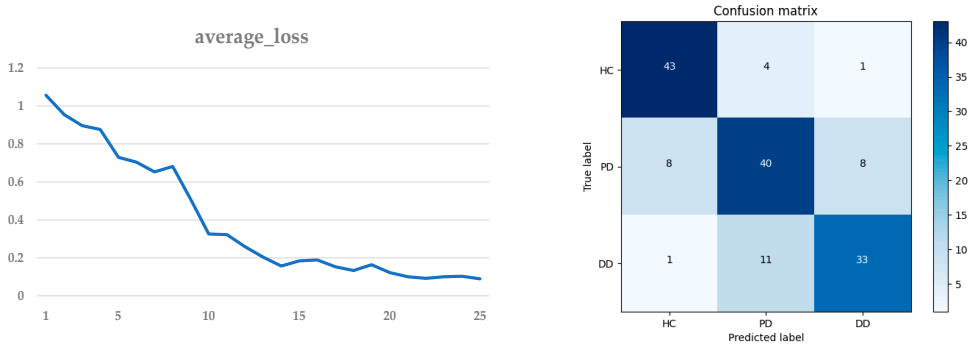
By carefully tuning these hyperparameters, the aim was to fine-tune the model for optimal performance, ensuring that it generalizes well from the training data and makes

accurate predictions for unseen data. Table 4 presents a list of the hyperparameters considered in the study, as well as an exploration of the range of values for each parameter. By systematically tuning these hyperparameters, it was a matter of finding the optimal configuration that delivered the best performance for the neural network model.

**Table 4.** Hyperparameters that were selected for optimization.

| Hyperparameter | Available Option |
|---|---|
| Hidden_size | 64, 128 |
| Lstm_layers | 1, 2 |
| Lr | 0.001 |
| Weight_decay | $10^{-5}$ |
| Lstm_dropout | 0.2, 0.5, 0.8 |
| Fc_dropout | 0.2, 0.5, 0.8 |

After experimental validation, the best results were obtained on both sets with 128 hidden layers and two LSTM layers, with a dropout rate of 0.2, where the accuracy was 78.23%, the precision was 82.13%, the recall was 75.71%, and the F1 value was 76.02%. Figure 10a,b show the loss function and confusion matrix, respectively. As the training loss decreased, the model became increasingly able to capture patterns and relationships in the training data, which helped avoid overfitting. Ideally, if the training loss decreased along with the validation loss, this might indicate that the model not only performed well on the training data but was also able to generalize to unseen data. A decrease in training loss was usually accompanied by an increase in model accuracy because the model was better at predicting the training data, as shown in Figure 11.




(**a**) (**b**)

**Figure 10.** Pd vs. hc vs. dd: (**a**) loss function; (**b**) confusion matrix.

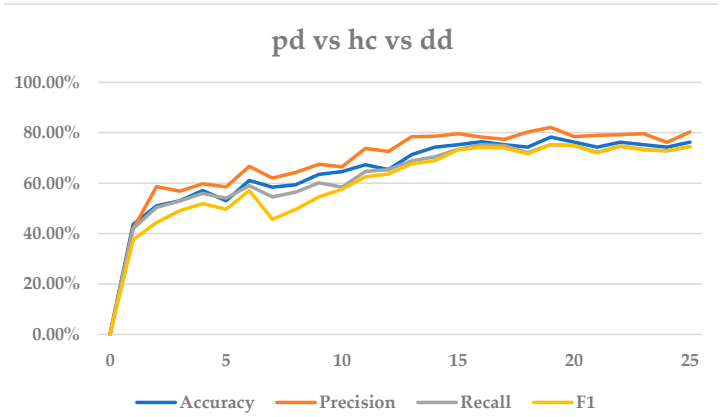



**Figure 11.** Metrics change curve. The performance of each metric with rounds is shown, and the stable metrics are reached after about 15 epochs are performed; as the rounds are performed, they occasionally drop slightly.

Boxplots are a very useful tool to help us quickly understand the performance distribution of different models and to make visual comparisons. Figure 12 shows the results obtained from different deep learning models in the analysis of various metrics. The comparison showed that the CA_LSTM_FCN model proposed in this paper also performed well in the triple classification task, which was basically higher than the model before improvement. The performance of the model was stable in different situations, with better central tendency and stability in all metrics.

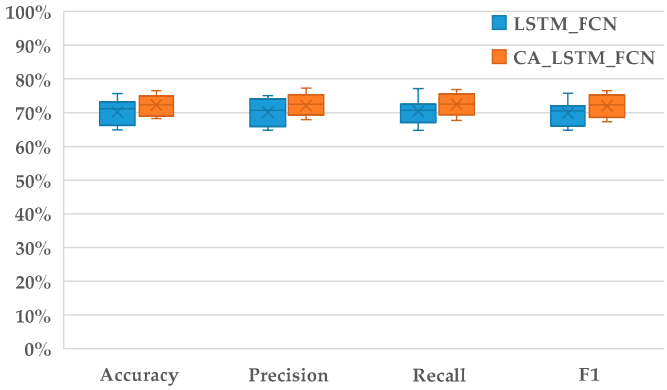



**Figure 12.** Comparison of CA_LSTM_FCN and LSTM_FCN.

## 4. Discussion

This study centered on determining the feasibility of employing sensor data in conjunction with deep learning models to detect Parkinson's disease within a given population. The experiment utilized sensor data extracted from the Pads dataset. In this study, two binary classification tasks were first set up, namely pd vs. hc and pd vs. dd. The sensor data from the interaction assessment performed by each person were connected along the *XYZ* axis of both hands, and the connected long-series data were input into the improved model in this paper. The experimental results showed that the model achieved good results, with accuracy rates of 92.15% and 85.38% for the two tasks, respectively. Notably, the performance of our model, although similar to the previous methods for the pd vs. hc task, was significantly higher than that for the pd vs. dd task, with an improvement of about 10%, which indicated that the model could successfully recognize Parkinson's patients.

LSTM was able to detect sudden changes in acceleration and angular velocity data and could recognize changes in the state of motion, such as changes in the rate of change or direction of acceleration [53]. The LSTM part had the ability to capture long-term dependencies in a time series and to understand dynamic changes and trends in time-series data. FCN extracted local texture and shape features in acceleration and angular velocity data through convolutional layers to recognize local changes in motion, such as local patterns of acceleration changes. With a deep convolutional network structure, FCN was able to capture a wider range of contextual information, which was important for understanding the spatial distribution of acceleration and angular velocity data [54]. To improve the detection performance, the model added a channel attention module to the convolutional network module, which helped the model automatically learn which channels (features) were important and which were not. For time-series data, the importance of the features may change at different time points.

In this study, based on the sensors, the datasets could be categorized into the acceleration dataset, the gyroscope dataset, and both datasets. The proposed model was validated on different datasets, and it was found that the best performance was obtained on the dataset that combined acceleration and gyroscope. It showed that the proposed model could work well in extracting the most critical features from the data from different sensors compared to other methods, which in turn improved the accuracy of the classification task. Enhancing a model's capabilities was achieved by integrating data from multiple sensors,

rather than depending on a single type of data. In deep learning, there is a prevailing notion that performance is directly proportional to the amount of data. However, this notion is based on the premise that the data and their corresponding labels are impeccable. In this study, the problem of limited data availability was addressed by utilizing data from multiple sensors to improve the overall performance of the model.

So far, most of the researchers studied the pd vs. hc task, and few of them studied pd vs. dd; there are almost no experiments to classify the three objects together. Therefore, on this dataset, all samples were put together to further validate the proposed model. The triple classification model needed to process more category information during the training process, which helped to improve the generalization ability of the model and made it perform better when facing new and unseen data.

Data collected by motion sensor devices, combined with deep learning models, could better differentiate between clinically diagnosed and pre-diagnosed Parkinson's disease. Utilizing this type of equipment allowed for passive, continuous collection to obtain reliable estimates of a person's motor function and ability, and to detect subtle changes early [55]. This approach was able to identify specific patterns associated with motor acceleration that were relevant to the future onset or existing confirmed diagnosis of Parkinson's disease. Deep learning techniques were able to process large-scale sensor data and extract features from it that were useful for PD diagnosis, which was crucial for early diagnosis and treatment.

The current study successfully showed that the CA_LSTM_FCN model was capable of identifying Parkinson's disease patients from sensor data. However, one significant limitation was that the experiments were conducted using only a few minutes of data. To address this, future work should aim to extend the data collection period to hours or even days. This would allow for a more comprehensive capture of participants' movements and pattern changes, providing a richer dataset for the classification model.

By collecting data over a longer period of time, it was possible to further assess the accuracy of the data collected by the model at different times and to study how its performance changes over time. This was crucial for understanding the model's reliability and consistency in real-world scenarios. Additionally, conducting longer experiments enabled the collection of data outside the controlled laboratory environment. Therefore, it was essential to evaluate the model's performance in more realistic settings to ensure its applicability in practical situations.

In summary, extending the duration of data collection and conducting experiments outside the laboratory would be important steps for future research. This would not only enhance the model's generalizability but also provide valuable insights into its long-term accuracy and reliability in detecting Parkinson's disease from sensor data.

## 5. Conclusions

In this paper, a new deep learning method CA_LSTM_FCN was proposed, which utilizes convolutional neural networks to extract spatial features and LSTM to extract temporal features to identify Parkinson's patients from acceleration and gyroscope data obtained from sensors. The model demonstrated impressive performance, with an accuracy of 92.15% for pd vs. hc, and an accuracy of 85.38% for pd vs. dd. Furthermore, it achieved a commendable accuracy of 78.23% in the more complex triple categorization task, which involved distinguishing between PD, HC, and DD. These results were particularly significant for accurately identifying diseases with subtle distinctions. The experimental results showed that the algorithm was effective and outperformed the combination of machine learning classifiers and feature extraction used by previous authors.

The method presented in this paper offers a promising approach for diagnosing Parkinson's disease. Wearable devices could provide continuous, objective data that could aid in early diagnosis and more accurate assessment of the condition. The effectiveness of the method could be improved by combining it with other sophisticated deep learning techniques. These advances may expand their applications beyond their current scope

to various areas of healthcare and biomedical research. Data collected through wearable devices could provide Parkinson's patients with more personalized treatment plans, thereby improving their quality of life.

**Author Contributions:** Conceptualization, all authors; methodology, J.Y. and K.M.; software, K.M.; validation, J.Y. and K.M.; formal analysis, all authors; investigation, J.Y.; resources, T.L.; data curation, X.W.; writing—original draft preparation, J.Y. and K.M.; writing—review and editing, all authors; visualization, T.L.; supervision, H.L.; project administration, J.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research is supported by National Natural Science Foundation of China (Grant Nos. 12422213, 12372008); the National Key R&D Program of China (Grant No. 2023YFE0125900), the Natural Science Foundation of Heilongjiang Province (Grant No. YQ2022A008), the Basic Research Programs of Heilongjiang Provincial Universities (Grant No. 2023KYYWF0980).

**Data Availability Statement:** Data are contained within the article.

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

## References

1. Mai, A.S.; Deng, X.; Tan, E.K. Epidemiology of early-onset Parkinson disease (EOPD) worldwide: East versus west. *Park. Relat. Disord.* **2024**, 107126. [CrossRef] [PubMed]
2. Cabestany, J.; Suppa, A.; Ólaighin, G. Editorial: Parkinson's disease: Technological trends for diagnosis and treatment improvement. *Front. Neurol.* **2023**, *14*, 1151858. [CrossRef] [PubMed]
3. Carvajal-Castaño, H.A.; Pérez-Toro, P.A.; Orozco-Arroyave, J.R. Classification of Parkinson's Disease Patients—A Deep Learning Strategy. *Electronics* **2022**, *11*, 2684. [CrossRef]
4. Goetz, C.G.; Tilley, B.C.; Shaftman, S.R.; Stebbins, G.T.; Fahn, S.; Martinez-Martin, P.; Poewe, W.; Sampaio, C.; Stern, M.B.; Dodel, R. Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. *Mov. Disord. Off. J. Mov. Disord. Soc.* **2008**, *23*, 2129–2170. [CrossRef] [PubMed]
5. Simpson, G.M.; Angus, J.W.; Frc, P. A rating scale for extrapyramidal side effects. *Acta Psychiatr. Scand.* **1970**, *45*, 11–19. [CrossRef]
6. Li, S.; Wu, Y.; Asghar, W.; Li, F.; Zhang, Y.; He, Z.; Liu, J.; Wang, Y.; Liao, M.; Shang, J.; et al. Wearable Magnetic Field Sensor with Low Detection Limit and Wide Operation Range for Electronic Skin Applications. *Adv. Sci.* **2024**, *11*, e2304525. [CrossRef]
7. Li, B.; Yao, Z.; Wang, J.; Wang, S.; Yang, X.; Sun, Y. Improved Deep Learning Technique to Detect Freezing of Gait in Parkinson's Disease Based on Wearable Sensors. *Electronics* **2020**, *9*, 1919. [CrossRef]
8. Talaei, F.; Kargar, S.M. Design and Fabrication of a Device for Reducing Hand Tremor in Parkinson Patients during Eating. *J. Med. Signals Sens.* **2023**, *13*, 21–28. [CrossRef]
9. Vidya, V.; Poornachandran, P.; Sujadevi, V.; Dharmana, M.M. IMU sensor based self stabilizing cup for elderly and parkinsonism. In Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), Udupi, India, 13–16 September 2017; pp. 2264–2269.
10. Bagheri, E.; Jin, J.; Dauwels, J.; Cash, S.; Westover, M.B. A fast machine learning approach to facilitate the detection of interictal epileptiform discharges in the scalp electroencephalogram. *J. Neurosci. Methods* **2019**, *326*, 108362. [CrossRef]
11. Hirano, R.; Asai, M.; Nakasato, N.; Kanno, A.; Uda, T.; Tsuyuguchi, N.; Yoshimura, M.; Shigihara, Y.; Okada, T.; Hirata, M. Deep learning based automatic detection and dipole estimation of epileptic discharges in MEG: A multi-center study. *Sci. Rep.* **2024**, *14*, 24574. [CrossRef]
12. Mei, J.; Desrosiers, C.; Frasnelli, J. Machine Learning for the Diagnosis of Parkinson's Disease: A Review of Literature. *Front. Aging Neurosci.* **2021**, *13*, 633752. [CrossRef] [PubMed]
13. Klaver, E.C.; Heijink, I.B.; Silvestri, G.; van Vugt, J.P.P.; Janssen, S.; Nonnekes, J.; van Wezel, R.J.A.; Tjepkema-Cloostermans, M.C. Comparison of state-of-the-art deep learning architectures for detection of freezing of gait in Parkinson's disease. *Front. Neurol.* **2023**, *14*, 1306129. [CrossRef] [PubMed]
14. Bukhari, S.N.H.; Ogudo, K.A. Ensemble Machine Learning Approach for Parkinson's Disease Detection Using Speech Signals. *Mathematics* **2024**, *12*, 1575. [CrossRef]
15. Roth, N.; Kuderle, A.; Ullrich, M.; Gladow, T.; Marxreiter, F.; Klucken, J.; Eskofier, B.M.; Kluge, F. Hidden Markov Model based stride segmentation on unsupervised free-living gait data in Parkinson's disease patients. *J. Neuroeng. Rehabil.* **2021**, *18*, 93. [CrossRef]
16. Rizek, P.; Kumar, N.; Jog, M.S. An update on the diagnosis and treatment of Parkinson disease. *Cmaj* **2016**, *188*, 1157–1165. [CrossRef]
17. Brzenczek, C.; Klopfenstein, Q.; Hahnel, T.; Frohlich, H.; Glaab, E.; Consortium, N.-P. Integrating digital gait data with metabolomics and clinical data to predict outcomes in Parkinson's disease. *NPJ Digit. Med.* **2024**, *7*, 235. [CrossRef]
18. Singh, S.; Xu, W. Robust Detection of Parkinson's Disease Using Harvested Smartphone Voice Data: A Telemedicine Approach. *Telemed. J. E Health* **2020**, *26*, 327–334. [CrossRef]

19. Terlapu, P.V.; Swetha, M.; Ram, J.S.; Srinivas, K.S.; Nataraj, B.S.; Lahari, M.; Sowjanya, G.; Deexitha, B.S.; Mohitha, M.R. Intelligent Parkinson's Disease Detection: Optimization Algorithm Implementation for SVM and MLP Classifiers on Voice Bio-Markers. In Proceedings of the International Conference on Computational Innovations and Emerging Trends (ICCIET 2024), Hyderabad, India, 4–5 April 2024; Advances in Computer Science Research; pp. 230–241.
20. Williamson, J.R.; Telfer, B.; Mullany, R.; Friedl, K.E. Detecting Parkinson's Disease from Wrist-Worn Accelerometry in the U.K. Biobank. *Sensors* **2021**, *21*, 2047. [CrossRef]
21. Lee, C.Y.; Kang, S.J.; Hong, S.-K.; Ma, H.-I.; Lee, U.; Kim, Y.J. A Validation Study of a Smartphone-Based Finger Tapping Application for Quantitative Assessment of Bradykinesia in Parkinson's Disease. *PLoS ONE* **2016**, *11*, e0158852. [CrossRef]
22. Kamble, M.; Shrivastava, P.; Jain, M. Digitized spiral drawing classification for Parkinson's disease diagnosis. *Meas. Sens.* **2021**, *16*, 100047. [CrossRef]
23. Dirkx, M.F.; Zach, H.; van Nuland, A.J.; Bloem, B.R.; Toni, I.; Helmich, R.C.J.B. Cognitive load amplifies Parkinson's tremor through excitatory network influences onto the thalamus. *Brain* **2020**, *143*, 1498–1511. [CrossRef]
24. Vlieger, R.; Daskalaki, E.; Apthorp, D.; Lueck, C.J.; Suominen, H. Evaluating Effects of Resting-State Electroencephalography Data Pre-Processing on a Machine Learning Task for Parkinson's Disease. *Stud. Health Technol. Inform.* **2024**, *310*, 1480–1481. [CrossRef] [PubMed]
25. Belvisi, D.; Conte, A.; Bologna, M.; Bloise, M.C.; Suppa, A.; Formica, A.; Costanzo, M.; Cardone, P.; Fabbrini, G.; Berardelli, A. Re-emergent tremor in Parkinson's disease. *Park. Relat. Disord.* **2017**, *36*, 41–46. [CrossRef] [PubMed]
26. Clark, L.N.; Louis, E.D. Essential tremor. *Handb. Clin. Neurol.* **2018**, *147*, 229–239. [CrossRef] [PubMed]
27. Kostikis, N.; Hristu-Varsakelis, D.; Arnaoutoglou, M.; Kotsavasiloglou, C. A Smartphone-Based Tool for Assessing Parkinsonian Hand Tremor. *IEEE J. Biomed. Health Inform.* **2015**, *19*, 1835–1842. [CrossRef]
28. Sajal, M.S.R.; Ehsan, M.T.; Vaidyanathan, R.; Wang, S.; Aziz, T.; Mamun, K.A.A. Telemonitoring Parkinson's disease using machine learning by combining tremor and voice analysis. *Brain Inform.* **2020**, *7*, 12. [CrossRef]
29. Gerraty, R.T.; Provost, A.; Li, L.; Wagner, E.; Haas, M.; Lancashire, L. Machine learning within the Parkinson's progression markers initiative: Review of the current state of affairs. *Front. Aging Neurosci.* **2023**, *15*, 1076657. [CrossRef] [PubMed]
30. Mukhamediev, R.; Kuchin, Y.; Yunicheva, N.; Kalpeyeva, Z.; Muhamedijeva, E.; Gopejenko, V.; Rystygulov, P. Classification of Logging Data Using Machine Learning Algorithms. *Appl. Sci.* **2024**, *14*, 7779. [CrossRef]
31. Xing, X.; Luo, N.; Li, S.; Zhou, L.; Song, C.; Liu, J. Identification and Classification of Parkinsonian and Essential Tremors for Diagnosis Using Machine Learning Algorithms. *Front. Neurosci.* **2022**, *16*, 701632. [CrossRef]
32. Wang, Z.; Xiong, C.; Zhang, Q. Enhancing the online estimation of finger kinematics from sEMG using LSTM with attention mechanisms. *Biomed. Signal Process. Control* **2024**, *92*, 105971. [CrossRef]
33. Karim, F.; Majumdar, S.; Darabi, H.; Chen, S. LSTM Fully Convolutional Networks for Time Series Classification. *IEEE Access* **2018**, *6*, 1662–1669. [CrossRef]
34. Varghese, J.; Brenner, A.; Fujarski, M.; van Alen, C.M.; Plagwitz, L.; Warnecke, T. Machine Learning in the Parkinson's disease smartwatch (PADS) dataset. *NPJ Park. Dis.* **2024**, *10*, 9. [CrossRef]
35. Brenner, A.; Fujarski, M.; Warnecke, T.; Varghese, J. Reducing a complex two-sided smartwatch examination for Parkinson's Disease to an efficient one-sided examination preserving machine learning accuracy. *arXiv* **2022**, arXiv:2205.05361.
36. Little, M.A.; Volotinen, S.; Sanderson, B.; Huopaniemi, U.; Mowlem, F.; Olt, J.; Byrom, B. Novel algorithms deriving clinical performance measures from smartphone sensor data collected under a walking test. *bioRxiv* **2021**, 465337. [CrossRef]
37. Yuan, Z.; Gao, X.; Yang, K.; Peng, J.; Luo, L. Performance Enhancement of Ultrasonic Weld Defect Detection Network Based on Generative Data. *J. Nondestruct. Eval.* **2024**, *43*, 102. [CrossRef]
38. Zhang, B.; Fang, J.; Li, Y.; Wang, Y.; Zhou, Q.; Wang, X. GFRENet: An Efficient Network for Underwater Image Enhancement with Gated Linear Units and Fast Fourier Convolution. *J. Mar. Sci. Eng.* **2024**, *12*, 1175. [CrossRef]
39. Bishop, C. Training with noise is equivalent to Tikhonov regularization. *Neural Comput.* **1995**, *7*, 108–116. [CrossRef]
40. Le Guennec, A.; Malinowski, S.; Tavenard, R. Data Augmentation for Time Series Classification using Convolutional Neural Networks. In Proceedings of the ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Riva Del Garda, Italy, 19 September 2016. Available online: https://shs.hal.science/halshs-01357973v1 (accessed on 30 August 2016).
41. Caramia, C.; Torricelli, D.; Schmid, M.; Munoz-Gonzalez, A.; Gonzalez-Vargas, J.; Grandas, F.; Pons, J.L. IMU-Based Classification of Parkinson's Disease from Gait: A Sensitivity Analysis on Sensor Location and Feature Selection. *IEEE J. Biomed. Health Inform.* **2018**, *22*, 1765–1774. [CrossRef]
42. Um, T.T.; Pfister, F.M.J.; Pichler, D.; Endo, S.; Lang, M.; Hirche, S.; Fietzek, U.; Kulić, D. Data augmentation of wearable sensor data for parkinson's disease monitoring using convolutional neural networks. In Proceedings of the Proceedings of the 19th ACM International Conference on Multimodal Interaction, Suzhou, China, 14–18 October 2019; Association for Computing Machinery: New York, NY, USA, 2017; pp. 216–220.
43. Kim, M.; Jeong, C.Y. Label-preserving data augmentation for mobile sensor data. *Multidimens. Syst. Signal Process.* **2021**, *32*, 115–129. [CrossRef]
44. Wang, Z.; Yan, W.; Oates, T. Time series classification from scratch with deep neural networks: A strong baseline. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–19 May 2017; pp. 1578–1585. [CrossRef]

45. Baydogan, M.G.; Runger, G. Time series representation and similarity based on local autopatterns. *Data Min. Knowl. Discov.* **2015**, *30*, 476–509. [CrossRef]
46. Tuncel, K.S.; Baydogan, M.G. Autoregressive forests for multivariate time series modeling. *Pattern Recognit.* **2018**, *73*, 202–215. [CrossRef]
47. Hochreiter, S.J.N.C.M.-P. *Neural Comput*; MIT-Press: Cambridge, MA, USA, 1997.
48. Kim, S.S. An Adaptive Time-Delay Recurrent Neural Network for Learning Spatiotemporal Correlations. In *Proceedings of the ICONIP: International Conference on Neural Information Processing*; The Institute of Electronics and Information Engineers: Seoul, Republic of Korea, 1994; pp. 1317–1322.
49. Yu, Z.; Lee, M. Real-time human action classification using a dynamic neural model. *Neural Netw.* **2015**, *69*, 29–43. [CrossRef] [PubMed]
50. Ioffe, S. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.
51. Pengtao, W. Based on Adam Optimization Algorithm: Neural Network Model for Auto Steel Performance prediction. *J. Phys. Conf. Ser.* **2020**, *1653*, 012012. [CrossRef]
52. Avuclu, E.; Elen, A. Evaluation of train and test performance of machine learning algorithms and Parkinson diagnosis with statistical measurements. *Med. Biol. Eng. Comput.* **2020**, *58*, 2775–2788. [CrossRef]
53. Chang, P.; Wang, C.; Chen, Y.; Wang, G.; Lu, A. Identification of runner fatigue stages based on inertial sensors and deep learning. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1302911. [CrossRef]
54. Rifaat, N.; Ghosh, U.K.; Sayeed, A. Accurate gait recognition with inertial sensors using a new FCN-BiLSTM architecture. *Comput. Electr. Eng.* **2022**, *104*, 108428. [CrossRef]
55. di Biase, L.; Pecoraro, P.M.; Pecoraro, G.; Shah, S.A.; Di Lazzaro, V. Machine learning and wearable sensors for automated Parkinson's disease diagnosis aid: A systematic review. *J. Neurol.* **2024**, *271*, 6452–6470. [CrossRef]