**A state-of-the-art machine learning review for Parkinson’s disease detection**

Nisha Kumari, Rahul Manipal, Altaf Siddique, Shreya Vaish, Sandeep Kumar\*, Mohit Arora

Department of Computer Science and Engineering, Lovely Professional University, Phagwara, INDIA

E-mail: devnishakumari823@gmail.com, 18rahulmanipal23@gmail.com, altafraza896@gmail.com, shreyavaish1516@gmail.com, sandeep.28953@lpu.co.in, mohit.15980@lpu.co.in

**Abstract**

It is significant to know how serious Parkinson's disease is and to get a diagnosis as soon as possible. The goal of this review article mainly focuses on the existing condition of research on Parkinson's disease. We did a systematic review of several research papers to figure out how to identify the Parkinson's disorder in early stages and treat people as soon as possible. In our analysis, we talk about how different machine learning (ML) algorithms can be used to spot Parkinson's disease. These algorithms can analyse large datasets from clinical and neuroimaging studies, Parkinson's telemonitoring vocal data collection, tappy keystroke datasets, the UCI ML repository, etc. Deep learning (DL) algorithms like deep convolutional neural networks, deep learning cascaded neural networks, long short-term memory, etc. provided significant accuracy for the diagnosis of PD from MRI images. Few studies present accuracy rates in the range of 95–99% to assess the accuracy of auditory speech characteristics in Parkinson’s victims.

**Keywords:** Parkinson disease (PD), CNN, machine learning, review*,* deep learning

1. **Introduction**

Parkinson's disease is a degenerative neurological disorder which attacks millions of people around the world. One potential way to detect Parkinson’s is the through use of ML [1] and DL techniques, that has the capabilities of analyzing huge quantities of data and recognize trends that humans may find challenging to detect. As per the report by the Parkinson's Foundation, almost ten million people globally are having Parkinson's disease, thus making this an alarming state of international affairs. Several studies exploring the use of machine learning and deep learning for PD detection. One approach is to analyze speech patterns, as PD can affect the vocal cords and cause changes in speech. A study published in the Journal of Neural Engineering used a deep learning algorithm to analyze speech data from people infected by Parkinson’s disease and was able to accurately distinguish between the two groups with a high level of accuracy. Another way is to look at how people walk, since PD can also affect how people move and stay balanced. A study published in the Journal of Neuroscience Methods used machine learning algorithms to look at gait data from people with PD and healthy controls. The algorithms were able to correctly put people into their correct groups. Other studies have looked at how machine learning [2] and deep learning could be used to diagnose PD using a combination of biomarkers, such as imaging data or genetic information. A study in the Journal of Parkinson's Disease used a machine learning algorithm to look at imaging data and genetic information about a group of patients and was able to accurately predict whether or not they had PD. Machine learning algorithms can swiftly and reliably analyze enormous amounts of data, enabling early disease identification and treatment. The accuracy and dependability of diagnoses can be increased by using machine learning algorithms to find patterns and links in data that are not immediately obvious to human observers. The algorithms that have been presented are based on a variety of methods, including CNN, LSTM [3], and ensemble classifiers[4]. According to numerous studies, 90% of PD patients have speaking and vocal issues, including dysphonia, monotone, and hypophonia.

1. **Literature Review**

Parkinson's disease is incurable, but early signs of problems can be found and treated with the help of a few ML algorithms[5]. The researchers used the best ML algorithms[6], such as KNN, SVM, GBM, RF[7], ANN[8][9], LSTM, etc., in several studies to look for PD. In a few studies, machine learning algorithms were used to turn multimodal brain imaging data into a diagnostic tool for PD. Nalini et.al [10] proposed a study to determine the accuracy of PD patients’ audial speech features using three types of deep-learning CNN models. Pooja Raundale et.al [11] proposed a study to predict the severity of PD using deep Neural Network and Random Forest Classifier. Wu et.al [12] studied to detect PD at an early stage, and Arti et.al focused on to address common PD symptoms such as Bradykinesia using SVM classifier that achieved the highest precision in audial characteristics for identification i.e., 91.4% accuracy. Iqra et.al [13] discussed Ensemble ML classifier studied the effectiveness of PD detection methods based on voice [14][15] using different ml techniques such as XGBoost, Logistic Regression (LR), Naive Bayes (NB), RF, KNN, Decision Tree (DT), SVM [16], and MLP. Recursive Feature Elimination and Min-Redundancy and Max-Relevance techniques were also used for selection of features. Jefferson et.al proposed a study on voice signal processing for detection of Parkinson’ using four ML methods i.e., OPF, KNN, MLP, and SVM. Satyabrata et al. [17] described a study on wearable assessment system with a combination of ML methods that can autonomously distinguish between PD patient’s raw accelerometer gait signals obtained by the wearable accelerometer-based sensor on and off modes. You et al. [18] conducted a study aiming to use machine learning (ML) techniques. The most important details in this text are the studies that have been conducted to identify Parkinson's Disease using ICD-code (G20). The best average prediction accuracy was shown by NN, GBM & RF Algorithms, with Male having better entire model performance metrics than Female. Ensemble ML [19] approaches such as boosting algorithms can significantly contribute to the PD detection. Kamal et.al studied that over 5 million people nationwide suffer with Parkinson's disease whereas only about 200,000 people in the United States. Lerina et al. [20] introduced a new way of research exploration for Parkinson's disease diagnosis and monitoring. This study used LSTM and Convolutional Neural Network to detect PD. Guidong et.al [21] used SSCL to classify vocal patterns in PD. Marta et.al used ML algorithms to differentiate matched fit particulars from Parkinson's patients. The Naive Bayes method produced 0.923 precision and the Random Forest achieved an ARUC 0.786, outperforming other examined ML algorithms in terms of PD stage detection. Alex Li et al. [22] proposed a machine learning classifier using force sensor data from 214 individuals with idiopathic PD and 92 healthy subjects from the PhysioNet dataset Gait in Parkinson's Disease. The classifier performed extremely well, earning an excellent recall value. Majid et.al proposed novel, effective common spatial- pattern-based methods to detect PD, with 99% sensitivity, specificity and accuracy. The greatest classification accuracy was found in features taken from the alpha and beta bands. The summarized related work is given in Table 1.

Table 1: Literature Review Summary

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Author Name** | **Algorithm Used** | **Best Performer** | **Performance Metrics** | **Datasets Used** |
| Nalini et al. [10] | Deep-learning cascaded neural network models (RNN, MLP, LSTM) | LSTM Model with (~99%) Accuracy | Accuracy (99%),  Loss (3.5)  Precision (100%),  Recall (97.95%),  F1 score (98.96%) | UCI machine learning repository |
| Pooja et al. [11] | XGBoost, ANN & Random Forest Classifier | XG Boost classifier with (~95%) Accuracy | Accuracy (95%) | Parkinson's Telemonitoring vocal data and Tappy keystroke |
| Wu et al. [12] | Deep learning model and Ensemble learning method | Ensemble Network method with (~96.45%) Accuracy | Accuracy (96.68%)  Precision (97.67%)  Sensitivity (97.52%)  Specificity (94.84%)  F1 score (97.58%) | Parkinson's Progression Markers Initiative datasets |
| Arti et al. [23] | KNN, SVM, ANN, CART, Naïve Bayes, decision tree, LR, etc,. | SVM classifier  (~91.4%)  Accuracy | Accuracy (91.4%) | Scientometric data [2022] |
| Iqra et al. [13] | MLP, Naive Bayes, RF, KNN, Decision Tree, Logistic Regression, XGBoost and SVM Algorithm | (~95.39%)  Accuracy performed by XGBoost  (mRMR feature-selection) | Accuracy (95.39%)  Precision (95%)  Recall (95%)  F1 score (95%) | UCI machine learning repository |
| Jefferson et al. [24] | KNN, MLP, OPF, and SVM | KNN | Accuracy (94.55%)  Specificity (94.26%)  Sensitivity (94.55%) | Datasets of speech and phonation |
| Satyabrata et al. [17] | RF, K Nearest Neighbor , SVM and Naive Bayes Machine Learning Algorithm | Random Forest classifier with (~96.72%)  Accuracy | Accuracy (96.72%)  Avg Precision (96.92%)  Avg F1 Score (0.97)  Avg Recall (97.35%) | 20 distinct Parkinson's disease patients data were gathered |
| You et al. [18] | Neural Network Algorithm, Gradient boosting Algorithm,  Random Forest Algorithm | Neural Network Algorithm with (77.9%) accuracy | Accuracy (77.9%)  Specificity (0.698)  Sensitivity (0.677)  Recall (0.677)  Precision (0.695) | National Health Insurance Service-Health Screening datasets |
| Anirudha et al. [25] | SVM, KNN, Neural Network & non-Neural based Machine Learning Algorithm | Convolutional Neural Network Algorithm | Accuracy (89%)  F1 score (0.61)  Precision (0.61) | PubMed Database of the US national library of medicine |
| Mirza et al. [26] | AdaBoost, Gradient Boosting, Light GBM XGBoost Algorithm | Light GBM | Accuracy (93.39%) | UCI machine learning repository |
| Kamal et al. [27] | Boosted Logistic Regression algorithm & Random Forest, Bayesian Network, Multilayer Perceptron (artificial neural network) | Boosted Logistic Regression | Accuracy (97.159%)  Recall (0.972)  Precision (0.974) | Biomedical voice of human |
| Lerina et al. [20] | Convolutional neural network & LSTM | LSTM | Accuracy (97%) | Italian Parkinson’s voice speech datasets |
| Guidong et al. [21] | SVM, KNN & Semi-supervised competitive learning algorithm | SSCL | Accuracy (83.8%)  Specificity (85%)  Precision (84.6%)  F1-Score (83.5%) | Parkinson Datasets with replicated acoustic features Datasets |
| Marta et al. [2] | Naive Bayes, RF algorithm | RF | Accuracy (84.6%)  Precision (92.3%)  Recall (80%) | Movement Disorder Clinics |
| Alex Li et al. [22] | Logistic regression, SVM, decision tree, KNN | SVM | Accuracy (85%)  Precision (0.86)  Recall (0.94) | Gait in Parkinson’s Disease datasets |
| Majid et al. [28] | Random forest, SVM, and  knn Algorithm | KNN | OFF-Medication Parkinson’s Diagnosis with Accuracy (99.41%)  Sensitivity (99.47%),  Specificity (99.35%),  F1-Score (99.40%),  On medication PD detection Accuracy (95% to 98%) | EEG Datasets divided in two form UNM and SanDiego datasets |
| Dafa et al. [29] | SVM classifier and ALFF-based radiomics method | SVM classifier with primary set dataset achieve highest accuracy (81.45%) | Accuracy (81.45%)  Sensitivity (86.86%),  Specificity (73.66%),  F1-Score (83.68%), | NEUROCON datasets (primary set and external validation set) |
| Ārti et al. [30] | SVM, KNN and Artificial neural network | ANN (artificial neural network) | Accuracy (96.7%)  F1-Score (87.01%), | Max Little of the University of Oxford provided datasets of taped speech patterns. |
| Mehrbakhsh et al. [31] | EM clustering, SOM ensemble algorithm, SVR algorithm, HGPA approach | (HGPA+EM+SVR) perform better than (HGPA+SOM+SVR) | Accuracy (56%) | Real-world PD datasets |
| Anant et al. [32] | Unsupervised and Supervised (Random Forest, LGBM, XGBoost) machine learning | Supervised ML algorithm | Accuracy (92%) | Data of longitudinal clinical collected from PPMI datasets |
| Ana et al. [33] | SVM and ANN | SVM | Accuracy (97.35%) | PRISMA methodology |
| Aditi Govindua et al. [34] | SVM, KNN, RF and Logistic Regression | Random Forest Classifier | Accuracy (91.83%)  Recall (0.86)  Precision (0.95) | Audio data from UCI machine learning Repository & PPMI |
| Zehra et al. [35] | ANN, SVM & Regression Trees algorithm | Support Vector Machines | Accuracy (93.84%) | Speech signal of 31 people dataset |

We have compared machine learning algorithms based on the accuracy in Figure 1.

Figure 1: Performance Evaluation of ML Algorithms

Figure 2: Datasets Used in the Literature Review of Existing Parkinson disease detection algorithms

In Figure 2, pie chart representation is shown in order to display the use of datasets in the literature.

1. **Conclusion & Future Scope**

By doing thorough review of the papers, meet with a conclusion that machine learning models have showed potential in detecting PD by employing extensive data sources such as voice recordings, handwriting samples, and gait analyses. These techniques have been trained to recognize patterns and features in the data that are indicative of PD, which may be utilized to establish accurate diagnostic methods. Overall, artificial intelligence algorithms hold great potential in improving the accuracy and efficiency of PD diagnosis, which could lead to earlier interventions and better outcomes for patients. KNN algorithm provides better performance than others on the various parameters of performance metrics such as Specificity, Accuracy, Sensitivity & F1-score but Accuracy is the most common in all the research paper. Most common dataset used in all the research paper are UCI Machine Learning Repository. Currently, no known cure is available for Parkinson's disease but, there exists several medicaments which can assists in controlling the ailments and improves the quality of life. Ongoing research and development of new therapies and treatments hold promise for better outcomes for those living with Parkinson's disease in the future. Still, more study is required to confirm the efficacy of such algorithms within clinical contexts, as well as to highlight the difficulties about data unpredictability and sample size.

**References**

[1] A. M. Maitín, A. J. García-Tejedor, and J. P. R. Muñoz, “Machine learning approaches for detecting parkinson’s disease from eeg analysis: A systematic review,” *Appl. Sci.*, vol. 10, no. 23, pp. 1–21, 2020, doi: 10.3390/app10238662.

[2] dkk 2018 ) richard oliver ( dalam Zeithml., “Machine learning models for Parkinson’s disease detection and stage classification based on spatial-temporal gait parameters Marta,” *Angew. Chemie Int. Ed. 6(11), 951–952.*, pp. 2013–2015, 2021.

[3] S. Bhati, L. M. Velazquez, J. Villalba, and N. Dehak, “LSTM siamese network for Parkinson’s disease detection from speech,” in *GlobalSIP 2019 - 7th IEEE Global Conference on Signal and Information Processing, Proceedings*, 2019, pp. 11–15. doi: 10.1109/GlobalSIP45357.2019.8969430.

[4] J. M. Templeton, C. Poellabauer, and S. Schneider, “Classification of Parkinson’s disease and its stages using machine learning,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-18015-z.

[5] L. Parisi, N. RaviChandran, and M. L. Manaog, “Feature-driven machine learning to improve early diagnosis of parKinson’s disease,” *Expert Syst. Appl.*, vol. 110, pp. 182–190, 2018, doi: 10.1016/j.eswa.2018.06.003.

[6] C. Mittal and A. Sharma, “Parkinson’s Disease Detection Using Different Machine Learning Algorithms,” *Int. J. Sci. Res. Publ.*, vol. 12, no. 2, pp. 23–27, 2022, doi: 10.29322/ijsrp.12.02.2022.p12205.

[7] K. Açıcı, Ç. B. Erdaş, T. Aşuroğlu, M. K. Toprak, H. Erdem, and H. Oğul, “A random forest method to detect parkinson’s disease via gait analysis,” *Commun. Comput. Inf. Sci.*, vol. 744, no. June 2018, pp. 609–619, 2017, doi: 10.1007/978-3-319-65172-9\_51.

[8] F. SoleimanianGharehchopogh and P. Mohammadi, “A Case Study of Parkinson&apos;s Disease Diagnosis using Artificial Neural Networks,” *Int. J. Comput. Appl.*, vol. 73, no. 19, pp. 1–6, 2013, doi: 10.5120/12990-9206.

[9] M. Belić, V. Bobić, M. Badža, N. Šolaja, M. Đurić-Jovičić, and V. S. Kostić, “Artificial intelligence for assisting diagnostics and assessment of Parkinson’s disease—A review,” *Clin. Neurol. Neurosurg.*, vol. 184, 2019, doi: 10.1016/j.clineuro.2019.105442.

[10] N. Chintalapudi, G. Battineni, M. A. Hossain, and F. Amenta, “Cascaded Deep Learning Frameworks in Contribution to the Detection of Parkinson’s Disease,” *Bioengineering*, vol. 9, no. 3, 2022, doi: 10.3390/bioengineering9030116.

[11] P. Raundale, C. Thosar, and S. Rane, “Prediction of parkinson’s disease and severity of the disease using machine learning and deep learning algorithm,” *2021 2nd Int. Conf. Emerg. Technol. INCET 2021*, no. May 2021, 2021, doi: 10.1109/INCET51464.2021.9456292.

[12] W. Wang, J. Lee, F. Harrou, and Y. Sun, “Early Detection of Parkinson’s Disease Using Deep Learning and Machine Learning,” *IEEE Access*, vol. 8, pp. 147635–147646, 2020, doi: 10.1109/ACCESS.2020.3016062.

[13] I. Nissar, D. R. Rizvi, S. Masood, and A. N. Mir, “Voice-based detection of parkinson’s disease through ensemble machine learning approach: A performance study,” *EAI Endorsed Trans. Pervasive Heal. Technol.*, vol. 5, no. 19, pp. 1–8, 2019, doi: 10.4108/eai.13-7-2018.162806.

[14] D. Braga, A. M. Madureira, L. Coelho, and R. Ajith, “Automatic detection of Parkinson’s disease based on acoustic analysis of speech,” *Eng. Appl. Artif. Intell.*, vol. 77, no. September 2018, pp. 148–158, 2019, doi: 10.1016/j.engappai.2018.09.018.

[15] A. G. Mohammadi, P. Mehralian, A. Naseri, and H. Sajedi, “Parkinson’s disease diagnosis: The effect of autoencoders on extracting features from vocal characteristics,” *Array*, vol. 11, no. May 2020, p. 100079, 2021, doi: 10.1016/j.array.2021.100079.

[16] S. Shetty and Y. S. Rao, “SVM based machine learning approach to identify Parkinson’s disease using gait analysis,” in *Proceedings of the International Conference on Inventive Computation Technologies, ICICT 2016*, 2016, vol. 2. doi: 10.1109/INVENTIVE.2016.7824836.

[17] S. Aich *et al.*, “A supervised machine learning approach to detect the On/Off state in Parkinson’s disease using wearable based gait signals,” *Diagnostics*, vol. 10, no. 6, pp. 1–18, 2020, doi: 10.3390/diagnostics10060421.

[18] Y. H. Park *et al.*, “Machine learning based risk prediction for Parkinson’s disease with nationwide health screening data,” *Sci. Rep.*, vol. 12, no. 1, pp. 1–11, 2022, doi: 10.1038/s41598-022-24105-9.

[19] Institute of Electrical and Electronics Engineers and V. & P. R. (5th : 2021 : J. O. International Conference on Imaging, “Detection of Parkinson’s Disease by Employing Boosting Algorithms,” in *2021 Joint 10th International Conference on Informatics, Electronics & Vision (ICIEV) & 2021 5th International Conference on Imaging, Vision & Pattern Recognition (icIVPR)*, 2021.

[20] L. Aversano, M. L. Bernardi, M. Cimitile, M. Iammarino, D. Montano, and C. Verdone, “A Machine Learning approach for Early Detection of Parkinson’s Disease Using acoustic traces,” in *IEEE Conference on Evolving and Adaptive Intelligent Systems*, 2022, vol. 2022-May, no. November. doi: 10.1109/EAIS51927.2022.9787728.

[21] G. Bao, M. Lin, X. Sang, Y. Hou, Y. Liu, and Y. Wu, “Classification of Dysphonic Voices in Parkinson’s Disease with Semi-Supervised Competitive Learning Algorithm,” *Biosensors*, vol. 12, no. 7, 2022, doi: 10.3390/bios12070502.

[22] A. Li and C. Li, “Detecting Parkinson’s Disease through Gait Measures Using Machine Learning,” *Diagnostics*, vol. 12, no. 10, pp. 1–9, 2022, doi: 10.3390/diagnostics12102404.

[23] A. Rana, A. Dumka, R. Singh, M. K. Panda, N. Priyadarshi, and B. Twala, “Imperative Role of Machine Learning Algorithm for Detection of Parkinson’s Disease: Review, Challenges and Recommendations,” *Diagnostics*, vol. 12, no. 8, 2022, doi: 10.3390/diagnostics12082003.

[24] J. S. Almeida *et al.*, “Detecting Parkinson’s disease with sustained phonation and speech signals using machine learning techniques,” *Pattern Recognit. Lett.*, vol. 125, pp. 55–62, 2019, doi: 10.1016/j.patrec.2019.04.005.

[25] A. S. Chandrabhatla, I. J. Pomeraniec, and A. Ksendzovsky, “Co-evolution of machine learning and digital technologies to improve monitoring of Parkinson’s disease motor symptoms,” *npj Digit. Med.*, vol. 5, no. 1, pp. 1–18, 2022, doi: 10.1038/s41746-022-00568-y.

[26] M. M. Nishat, T. Hasan, S. M. Nasrullah, F. Faisal, M. A. A. R. Asif, and M. A. Hoque, “Detection of Parkinson’s Disease by Employing Boosting Algorithms,” in *2021 Joint 10th International Conference on Informatics, Electronics and Vision, ICIEV 2021 and 2021 5th International Conference on Imaging, Vision and Pattern Recognition, icIVPR 2021*, 2021, no. October. doi: 10.1109/ICIEVICIVPR52578.2021.9564108.

[27] K. N. R. Challa, V. S. Pagolu, G. Panda, and B. Majhi, “An improved approach for prediction of Parkinson’s disease using machine learning techniques,” in *International Conference on Signal Processing, Communication, Power and Embedded System, SCOPES 2016 - Proceedings*, 2017, pp. 1446–1451. doi: 10.1109/SCOPES.2016.7955679.

[28] M. Aljalal, S. A. Aldosari, K. Alsharabi, A. M. Abdurraqeeb, and F. A. Alturki, “Parkinson’s Disease Detection from Resting-State EEG Signals Using Common Spatial Pattern, Entropy, and Machine Learning Techniques,” *Diagnostics*, vol. 12, no. 5, 2022, doi: 10.3390/diagnostics12051033.

[29] D. Shi *et al.*, “Machine Learning for Detecting Parkinson’s Disease by Resting-State Functional Magnetic Resonance Imaging: A Multicenter Radiomics Analysis,” *Front. Aging Neurosci.*, vol. 14, no. March, pp. 1–12, 2022, doi: 10.3389/fnagi.2022.806828.

[30] A. Rana, A. Dumka, R. Singh, M. Rashid, N. Ahmad, and M. K. Panda, “An Efficient Machine Learning Approach for Diagnosing Parkinson’s Disease by Utilizing Voice Features,” *Electron.*, vol. 11, no. 22, pp. 1–20, 2022, doi: 10.3390/electronics11223782.

[31] M. Nilashi *et al.*, “Predicting Parkinson’s Disease Progression: Evaluation of Ensemble Methods in Machine Learning,” *J. Healthc. Eng.*, vol. 2022, pp. 9–11, 2022, doi: 10.1155/2022/2793361.

[32] A. Dadu *et al.*, “Identification and prediction of Parkinson’s disease subtypes and progression using machine learning in two cohorts,” *npj Park. Dis.*, vol. 8, no. 1, pp. 1–12, 2022, doi: 10.1038/s41531-022-00439-z.

[33] A. M. Maitin, J. P. Romero Muñoz, and Á. J. García-Tejedor, “Survey of Machine Learning Techniques in the Analysis of EEG Signals for Parkinson’s Disease: A Systematic Review,” *Appl. Sci.*, vol. 12, no. 14, 2022, doi: 10.3390/app12146967.

[34] A. Govindu and S. Palwe, “Early detection of Parkinson’s disease using machine learning,” *Procedia Comput. Sci.*, vol. 218, no. 2022, pp. 249–261, 2023, doi: 10.1016/j.procs.2023.01.007.

[35] Z. Karapinar Senturk, “Early diagnosis of Parkinson’s disease using machine learning algorithms,” *Med. Hypotheses*, vol. 138, p. 109603, 2020, doi: 10.1016/j.mehy.2020.109603.