An Overview of Various Datasets Used in the Early Detection of Parkinson's Disease

Ishita Jaiswal, Bramah Hazela(✉), Shikha Singh, and Vineet Singh

Amity University, Uttar Pradesh, Lucknow Campus 226010, India
vsingh@lko.amity.edu

Abstract. Parkinson′s disease (PD) is a progressive neurological disorder characterized by motor and non-motor symptoms resulting from the degeneration of dopamine-producing neurons in the brain′s substantia nigra. Common motor symptoms include tremors, rigidity, postural instability, and bradykinesia. While PD has no cure, medications can help manage symptoms. Early detection is crucial for timely intervention. This paper surveys datasets and machine learning (ML) approaches for PD detection, comparing their effectiveness. ML models like Artificial Neural Networks (ANN), Support Vector Machines (SVM), k-Nearest Neighbors (KNN), and XGBoost have been employed. ANN and XGBoost achieved accuracies of 96.7% and 95%, respectively, using extensive voice datasets, while SVM attained 99% accuracy with smaller datasets.

Parkinson′s disease, a progressive neurological disorder, manifests motor symptoms due to dopamine neuron degeneration. Medications can mitigate symptoms, emphasizing the importance of early detection. This study evaluates ML techniques for PD detection, highlighting ANN, SVM, KNN, and XGBoost. ANN and XGBoost excel with large voice datasets, achieving 96.7% and 95% accuracy, respectively. SVM outperforms with smaller datasets, achieving 99% accuracy. Early prediction facilitated by ML classification models is crucial for timely intervention in Parkinson′s disease management.

Keywords: Artificial Nural Network · K-Nearest Neighbor · XG Boost · SVM · Neural Networks

1 Introduction

PD is a progressive neurological disorder that deteriorates over time and affects speech as well as other movements. Neurodegenerative diseases are a group of disorders that include both inherited and randomly occurring problems that cause the neurological system to gradually deteriorate [1]. PD remains the second most prevalent neurodegenerative disorder, following Alzheimer′s disease, brain cancer, degenerative nerve diseases, and epilepsy [2].

PD primarily stems from the progressive loss of neurons that produce dopamine in the "substantia nigra" region of the midbrain that is recognized as the "movement control center" as shown in Fig. 1. This dopamine loss leads to uncontrolled movements

P. Pareek et al. (Eds.): IC4S 2024, LNICST 598, pp. 191–205, 2025.
https://doi.org/10.1007/978-3-031-77078-4_17

known as hypokinetic movement disorders [3]. While PD is relatively easy to diagnose in advanced stages, effective treatment remains a considerable challenge, with no known cure or definitive medical intervention available.

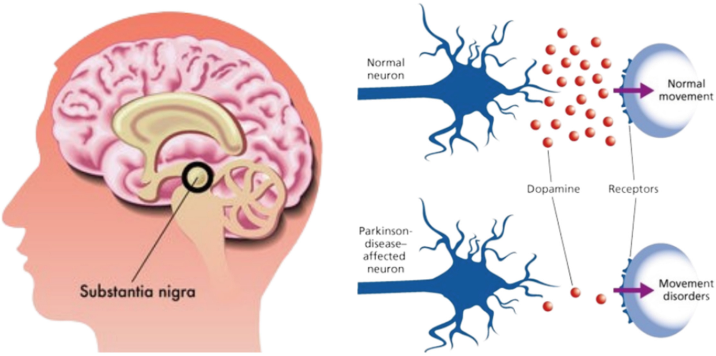

Fig. 1. PD: Typical versus abnormal movement

Despite many years of research, the precise cause of PD remains elusive. Numerous researchers believe that a blend of genetic factors [4] and environmental influences [5] contribute to PD′s development. These environmental factors may include exposure to toxins, head injuries, rural living conditions, water quality, manganese exposure, and pesticide exposure, although their impact can vary from person to person. Additionally, PD manifests differently in each affected individual. Table 1 outlines the different stages of PD. The main motor symptoms include bradykinesia (slowness of movement), limb rigidity, trunk rigidity, tremors in the hands, arms, legs, jaw, and face, and poor balance and coordination. [6]. Non-motor symptoms, such as depression and memory loss can also occur and have a notable impact on the quality of life of an individual [7][8]. Early and precise diagnosis of PD is crucial due to the challenges associated with its advanced-stage treatment, which may be less effective in halting disease progression. This emphasizes the importance of early detection to improve patient outcomes. Table 1 shows the progression of PD.

Table 1 explains that severity of the disease increases many folds as the stage of the disease progresses. Early detection of Parkinson Disease will be helpful for the timely treatment of disease. Presently, Machine learning based predictive models have been accurately predicting the medical conditions of the patients. A lot of research has been done on the dataset of PD to predict the condition of an individual showing the symptoms of the disease in the early stages.

This paper reviews the Machine learning based models using the Voice dataset, Time series Gait dataset, Parkinson disease mobile data and such other databases to detect Parkinson at early stage. Early detection of PD will be helpful to decide the course of treatment to prevent the further deterioration of patient's condition. Additionally, physical, and occupational therapies, along with lifestyle modifications, can improve the standard of living for those with PD. Early detection of PD offers a critical window for implementing interventions aimed at slowing its progression. Through a

Table 1. Progression Phases of PD

Stages	Symptoms
Initial Phase (Phase 1)	At this point, individuals with PD experience minimal disruption in daily activities. Symptoms, including tremors, are predominantly observed on just one side of the body
Intermediate Phase (Phase 2)	Here, manifestations such as rigidity and tremors are evident on each side of the body. Additionally, patients' facial expressions might alter
Middle Phase (Phase 3)	This stage showcases significant symptoms like compromised balance and reduced reflexes, along with the symptoms from intermediate stage
Progressive Phase (Phase 4)	Symptoms intensify during this phase, making mobility challenging without external aids like walking assistance
Severe Phase (Phase 5)	This phase represents the most severe and incapacitating stage. Leg rigidity may lead to episodes of freezing while standing, increasing the risk of falls. Additionally, patients may encounter hallucinatory experiences and sporadic delusions

combination of medication management, physical therapy, speech therapy, occupational therapy, dietary adjustments, lifestyle modifications, and regular monitoring, individuals can effectively manage symptoms and maintain functionality. By initiating treatment promptly upon diagnosis, individuals can potentially delay the onset of advanced stages of the disease, thereby improving overall quality of life and maximizing functional independence. Collaboration between patients, caregivers, and healthcare professionals is essential in developing personalized strategies that address the multifaceted aspects of PD management from its early stages.

2 Datasets Used in Various Studies

Researchers have worked on different dataset for detection of PD. Few of the datasets discussed in this research paper are UCI Voice Dataset [9][10], Time Series Gait Dataset [11], The mpower study [12] Parkinson disease mobile data, HandPDMultiMC dataset [13], PaHaW dataset [14], "Gait in Parkinson′s Disease" database [15], Dataset of DaTscan SPECT images [16].

2.1 Voice Data Importance in Parkinson's Disease Detection

Multiple research studies have consistently identified that more than 90% of PD patients encounter speech and vocal difficulties, encompassing issues such as dysphonia, dysarthria, monotone, and hypophonia [17][18]. Consequently, one of the initial noticeable indications in individuals with Parkinson is the deterioration of their voice.

The assessment of voice presents a straightforward and non-invasive means of analysis so voice analysis can be used to track the advancement of PD [19][20]. To gauge the advancement of PD, numerous vocal tests have been developed, including sustained phonation and continuous speech assessments [21]. Typically, individuals with Parkinson′s Disease (PD) face challenges in their speech, which can be divided into two main categories: hypophonia and dysarthria. Hypophonia refers to a person speaking in a very soft and faint voice, while dysarthria signifies slow and unclear speech due to central nervous system damage. Consequently, many healthcare professionals who cater to PD patients primarily notice dysarthria and aim to enhance their vocal strength

Table 2. Description of the data sets

Dataset	**Type of Data**	**Detailed Description**
UCI Voice Dataset	Voice data	This dataset comprises biomedical voice measurements collected from 31 individuals, with 8 individuals forming the control group and 23 individuals diagnosed with PD. It includes a total of 195 biomedical voice measures
Time Series Gait Dataset	Gait data	It comprises records from patients diagnosed with PD (15 individuals), Huntington′s disease (20 individuals), and amyotrophic lateral sclerosis (13 individuals). Records from 16 healthy individuals serving as controls are included
The mPower study, Parkinson disease mobile data	Mobile data	The research gathered information from a sample of 14,684 individuals, comprising both patients with PD and healthy participants
HandPDMultiMC dataset	Handwriting data	The dataset contains handwriting samples collected from 42 subjects, with 21 diagnosed with PD and 21 serving as controls. Additionally, it includes time series signals
PaHaW dataset	Handwriting data	The Parkinson's Disease Handwriting Database (PaHaW) comprises numerous handwriting samples obtained from 37 individuals diagnosed with PD alongside 38 controls who are matched in terms of gender and age
Dataset of DaTscan SPECT images	SPECT images	The dataset gathered for a study encompasses 659 distinct patient DaT SPECT images, classified into one of two categories: PD with a count of 449 images and non-PD with 210 images

through specific therapeutic interventions [22][23]. Large datasets are available to train the Machine learning algorithm for classification. Table 2 shows a description of the datasets. Table 3 shows the voice measures of the UCI voice dataset.

Figure 2 shows the comparison between the various datasets that are used for the prediction the Parkinson Disease. It shows that reliability of the SPECT images for the accurate prediction of the disease.

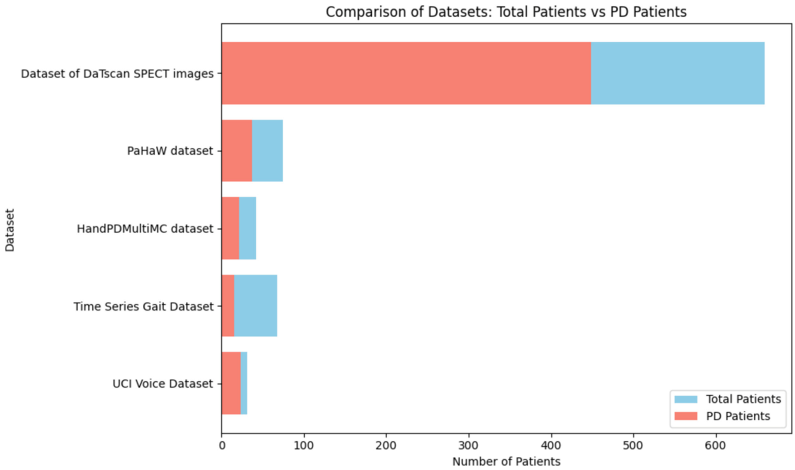

Fig. 2. Comparison of Datasets based on the number of PD records.

Table 3 shows the various sources of Parkinson Dataset that has been used for the predictive analysis.

3 Various ML Approaches

A lot of research is focused on the early prediction of PD using the Machine learning based models. These models have used a large dataset through reliable resources. Different methods have been applied using the datasets such as brain MRI scans, voice recordings, posture images, sensor data, and handwritten samples. This section discusses the most prominent Machine Learning based predictive models.

Prediction of Parkinson's Disease Using Biomedical Voice Measurements Dataset and XG Boost algorithm

The research by Oduntan [24] utilized biomedical voice measurements to discern patients with PD from healthy individuals, employing the Knowledge Discovery in Databases (KDD) process through XG Boost. XG boost make use of decision trees to minimize a cost function through gradient optimization. It is capable of sparsity -aware split finding to improve efficiency. It can split trees with greater precision with techniques to minimize overfitting and provide accurate results. This model is more suitable for the larger datasets. By leveraging the Extreme Gradient Boosting (XGBoost) algorithm, the

Table 3. Voice Measures of UCI voice Dataset

Feature No	Voice_Measure	Meaning
1	MDVP: Fhi (Hz)	Maximum vocal fundamental frequency
2	MDVP: Fo (Hz)	Average vocal fundamental frequency
3	MDVP: Jitter (%)	Several measures of variation in data
4	MDVP:Flo (Hz)	Minimum vocal fundamental frequency
5	MDVP:Jitter (Abs)	fundamental frequency
6	MDVP: PPQ	MDVP five-point period perturbation quotient
7	MDVP: RAP	MDVP relative amplitude perturbation
8	Jitter:DDP	Average absolute difference of differences between jitter cycles
9	MDVP:Shimmer	Several measures of variation in amplitude
10	HNR	components in the voice
11	NHR	Two measures of ratio of noise to tonal
12	D2	measures
13	RPDE	Two nonlinear dynamical complexity
14	spread1	Three nonlinear measures of fundamental
15	DFA	Signal fractal scaling exponent
16	status	Health status of the subject: (1) Parkinson's, (0) healthy
17	spread2	frequency variation

study achieved an impressive 95% accuracy in classification. As we can see in Fig. 3, the model exhibited a 100% precision rate for identifying healthy subjects and a 94% recall rate for detecting PD patients. This underscores the potential of vocal features as a valuable asset in the early detection of PD, offering a promising avenue for improving diagnostic efficiency, especially in underserved regions.

This research utilized a larger dataset and achieved accurate results that establishes the role of voice dataset in the early detection of PD.

SVM Based Machine Learning Approach to Identify Parkinson's Disease Using Gait Analysis

Parkinson disease affects the walking ability of the patients, and the change in the walking pattern can be a feature for the detection of PD. Collection of gait data is complicated so a huge data set is difficult to achieve. In the research by Hausdorff JM, Mitchell SL [25], gait-related features from PD patients, healthy controls, "Amyotrophic lateral sclerosis" (ALS) individuals, and Huntington′s disease (HD) patients were analysed using a SVM classifier with a Gaussian radial basis function kernel. SVM is a classifier that works well with smaller dataset. In SVM dataset is classified through a hyperplane that divides the into suitable classes. The dataset, comprising 12 gait metrics, underwent

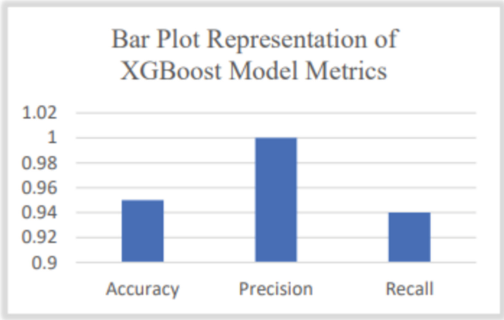

Fig. 3. Bar Plot Representation of the result

feature selection, leading to the identification of seven pivotal features. The optimized SVM model achieved an accuracy of 83.33%, effectively distinguishing PD patients with a 75% true positive rate. This study underscores the potential of gait dynamics in differentiating PD from other neurodegenerative conditions, emphasizing the importance of gait analysis for non-invasive neurological diagnostics. Figure 4 shows the result of classification with SVM based model based on various performance parameters.

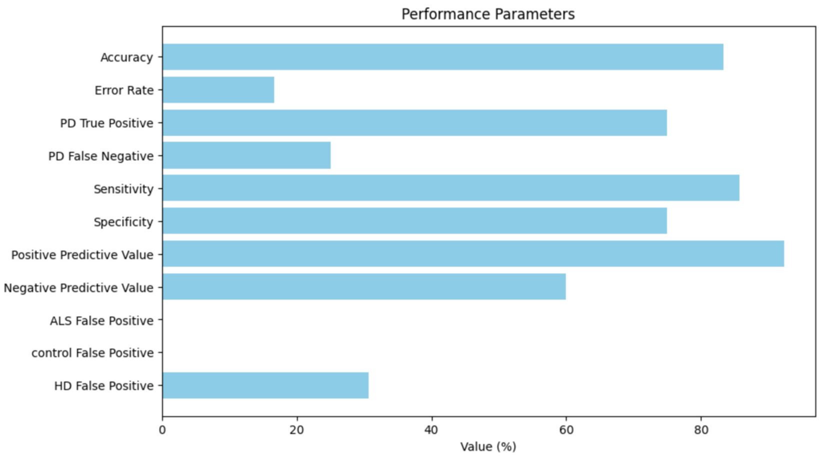

Fig. 4. Results of classification with SVM

Parkinson's Disease Identification using KNN and ANN Algorithms based on Voice Disorder

In this work, authors explored the diagnostic potential of voice analysis in detecting PD [26]. For this purpose, the study employed a dataset that focused on voice-based features commonly used in acoustic analysis, such as fundamental frequency (F0), jitter, shimmer, and harmonics-to-noise ratio (HNR). Two classification methodologies were adopted: ANN and KNN. Artificial Neural Network consist of three layers namely input layer, output layer and hidden layer. These layers are interconnected with each other in such a way where input layer receives input, hidden layer transforms the data to be passed onto output layer. Every connection has some weight that determines the influence of one layer on other. KNN searches the nearest neighbours surrounding the data. As the value of k increases smoothness in the classification increases. The ANN method demonstrated superior performance with a 96.7% accuracy rate, outperforming the KNN method, which achieved 79.31%. The findings suggest that voice-based ANN systems hold promise for enhancing telemedicine-based PD diagnosis.

Parkinson's Disease Diagnosis Using Machine Learning and Voice
In this work, T. J. Wroge utilized voice data from the mPower study to diagnose PD [27] using the mRMR and GeMaps algorithms for feature extraction and various ML models for classification. mRMR stands for "Minimum redundancy Maximum Relevance" that extracts the most relevant features to reduce the redundancy. It is most important pre-processing tool. GeMAPS stands for Geneva Minimalistic Acoustic Acoustic Parameter Set that works with the low-level features such as Harmonic Difference F0, Harmonic Difference (F0- A3), Alpha ratio, Hammarberg Index. This study achieved an impressive 86% accuracy in distinguishing PD cases, with the Gradient Boosted Decision Tree and ANN performing notably well. These findings highlight the promising potential of voice-based biomarkers in PD diagnosis and signify a significant step towards more personalized detection. Figure 5 shows the precision-recall curve obtained. Figure 6 and 7 show the result of classifiers for AVEC dataset and GeMaps dataset respectively.

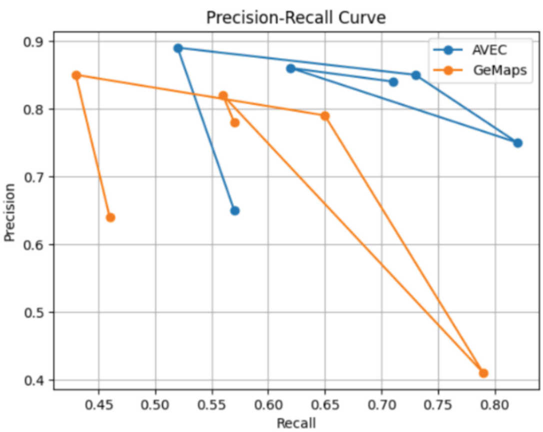

Fig. 5. Precision-Recall Curve

An Efficient Machine Learning Approach for Diagnosing Parkinson's Disease by Utilizing Voice Features
This research paper investigated the efficacy of ML techniques in detecting PD through voice analysis [28]. Utilizing a dataset encompassing voice recordings from 195 participants, both diagnosed with PD and healthy, the study examined a spectrum of acoustic

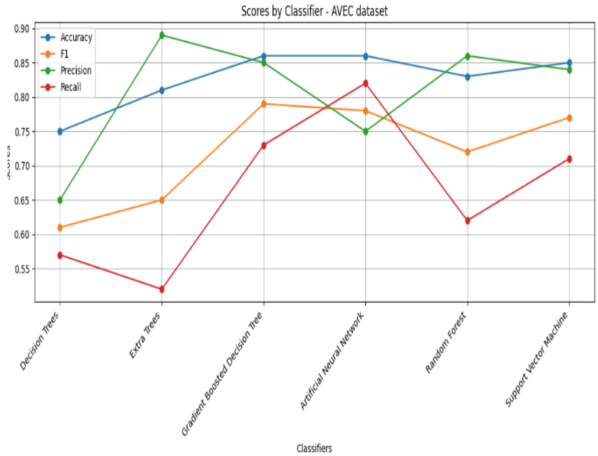

Fig. 6. Scores by classifier for AVEC dataset

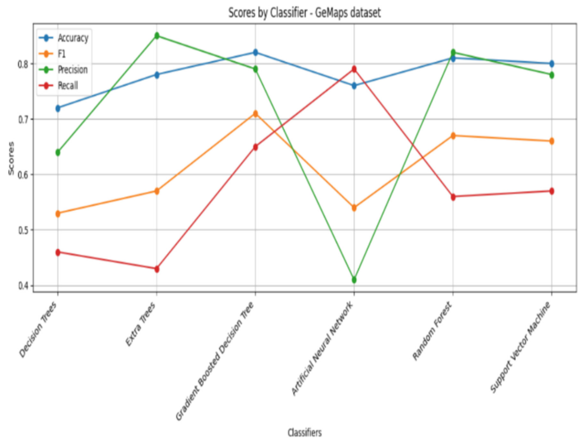

Fig. 7. Scores by classifier for GeMaps dataset

speech metrics to pinpoint relevant indicators of PD. Methodologically, the research prioritized feature selection to enhance diagnostic precision and deployed four machine learning classifiers: SVM, Naive Bayes, KNN and ANN. Evaluation metrics such as accuracy, F1-score (Fig. 8), sensitivity, specificity, and Matthews's Correlation Coefficient (MCC) were employed. Results indicated varying levels of accuracy across classifiers: SVM achieved 87.17%, Naïve Bayes recorded 74.11%, K-NN matched SVM at 87.17%, but ANN emerged as the frontrunner with an accuracy of 96.7%. Thus, ANN was identified as the most proficient classifier for PD detection based on voice analysis, underscoring its pivotal role in advancing non-invasive diagnostic techniques for neurodegenerative disorders. Figure 9 shows a bar chart representing performance comparison across classifiers.

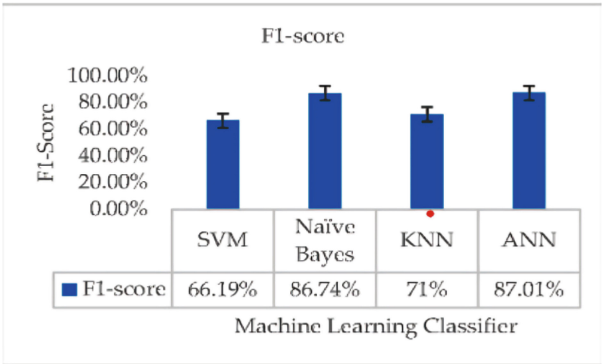

Fig. 8. F1 Score of various machine Learning Classifiers [29]

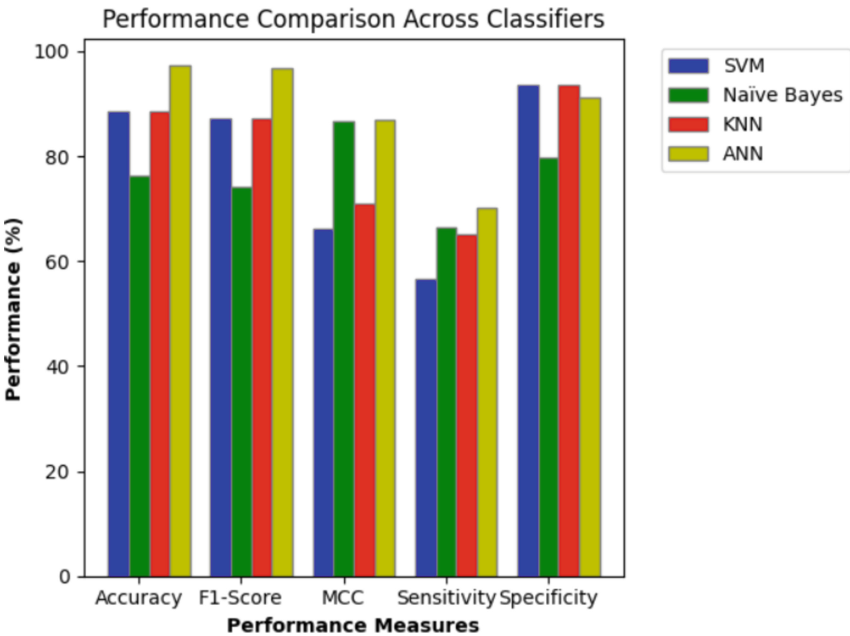

Fig. 9. Performance comparison across classifiers

Detection of Parkinson's disease from handwriting using deep learning: a comparative study

The research paper by Taleb [29] aimed to leverage deep learning, specifically CNN and CNN combined with Bidirectional Long Short-Term Memory (CNN-BLSTM), for PD detection using handwriting data from the HandPDMultiMC dataset. The study innovatively represented time series handwriting data through both Concatenation and Spectrogram approaches, the latter utilizing Short Time Fourier Transforms (STFT). Two model architectures were tested: a standalone CNN and a hybrid CNN-BLSTM. To prevent overfitting, the study employed transfer learning from the PaHaW dataset and data augmentation techniques like synthetic data generation and jittering.

In terms of results, the CNN-BiLSTM model, when enhanced with synthetic data and jittering, exhibited a standout accuracy of 97.62%. This accuracy notably surpassed previous SVM-based methodologies. The study conclusively demonstrated the potential of deep learning in enhancing the accuracy of PD detection through handwriting dynamics.

Parkinson's Disease Detection from Gait Pattern

The paper [30] by A. -G. Andrei introduces an automated system leveraging force sensor data to discern gait patterns indicative of PD. Using dataset from the "Gait in Parkinson′s Disease" database, which comprises records from 93 PD patients and 73 healthy individuals, the study employs SVM algorithm for classification. Initial data preprocessing involves noise filtration, followed by feature extraction emphasizing parameters such as step length and various force metrics. Results indicate SVM′s effectiveness, achieving up to 100% accuracy in certain test groups and consistently outperforming traditional threshold-based methods. The research underscores the potential of combining gait analysis with machine learning for improved PD detection.

Detection of Parkinson's Disease from gait using Neighbourhood Representation Local Binary Patterns

The research by Yurdakul introduces a pioneering computational approach aimed at early Parkinson′s Disease (PD) detection, leveraging gait signals as a potential biomarker [31]. Utilizing the "Neighborhood Representation Local Binary Pattern" (NR-LBP) framework, "Vertical Ground Reaction Force" (VGRF) readings, indicative of gait characteristics, are processed and transformed to capture intricate gait patterns. Subsequently, significant statistical features, identified through the student's t-test, are fed into ANN for classification. Impressively, the proposed method achieved a classification accuracy of 98.3% and a Matthews Correlation Coefficient of 0.959, surpassing conventional techniques. This underscores the potential of machine learning in revolutionizing PD diagnostics, offering a robust and efficient alternative to traditional clinical methods.

Assisted Diagnosis of Parkinsonism Based on the Striatal Morphology

The research paper by Segovia aimed to devise a method leveraging striatal morphology as a biomarker for Parkinsonism diagnosis, utilizing neuroimaging data from DaTSCAN images of patients and controls [32]. Employing a SVM classifier and feature selection techniques, the study achieved an impressive accuracy rate exceeding 94% in distinguishing between Parkinsonian patients and controls. Notably, volumetric features of the striatum, such as volume and centroid, emerged as pivotal indicators for differentiation. Moreover, the Bhattacharyya distance was identified as the most effective feature selection method, outperforming Fisher's discriminant ratio and relative entropy. The study's results highlight the potential of striatal morphological measures as robust diagnostic markers, although its capability to discern among various parkinsonian syndromes remains a challenge due to overlapping neuroimaging profiles. Table 4 presents comparisons of the accuracy achieved in various studies discussed above. Figure 10 draws a comparison between accuracies achieved by different researchers.

Table 4. Comparison between various Machine Learning approaches

S.NO	Author Name	Year	Methodology	Input Data	Performance
1	Oduntan	2021	KDD & XGBoost	Biomedical voice measurements	Accuracy: 95%; Precision: 100% (Healthy), 94% (PD)
2	Hausdorff JM, Mitchell SL	2007	SVM with Gaussian RBF Kernel	Gait metrics from PD, ALS, HD patients, and controls	Accuracy: 83.33%; PD True Positive: 75%
3	O. Asmae, R. Abdelhadi	2020	ANN & KNN	Voice-based features (F0, jitter, shimmer, HNR)	ANN: 96.7% accuracy; KNN: 79.31% accuracy
4	T. J. Wroge	2020	mRMR, GeMaps & Various ML models	Voice data from mPower study	Accuracy: 86%; Gradient Boosted DT & ANN notable
5	Rana, A	2022	SVM, Naïve Bayes, K-NN, ANN	Voice recordings from diagnosed PD patients and healthy	ANN: 96.7%; K-NN: 87.17%
6	Taleb	2020	CNN & CNN-BLSTM with transfer learning & data augmentation	Handwriting data from HandPDMultiMC & PaHaW datasets	CNN-BLSTM: 97.62% accuracy
7	A. -G. Andrei	2019	SVM	Force sensor data from "Gait in Parkinson′s Disease" database	Up to 99% accuracy with SVM
8	Yurdakul	2020	NR-LBP & ANN	Vertical Ground Reaction Force (VGRF) readings	Accuracy: 98.3%; MCC: 0.959
9	Segovia	2019	SVM & Feature Selection	Neuroimaging data from DaTSCAN images of patients and controls	Accuracy: > 94%; Striatal volumetric features significant

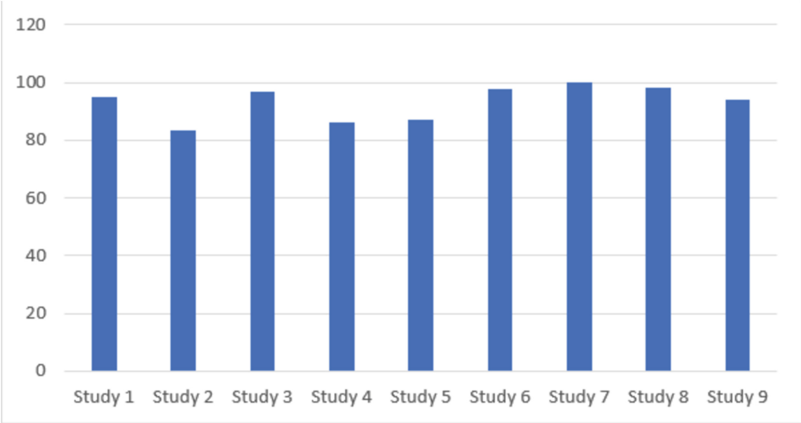

Fig. 10. Bar chart comparing accuracies of different methods.

4 Conclusion

In this comprehensive review, we delved into a multitude of machine learning approaches applied to the diagnosis of PD. From voice and gait analysis to neuroimaging techniques, each method exhibited its unique strengths and potential in enhancing early detection and differentiation of PD from other neurological disorders. Notably, the use of ML algorithms, such as XG Boost, SV M and ANN, showcased remarkable accuracies in various studies. The comparison reveals that different ML algorithms exhibit varying levels of performance depending on the type of dataset used. While ANN performed exceptionally well with voice datasets, SVM showed high accuracy when analyzing gait data. ML models trained on voice datasets have achieved high accuracy rates, with algorithms such as ANN demonstrating effectiveness in PD detection. The advancements in these methodologies not only emphasize the transformative potential of ML in healthcare but also underscore the urgent need for continued interdisciplinary research. As we move forward, integrating these insights into clinical practice could usher in a new era of personalized care for individuals with PD.

References

1. Parkinson, J.: An essay on the shaking palsy. J. Neuropsychiatry Clin. Neurosci. **14**(2), 223–222 (2002)
2. Samii, A., Nutt, J.G., Ransom, B.R.: Parkinson's disease. Lancet (London, England) **363**(9423), 1783–1793 (2004)
3. de Lau, L.M., Breteler, M.M.: Epidemiology of Parkinson′s disease. Lancet Neurol. **5**(6), 525–535 (2006)
4. Morris, E.M.: Movement disorder in people with Parkinson disease: a model for physical therapy. Phys. Ther. **80**, 578–597 (2000)
5. Schrag, A., et al.: Differentiation of atypical parkinsonian syndromes with routine MRI. Neurology **54**(3), 697–702 (2000)

6. Angel, R.W., Alston, W., Higgins, J.R.: Control of movement in Parkinson′s disease. Brain: J. Neurol. **93**(1), 1–14 (1970)
7. Wu, S.L., Liscic, R.M., Kim, S., Sorbi, S., Yang, Y.H.: Nonmotor symptoms of Parkinson′s disease. Parkinson′s Dis. (2017)
8. Yousaf, T., Wilson, H., Politis, M.: imaging the nonmotor symptoms in parkinson′s disease. In: International Review of Neurobiology, pp. 179–257 (2017)
9. UCI Machine Learning Repository, 'Parkinson Data Set'. Center for Machine Learning and Intelligent Systems. https://archive.ics.uci.edu/ml/datasets/parkinsons
10. Little, M.A., McSharry, P.E., Hunter, E.J., Spielman, J., Ramig, L.O.: Suitability of dysphonia measurements for telemonitoring of Parkinson's disease. IEEE Trans. Biomed. Eng. **56**(4), 1015 (2009)
11. Hausdorff, J.M., Lertratanakul, A., Cudkowicz, M.E., Peterson, A.L., Kaliton, D., Goldberger, A.L.: Dynamic markers of altered gait rhythm in amyotrophic lateral sclerosis. J. Appl. Physiol. **88**(6), 2045–2053 (2000)
12. Bot, B., Suver, C., Neto, E., et al.: The mPower study, Parkinson disease mobile data collected using researchKit. Sci Data **3**, 160011 (2016)
13. Taleb, C., Khachab, M., Mokbel, C., Likforman-Sulem, L.: Feature selection for an improved Parkinson′s disease identification based on handwriting. In: 1st International Workshop on Arabic Script Analysis and Recognition (ASAR), pp. 52–56 (2017)
14. Drotár, P., Mekyska, J., Rektorová, I., Masarová, L., Smékal, Z., Faundez-Zanuy, M.: Evaluation of handwriting kinematics and pressure for differential diagnosis of Parkinson's disease. Artif. Intell. Med. **67**, 39–46 (2016)
15. Hausdorff, J.M., Lowenthal, J., Herman, T., Gruendlinger, L., Peretz, C., Giladi, N.: Rhythmic auditory stimulation modulates gait variability in Parkinson's disease. Eur. J. Neurosci. **26**(8), 2369–2375 (2007)
16. Towey, D.J., Bain, P.G., Nijran, K.S.: Automatic classification of 123I-FP-CIT (DaTSCAN) SPECT images. Nucl. Med. Commun. **32**(8), 699–707 (2011)
17. Gelb, D.J., Oliver, E., Gilman, S.: Diagnostic criteria for Parkinson disease. Arch. Neurol. **56**(1), 33–39 (1999)
18. Tsanas, A., Little, M.A., McSharry, P.E., Ramig, L.O.: Accurate telemonitoring of Parkinson's disease progression by noninvasive speech tests. IEEE Trans. Biomed. Eng. **57**(4), 884–893 (2010)
19. Sapir, S., Spielman, J.L., Ramig, L.O., Story, B.H., Fox, C.: Effects of intensive voice treatment (the lee silverman voice treatment [LSVT]) on vowel articulation in dysarthric individuals with idiopathic Parkinson disease: acoustic and perceptual findings. J. Speech Lang. Hear. Res. **50**(4), 899–912 (2007)
20. Rahn, D.A., 3rd., Chou, M., Jiang, J.J., Zhang, Y.: Phonatory impairment in Parkinson's disease: evidence from nonlinear dynamic analysis and perturbation analysis. J. Voice: Official J. Voice Found. **21**(1), 64–71 (2007)
21. Baken, R., Orlikoff, R.F.: Clinical measurement of speech & voice. Singular Publishing Group, Albany, NY (1999)
22. Wroge, T.J., Ghomi, R.H.: An analysis of automated Parkinson's diagnosis using voice: methodology and future directions. In: Obeid, I., Selesnick, I., Picone, J. (eds.) Signal Processing in Medicine and Biology. Springer, Cham (2020)
23. Bang, C., Bogdanovic, N., Deutsch, G., et al.: Machine learning for the diagnosis of Parkinson's disease using speech analysis: a systematic review. Int. J. Speech Technol. **26**, 991–998 (2023)
24. Oduntan, I.: Prediction of Parkinson′s disease using biomedical voice measurements dataset (2021)
25. Hausdorff, J.M., et al.: Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. J Appl. Physiol. 262–269 (1997)

26. Asmae, O., Abdelhadi, R., Bouchaib, C., Sara, S., Tajeddine, K.: Parkinson's disease identification using KNN and ANN algorithms based on voice disorder. In: 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Meknes, Morocco (2020)
27. Wroge, T.J., Özkanca, Y., Demiroglu, C., Si, D., Atkins, D.C., Ghomi, R.H.: Parkinson's Disease Diagnosis Using Machine Learning and Voice. IEEE Signal Processing in Medicine and Biology Symposium (SPMB), Philadelphia, PA, USA (2018)
28. Rana, A., Dumka, A., Singh, R., Panda, M.K., Priyadarshi, N.: A Computerized analysis with machine learning techniques for the diagnosis of Parkinson's disease: past studies and future perspectives. Diagnostics **12**, 2708 (2022)
29. Taleb, C., Likforman-Sulem, L., Mokbel, C., et al.: Detection of Parkinson's disease from handwriting using deep learning: a comparative study. Evol. Intel. **16**, 1813–1824 (2023)
30. Andrei, A.G., Tăuțan, A.M., Ionescu, B.: Parkinson's disease detection from gait patterns. In: 2019 E-Health and Bioengineering Conference (EHB), Iasi, Romania (2019)
31. Yurdakul, O.C., Subathra, M., George, S.T.: Detection of Parkinson's disease from gait using neighborhood representation local binary patterns. Biomed. Sig. Process. Control **62** (2020)
32. Segovia, F., Gorriz, J.M., Ramirez, J., Martinez-Murcia, F.J., Castillo-Barnes, D.: Assisted diagnosis of parkinsonism based on the striatal Morphology. Int. J. Neural Syst. **29** (2019)