# Parkinson's Disease Detection using Spiral Drawings

A. Anisha, B.E., M.E [1]
Femima Shelly. A. T[2]
Benitta. R. K[3]
Amala Selciya. T.L[4]
Dept. of Computer Science and Engineering, St. Xavier's Catholic College of Engineering,
Chunkankadai, Kanyakumari, Tamil Nadu.

**Abstract:- Parkinson's disease is a neurological disorder that primarily affects people over the age of 60, often leading to motor impairment (MI) such as tremors, rigidity, and slowness. The disease's severity has been found to be linked to a decline in handwriting quality, with patients exhibiting reduced speed and pressure while writing. Biomarkers can aid in the diagnosis, monitoring, and prediction of the disease's progression, making it critical to accurately identify them. A convolutional neural network (CNN) is used in this study to analyze spiral drawing patterns from Parkinson's patients and healthy individuals, with the aim of creating a system that can effectively differentiate between the two groups and predict the PD stage. The model was trained on data from 280 patients and achieved an overall accuracy of 94.2%. Identifying biomarkers could provide valuable insights into the disease's causes and lead to better diagnosis and treatment outcomes.**

*Keywords:- Parkinson's Disease, CNN, Deep Learning, Machine Learning, Cat Boost Classifiers, VGG-16 Model.*

## I. INTRODUCTION

Parkinson's disease is a complex and progressive nervous system disorder that often affects people over the age of 60. Parkinson's disease is a complex and progressive neurological disorder that can have a significant impact on a person's daily life. It is commonly seen in people over the age of 60, but it can also affect younger individuals. The disease is characterized by motor impairments, such as rigidity, tremors, and slow movement that can make it difficult for individuals to perform even the most basic activities of daily living. In addition to these motor symptoms, Parkinson's disease patients often exhibit a decline in handwriting quality, with letters becoming smaller and more cramped, and writing becoming slower and more uneven. These changes in handwriting can be an early sign of Parkinson's disease and can help in the early detection and diagnosis of the disease. Biomarkers are measurable indicators of a person's health status that can help diagnose, monitor, and predict the progression of Parkinson's disease. Identifying the correct biomarkers can provide valuable insights into the disease's underlying causes and help develop more effective treatments. Biomarkers can be obtained from various sources, including blood, cerebrospinal fluid, and imaging techniques, and can be used to track changes in the disease over time. In conclusion, Parkinson's disease is a complex and debilitating disorder that affects millions of people worldwide. Identifying the correct biomarkers can help in the early detection, diagnosis, and monitoring of the disease and may lead to more effective treatments in the future.

The proposed system design uses convolutional neural networks (CNN) to analyze spiral drawing patterns in both Parkinson's disease patients and healthy individuals. The study aims to develop a system that can distinguish between spiral sketches from the two groups and predict the stage of Parkinson's disease in patients. The model achieved an overall accuracy of 94.2% after being trained on data from 280 individuals (70 healthy, 70 mild, 70 moderate, and 70 severe patients). This study is a significant step forward in using technology to aid in the diagnosis and monitoring of Parkinson's disease. By accurately identifying biomarkers and analyzing spiral drawing patterns, clinicians can gain valuable insights into the disease's progression and improve patient outcomes.

The potential benefits of using technology to aid in the diagnosis and monitoring of Parkinson's disease are immense. Improved diagnostic accuracy can lead to earlier detection and intervention, potentially slowing the progression of the disease and improving patient outcomes. Additionally, personalized treatment options that target specific disease pathways can be developed based on accurate biomarker identification. Overall, this study provides a promising approach to using technology to aid in the diagnosis and monitoring of Parkinson's disease, potentially leading to better patient outcomes and improved quality of life.

## II. RELATED WORKS

We have reviewed several studies related to the implementation of machine learning techniques for the development of automated systems in Parkinson's disease (PD) detection. In one study, the authors focused on distinguishing PD subjects at different stages based on speed and pen pressure while performing sketches. They extracted features from the sketches and proposed a method to establish a correlation between these features and the severity level of Parkinson's disease (PD). The study validated their approach using statistical tests which revealed significant differences in the correlation factor at different PD stages.

Another study examined how the movement of the pen tip on a paper while drawing the spiral can be used to differentiate between individuals with Parkinson's disease (PD) and those without PD. Features were extracted from the drawings and used to train various machine learning classifiers, including Naïve Bayes, Logistic Regression, Random Forest, K-Neighbors, Decision Tree, Gradient Boosting, XGBoost, LightGBM, and Cat Boost Classifiers. Performance metrics such as accuracy, loss, and Area Under the Curve (AUC) were employed. The study achieved an accuracy of 94.2%.

In a separate study, the researchers collected spiral data using telemetry touch screen devices in home environments to distinguish off episodes and peak dose dyskinesia in PD patients. Features were extracted from the data and used as input for machine learning classifiers, including Logistic Regression, Random Forest, K-Neighbors, and Cat Boost Classifiers. The study found that Cat Boost performed well among the classifiers, achieving an accuracy of 82%.

Additionally, we proposed a study that utilized an image dataset to distinguish PD patients from others. Different feature selection techniques, such as principal component analysis (PCA), were employed to identify the best features for training various classifiers. A performance comparison study was conducted using original feature sets and PCA-based feature sets with nonlinear decision tree-based classifiers. The results indicated that the Random Forest classifier (RFC) outperformed the others, and the PCA-based feature set exhibited better performance compared to the original feature sets. The study achieved the highest accuracy of 96.83% using RFC and PCA-based feature sets.

These studies collectively showcase the effectiveness of machine learning techniques in detecting Parkinson's disease (PD) using diverse data types and feature selection methods. The findings emphasize the promising potential of automated systems in aiding early diagnosis and continuous monitoring of PD patients.

➤ *Literature Gaps*

Considering the profound impact Parkinson's disease (PD) can have on individuals' lives and the potential for long-term hardships, it is crucial to prioritize achieving higher precision and accuracy in PD detection. While the papers discussed various data collection techniques for the control and detection of PD subjects, additional research is necessary to determine the optimal algorithms for precise and accurate PD diagnosis. Presently, research predominantly concentrates on exploring deep learning models, while conventional algorithms like K-nearest neighbors, Naïve Bayes, and Random Forests have been commonly employed for this purpose. Indeed, CatBoost classifiers offer several advantages, including built-in handling of categorical features, advanced regularization techniques, and competitive performance. While the choice of classifier ultimately depends on the specific dataset and problem, CatBoost often stands out as a powerful and efficient option, especially when confronted with complex and diverse data, which is crucial for accurately elucidating PD detection.

## III. METHODOLOGY

The methodology proposed for our project utilizes Convolutional Neural Networks (CNN) to analyze the drawing patterns observed in spiral sketches. CNNs are a type of deep learning model specifically designed for processing grid-like data such as images. In this case, the CNN is employed to extract meaningful features from the spiral sketches, enabling the detection of patterns and characteristics relevant to Parkinson's disease.
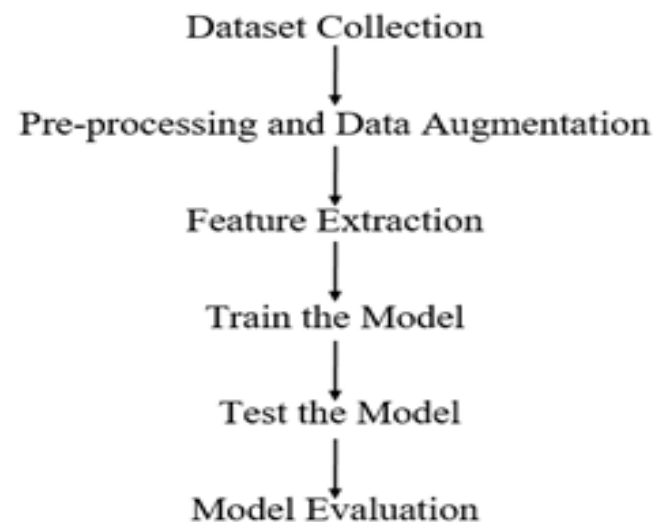


Fig 1 Work Flow Diagram

➤ *Dataset Collection*

In this study, an image dataset was curated by collecting data from 280 individuals at Rhock Hospital in Tirunelveli, India. The dataset was categorized into four distinct groups:
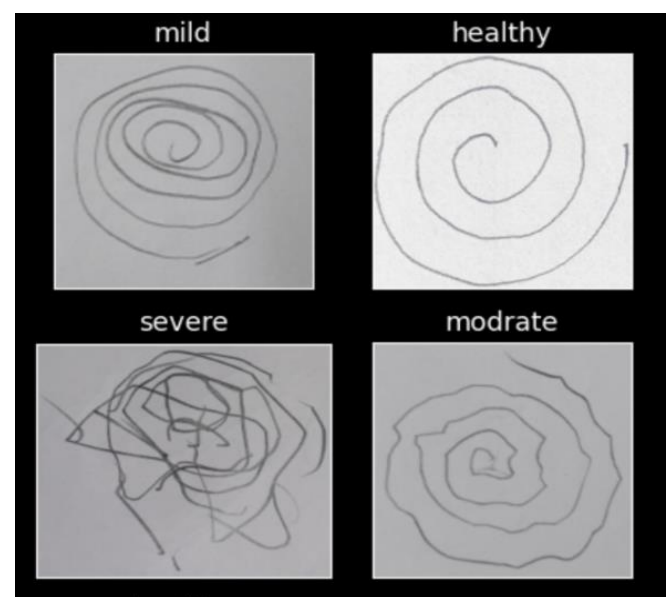


Fig 2 Sample Images for Spiral Sketches

- *The first group comprised 70 samples of healthy individuals.*
- *The second group consisted of 70 cases in the mild stage of Parkinson's disease (PD).*
- *The third group included 70 cases in the moderate stage of PD.*
- *The fourth group encompassed 70 cases in the severe stage of PD.*
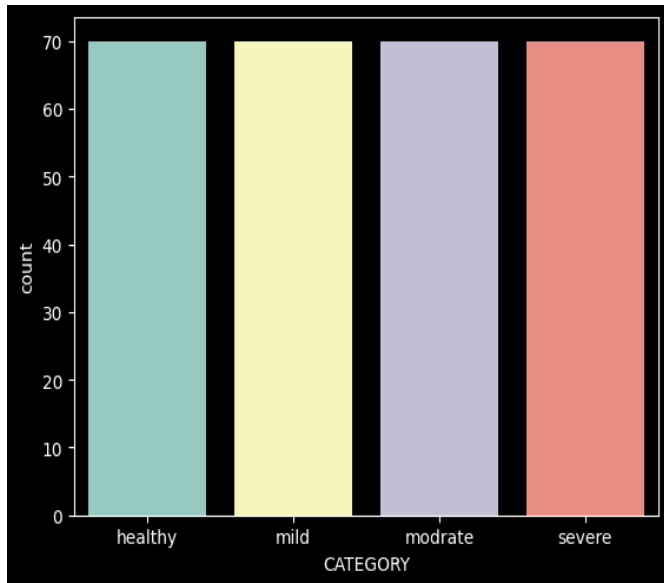


Fig 3 Dataset Group and the Count of Data

Consequently, 25% of the entire dataset consisted of samples from healthy individuals, while the remaining 75% represented individuals with PD. During the data collection process, each patient was requested to draw a spiral on paper. Subsequently, the spiral drawings were collected, scanned, and assembled to form an image dataset.

➤ *Pre-Processing and Data Augmentation*

The images from each group are read, and if they fall under the same category, they are preserved with that label (healthy, mild, moderate, or severe). The images are resized to 256×256 for use in Transfer Learning (TL) models. In order to ensure compatibility with the requirements of certain transfer learning (TL) models, the data segmentation is performed in advance, specifically focusing on adjusting the image sizes. This proactive approach helps to prevent potential issues or conflicts that may arise during the later stages of the PD detection process using TL models.

By appropriately resizing and standardizing the segmented images, any potential compatibility problems with the TL models are addressed beforehand, allowing for a smoother and more efficient workflow. To meet TL model requirements, the images in the study are read as RGB, aligning with the colored image training data. Additionally, the pixel values are scaled down from 0 to 255 to a range of 0 to 1 by dividing them by 256. This rescaling optimizes the performance and compatibility of the images with deep learning models, which tend to work more effectively within the 0-1 value range.

Table 1 Data Augmentation Parameters for Spiral Drawings

| Augmentation Parameters | Settings |
|---|---|
| Horizontal Flip | True |
| Vertical Flip | True |
| Brightness Range | ( 0.2, 0.9 ) |
| Rotation Range | 50 |
| Shear Range | 0.7 |
| Zoom Range | 0.7 |
| Rescale | 1./255 |

➤ *Feature Extraction*

Feature extraction plays a crucial role in Parkinson's disease (PD) detection using Convolutional Neural Networks (CNNs) by transforming raw input data, such as images or signals, into meaningful representations that capture relevant information specific to PD. In PD detection using CNN, feature extraction is essential for selecting discriminative features that differentiate between healthy individuals and PD patients according to the progression of the disease. This process entails converting raw text data into numerical representation sets that effectively capture significant characteristics of the image. These features are then used to train the machine learning model, enabling it to learn patterns and accurately classify individuals as healthy or having PD. In this process, the data will be split into a training dataset and a testing dataset, and feature extraction is carried out using the VGG16 transfer learning model. The splitting is done using the CatBoost Classifier algorithm, which is a powerful gradient-boosting algorithm specifically designed to handle categorical features effectively. By utilizing this algorithm, the data is partitioned into training and testing sets, allowing for the model to be trained on the training dataset and evaluated on the testing dataset.

➤ *Train the Model*

Training the CNN model involves using the prepared image dataset and training the machine learning algorithms on the pre-processed and feature-extracted data. In this process, a feature matrix is created from the pre-processed data, which serves as the input for training the model. The feature matrix represents the dataset in a structured format, allowing the model to learn patterns and make predictions based on the extracted features. The feature matrix in this scenario consists of two columns: the "category" column and the "png" column. The "category" column represents the category of the data, with values ranging from 0 to 3 representing healthy, mild, moderate, and severe respectively. The "png" column contains the image data, which has been transformed into a float data type. The feature matrix encapsulates both the category labels and the corresponding image data in a structured format for further analysis and model training. The CatBoost classifier, along with other classifiers and regressors, can be used to train the model. The training process involves fitting the model to the training data using the `fit()` method, allowing the model to learn patterns and relationships within the data. Once the model is trained, it can be used to make predictions between healthy individuals and the progression of the PD based on the learned patterns.

➢ *Test the Model*

Testing the CNN model involves assessing the model's performance on a separate testing dataset, which was not used during training. This provides an unbiased measure of the model's accuracy and ability to detect PD. During testing, the CNN model predicts the stage of Parkinson's disease using the test data. Comparing these predictions with the true labels, evaluation metrics like Accuracy and Precision are calculated. Accuracy measures overall correctness, while Precision assesses correctly predicted positive cases among all predicted positives. These metrics evaluate the model's performance in identifying between healthy individuals and Parkinson's disease stages.

➢ *Model Evaluation*

Performance metrics, including accuracy, precision, recall, F1-score, and AUC-ROC, are calculated to quantify the effectiveness of the model in Parkinson's disease (PD) detection. These metrics provide valuable insights into the model's performance and can be used to refine and optimize the system based on the evaluation results. Evaluation is essential to assess the model's ability to accurately identify the PD stage, and the metrics are derived from the model's performance on the testing set, providing a reliable measure of its capabilities.

## IV. RESULTS AND DISCUSSION

The developed multistage classifier, combining convolutional neural networks and machine learning algorithms, demonstrated promising results in the detection of Parkinson's disease from Spiral Sketches. The model achieved an accuracy of 94.2%, an average recall of 94%, an average precision of 93.5%, and an average F1 score of 94.91%. The performance evaluation metrics indicate the model's effectiveness in accurately identifying Parkinson's disease cases.

Additionally, K-Fold cross-validation was conducted on the complete dataset to assess the model's generalizability. The results showed that the model performed consistently well across different folds of training, validation, and testing, with a consistent tendency towards precision and recall. The performance aligns with the initial hypothesis, suggesting that the model can successfully differentiate between healthy individuals and those with Parkinson's disease. However, it is important to note that individuals predicted as healthy but having Parkinson's disease may be in the early stages of the condition, as the healthy subjects were selected to match the same age group.

➢ *Classification of PD Using VGG-16*

In Figure 4, the performance of the VGG16 model is visually represented in terms of accuracy and loss. The purple line represents the training accuracy, which is measured at 90.63%, while the grey line represents the testing accuracy, measured at 91.63%. It is observed that there is minimal difference between the testing and training accuracy, indicating that the model generalizes well to unseen data. The graph also illustrates the training loss (red

line) and testing loss (blue line), both of which decrease at a similar rate with little disparity. This suggests that the VGG16 model is well-suited for training and testing the handwritten image dataset, demonstrating a good fit.
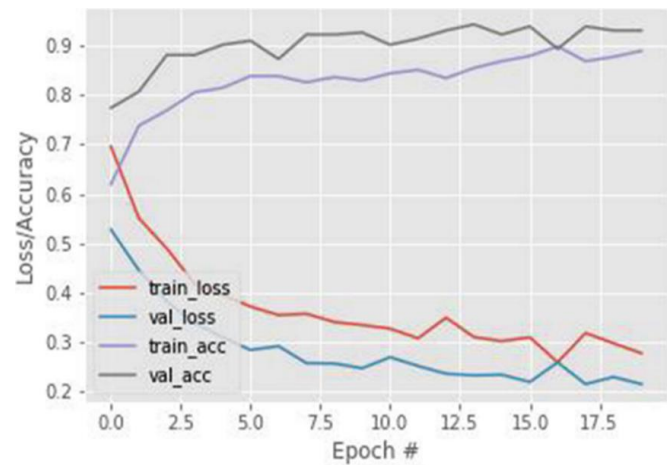


Fig 4 Training Loss and Accuracy on PD

➢ *K-fold Cross Validation*

Cross-validation is a widely used technique for evaluating machine learning models. In k-fold cross-validation, the dataset is divided into k subsets of equal size. The model is trained and tested k times, with each subset serving as the testing set once and the remaining subsets as the training set. This process helps assess the model's performance on different subsets of the data and reduces the impact of subject dependence.

By utilizing under-sampling with replacement, we ensure that each training set contains a diverse representation of the data. This helps prevent over-fitting and improves the generalization ability of the model. K-fold cross-validation provides a more robust evaluation of the model's effectiveness compared to a single train-test split. Additionally, it is cost-effective since it makes efficient use of the available data.

Table 2 Result of Different Parameters

| Fold | Training Accuracy in % | Testing Accuracy in % |
|---|---|---|
| 1 | 88.69 | 88.16 |
| 2 | 88.37 | 88.37 |
| 3 | 90.40 | 91.35 |
| 4 | 89.79 | 90.12 |
| Average | 89.31 | 89.50 |

In our study, we applied k-fold cross-validation with k = 4 to evaluate our deep learning model on our combined dataset. This approach allowed us to assess the model's performance across different subsets of the data and obtain more reliable performance estimates. Table 2 presents the results of our model using four-fold cross-validation. The average accuracy and the best accuracy were calculated based on the performance of the model across the four folds.

The average accuracy achieved was 89.50%, indicating the overall correctness of the model's predictions across the different subsets of the data. These results demonstrate the effectiveness of our model in accurately classifying the instances in the dataset.

➤ *Performance Metrics*

Commonly used metrics like accuracy, precision, sensitivity, and specificity may not be suitable for evaluating classifiers in imbalanced classification problems. They heavily rely on the majority class and can be misleading in detecting minority-class samples. Sensitivity and precision overlook the significance of true negatives, which is crucial in medical diagnoses. To address these limitations, alternative metrics such as F1-score and geometric mean (Gmean) are used to balance sensitivity and precision. However, these metrics do not consider true negatives and individual class contributions. Therefore, advanced metrics like the index of balanced accuracy (IBA) and Area Under the Curve (AUC) are included. IBA provides a balanced assessment by considering each class's contribution, while AUC measures the classifier's performance across different thresholds. Incorporating these advanced metrics alongside traditional ones enables a more accurate and comprehensive evaluation of classifiers in imbalanced data scenarios.

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

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

$$Sensitivity = TPR = \frac{TP}{TP + FN}$$

$$Specificity = TNR = \frac{TN}{TN + FP}$$

$$F1 = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity}$$

$$Gmean = \sqrt{Sensitivity \times Specificity}$$

Fig 5 Formulation for the Evaluation Metrics

In figure 5, where TP, FP, TN, FN, TPR, and TNR, refer respectively to, true positive, false positive, true negative, false negative, true positive rate, and true negative rate.

## V. CONCLUSION

This study aimed to improve the early detection of Parkinson's disease, which is crucial for understanding the causes, initiating interventions, and developing effective treatments. To achieve this, the study proposed a deep learning model that could distinguish between healthy individuals and those with PD based on their handwriting and analyze the stage of the disease. The proposed deep learning model achieved a high level of accuracy, with a detection rate of 94.2%.

The study demonstrates that the proposed method, which uses CNNs and spiral drawings, has the potential to serve as a reliable tool for diagnosing Parkinson's disease and predicting the stage of the disease, as it achieved high accuracy. Additionally, the study showed that the proposed method had higher accuracy compared to other machine learning techniques, such as Support Vector Machines (SVM) and Random Forest (RF). However, it is crucial to note that the study's reported accuracy is based on a relatively small dataset collected from 280 individuals (70 healthy, 70 mild, 70 moderate, and 70 severe patients) and may not be representative of more extensive or diverse populations. Further research employing larger and more diverse datasets is necessary to validate the effectiveness and generalizability of the proposed method. The potential of deep learning to handle increasingly complex and larger datasets is expected to be demonstrated in the future. Therefore, the outcomes of this study can be viewed as a promising initial step towards utilizing cutting-edge research for the early detection of diseases and analyzing the stage of the disease progression.

➤ *Acknowledgment*

## REFERENCES

[1]. Sura Mahmood Abdullah, Thekra Abbas, Munzir Hubiba Bashir "Deep Transfer Learning Based Parkinson's Disease Detection Using Optimized Feature Selection". IEEE Access Volume 11 – 2023

[2]. Wu Wang, Junho Lee, Fouzi Harrou, Ying Sun "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning". IEEE Access Volume 8 - 2020

[3]. Protima Khan, Md.Fazlul Kader, S.M.Riazul Islam, Aisha B.Rahman "Machine Learning and Deep Learning Approaches for Brain Disease Diagnosis: Principles and Recent Advances". IEEE Access Volume 9 - 2021

[4]. Fatih Demir, Abdulkadir Sengur, Ali Ari, Kamran Siddique "Feature Mapping and Deep Long Short Term Memory Network-Based Efficient Approach for Parkinson's Disease Diagnosis". IEEE Access Volume 9 - 2021

[5]. Lin Meng, Jun Pang, Yifan Yang, Lei Chen, Rui, Xu "Inertial-Based Gait Metrics During Turning Improve the Detection of Early-Stage Parkinson's Disease Patients". IEEE Transactions on Neural Systems and Rehabilitation Engineering Volume 31 - 2023

[6]. Mohammod Abdul Motin, Nemuel Daniel Pah, Sanjay Raghav, Dinesh Kant Kumar "Parkinson's Disease Detection Using Smartphone Recorded Phonemes in Real World Conditions". IEEE Access Volume 10 - 2022

[7]. Shaohua Wan, Yan Liang, Yin Zhang, Mohsen Guizani "Deep Multi-Layer Perceptron Classifier for Behavior Analysis to Estimate Parkinson's Disease Severity Using Smartphones". IEEE Access Volume 11 - 2018

[8]. Wu Wang, Junho Lee, Fouzi Harrou, Ying Sun "Early Detection of Parkinson's Disease Using Deep Learning and Machine Learning". IEEE Access Volume 8 - 2020

[9]. Manuel Gil-Martín, Juan Manuel Montero, Rubén San-Segundo "Parkinson's Disease Detection from Drawing Movements Using Convolutional Neural Networks". https://www2.mdpi.com/2079-9292/8/8/907 - 2019

[10]. B.R.Bloem, W.J.Marks Jr, A.L.Silva de Lima, M.L.Kuijf, T.Van Laar "The Personalized Parkinson Project: Examining disease progression through broad biomarkers in early Parkinson's disease". https://pubmed.ncbi.nlm.nih.gov /31315608/ - 2019

[11]. Mohamed Shaban "Deep Learning for Parkinson's Disease Diagnosis: A Short Survey". https://www2.mdpi.com/2073-431X/12/3/58 - 2023

[12]. Alex Li, Chenyu Li "Detecting Parkinson's Disease through Gait Measures Using Machine Learning". https://www2.mdpi.com/2075-4418/12/10/2404 - 2022

[13]. Himanish Skekhar Das, Akalpita Das, Anupal Neog, Saurav Mallik, Kangkana Bora, Zhongming Zhoa "Early Detection of Parkinson's Disease Using Fusion of Discrete Wavelet Transformation and Histograms of Oriented Gradients". https://www2.mdpi.com/2227-7390/10/22/4218 - 2022

[14]. Claudio Gallicchio, Alessio Micheli, Luca Pedrelli "Deep Echo State Networks for Diagnosis of Parkinson's Disease". https://arxiv.org/abs/1802.06708 - 2018

[15]. L. Cunningham, S. Mason, C. Nugent, G. Moose, D. Finlay, D. Craig "Home-Based Monitoring and Assessment of Parkinson's Disease". IEEE Transactions on Information Technology in Biomedicine Volume 15 - 2010

[16]. Sharif Noor Zisad, Mohammad Shahadat Hossain, Karl Andersson "Speech Emotion Recognition in Neurological Disorders Using Convolutional Neural Network". https://link.springer.com/chapter/10.1007/978-3-030-59277-6_26 - 2020

[17]. Ghayth AlMahadin, Ahmad Lotfi, Philip Breedon "Enhanced Parkinson's Disease Tremor Severity Classification by Combining Signal Processing with Resampling Techniques". SpringerLink. SN Computer Science - 2021

[18]. Srikanth Tammina "Transfer learning using VGG-16 with Deep Convolutional Neural Network for Classifying Images". International Journal of Scientific and Research Publications, Volume 9, Issue 10 – 2019

[19]. Tzu-Tsung Wong, Nai-Yu Yang "Dependency Analysis of Accuracy Estimates in k-Fold Cross Validation". IEEE Transactions on Knowledge and Data Engineering – 2017

[20]. Tzu-Tsung Wong, Po-Yang Yeh "Reliable Accuracy Estimates from k-Fold Cross Validation". IEEE Transactions on Knowledge and Data Engineering - 2019