Salud, Ciencia y Tecnología - Serie de Conferencias. 2025; 4:1353
doi: 10.56294/sctconf20251353

SYSTEMATIC REVIEW

Deep learning vs. conventional methods for Parkinson's disease diagnosis: A Systematic Review

Aprendizaje profundo vs. métodos convencionales para el diagnóstico de la enfermedad de Parkinson: Una Revisión Sistemática

Sowjanya Vuddanti[1], A. Srinagesh[2]

[1]Department of Computer Science and Engineering, Acharya Nagarjuna University, Guntur(Dt).
[2]R.V.R. & J.C. College of Engineering.

Cite as: Vuddanti S, Srinagesh A. Deep learning vs. conventional methods for Parkinson's disease diagnosis: a systematic review. Salud, Ciencia y Tecnología - Serie de Conferencias. 2025; 4:1353. https://doi.org/10.56294/sctconf20251353

Submitted: 29-05-2024 **Revised:** 08-09-2024 **Accepted:** 23-02-2025 **Published:** 24-02-2025

Editor: Prof. Dr. William Castillo-González

Corresponding Author: Sowjanya Vuddanti

ABSTRACT

A neurological condition called Parkinson's disease (PD) primarily affects movement, but it can also have an impact on speaking, thinking, and a host of other bodily processes. Machine learning models can be trained by systems to examine clinical data, genetic information, speech patterns, and even speech patterns in order to identify early indicators of Parkinson's disease before symptoms manifest. One of the main issues with machine learning models is their inability to handle inconsistent, noisy, or missing input, which can have a negative effect on the model's performance. By building a system that supports both transfer learning techniques and multi-modal fusion, these shortcomings can be addressed. In order to determine the model's efficacy, this study examines many deep learning techniques based on speech, image, and handwritten patterns. In order to improve diagnosis accuracy, deep learning techniques can look at complex data patterns from a range of sources, such as speech, signals, images of medical conditions, and walking patterns. By using convolutional neural networks, recurrent neural networks, and transfer learning, deep learning models are able to identify Parkinson's disease early on, monitor its progression, and offer personalized treatment. Traditional Parkinson's disease diagnosis techniques rely on manually defined features extracted from a range of data sources, such as speech, gait, and medical images. These characteristics are subsequently incorporated into machine learning models. To automatically detect and extract aspects of Parkinson's disease, deep learning approaches make use of transfer learning and end-to-end learning.

Keywords: Movement Disorder; Transfer Learning; End-to-End Learning; Genetic Information; Gait Patterns; Class Imbalance; Disease Progression Tracking.

RESUMEN

Una afección neurológica llamada enfermedad de Parkinson (EP) afecta principalmente al movimiento, pero también puede repercutir en el habla, el pensamiento y otros muchos procesos corporales. Los modelos de aprendizaje automático pueden ser entrenados por sistemas para examinar datos clínicos, información genética, patrones del habla e incluso del habla con el fin de identificar indicadores tempranos de la enfermedad de Parkinson antes de que se manifiesten los síntomas. Uno de los principales problemas de los modelos de aprendizaje automático es su incapacidad para manejar datos inconsistentes, ruidosos o ausentes, lo que puede tener un efecto negativo en el rendimiento del modelo. La creación de un sistema compatible con técnicas de aprendizaje por transferencia y fusión multimodal permite subsanar estas deficiencias. Para determinar la eficacia del modelo, este estudio examina muchas técnicas de aprendizaje profundo basadas en patrones de voz, imagen y escritura. Con el fin de mejorar la precisión del diagnóstico, las técnicas de aprendizaje profundo pueden examinar patrones de datos complejos procedentes de diversas fuentes, como el habla, las señales, las imágenes de afecciones médicas y los patrones de marcha. Mediante el uso de redes neuronales convolucionales, redes neuronales recurrentes y aprendizaje por transferencia, los modelos de aprendizaje profundo son capaces de identificar la enfermedad de Parkinson en una fase temprana, controlar su progresión y ofrecer un tratamiento personalizado. Las técnicas tradicionales de diagnóstico de la enfermedad de Parkinson se basan en características definidas manualmente y extraídas de una serie de fuentes de datos, como el habla, la marcha y las imágenes médicas. Estas características se incorporan posteriormente a modelos de aprendizaje automático. Para detectar y extraer automáticamente aspectos de la enfermedad de Parkinson, los enfoques de aprendizaje profundo hacen uso del aprendizaje por transferencia y del aprendizaje de extremo a extremo.

Palabras clave: Trastorno del Movimiento; Aprendizaje de Transferencia; Aprendizaje de Extremo a Extremo; Información Genética; Patrones de la Marcha; Desequilibrio de Clases; Seguimiento de la Progresión de la Enfermedad.

INTRODUCTION

PD detection utilizing ML methods is indispensable for the analysis of non-image data, including speech patterns, motor activity, and ubiquitous sensor data.(20) Highly successful in processing structured, tabular data, ML models include SVMs, DT, RF, & Gradient Boosting methods differ from CNNs, which are specialized for image-based data. These models may operate using data gathered from clinical assessments such tests measuring motor performance, cognitive capacity, or speech abnormalities frequent in Parkinson's disease patients. ML models are also very important for analyzing data from wearable devices that keep track of a patient's walking, shakes, moves, and other motor-related activities over time. This real-time information can then be utilized to identify alterations that could point to the start of Parkinson's disease or follow the course of the illness in already diagnosed individuals. ML is also good for finding PD because it can handle big and complicated datasets and check to see if a patient shows signs of the illness by selecting features and classifying them.

Even if ML is becoming more and more successful in PD identification, certain constraints still exist. One main restriction is the calibre and volume of the accessible data. Large datasets are necessary for training many ML algorithms in order to guarantee accuracy; yet, as longitudinal patient data sets are uncommon, models may over fit or perform poorly when applied to new data.(19) The great complexity of clinical & sensor data is another restriction that could complicate feature selection & raise the danger of incorporating duplicate or useless characteristics. Moreover, ML models usually need manual feature extraction, in which case specialists must pre-select relevant data for analysis a labor-intensive and prone to human error process. Variations in the positioning, quality, or calibration of wearable sensors may add noise and impact the performance of the model.

Detection using Deep Learning Techniques

The intricacy of PD and its impact on brain anatomy and motor function has made CNN-based PD identification all the more crucial. several times resulting in structural or functional abnormalities shown in medical pictures such as MRIs, DaTscans, or functional MRIs, PD affects several parts of the brain.(11) Being very good at processing visual data, CNNs can examine these medical photos and find minor characteristics that can point to early illness symptoms. Because of their hierarchical structure, CNNs are especially well-suited for the complex job of PD detection as they can automatically extract both high-level and low-level information without requiring human intervention. This extraction of features methodology is essential as early phases of Parkinson's disease may show patterns too faint for human interpretation or conventional image analysis techniques. Additionally, CNNs are better at generalising on big datasets because they can learn from millions of tiny details at the pixel level, which makes recognition more accurate.(12)

A DNN classifier using speech difficulties, a major early sign of Parkinson's disease. Two primary components make up the DNN: a softmax classifier for classifying and a sAE for feature extraction. As the sAE lowers the number of dimensions in the input data, it becomes easier to learn hidden features. The softmax classifier, on the other hand, gives odds for classification. The Limited Memory BFGS optimization technique is used to optimize the network's parameters as it is worked on speech info to PD patients & healthy controls. For a autoencoders, the training procedure consists of unsupervised learning; for the softmax layer, it comprises supervised learning. Two datasets the PSD which contains a variety of sound recordings (PSD) and the OPD dataset are used in testing this approach. By means of speech analysis, the DNN shows higher accuracy and the capacity to identify PD at early stages than conventional classifiers like SVM, DT, & Naive Bayes. The working of the detection mechanism is shown in figure 1.

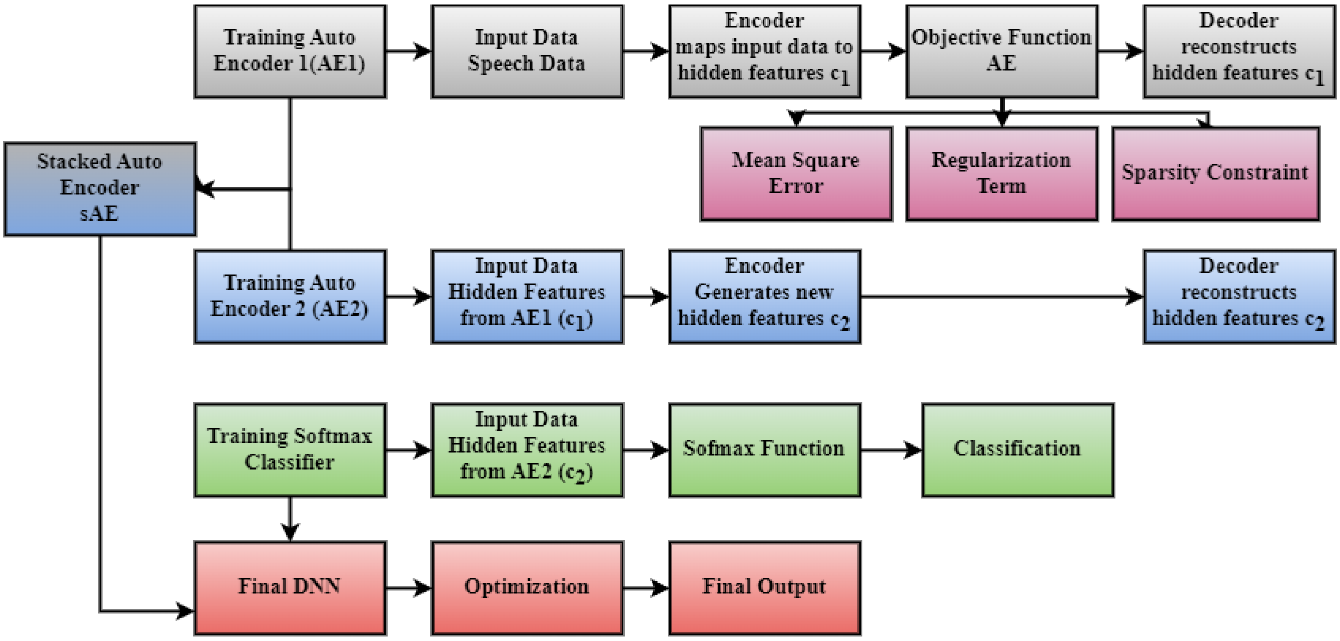

Figure 1. Working of Stacked Auto Encoder

By utilizing RA and ANN in concert to effectively identify PD. The methodology starts with regression analysis preprocessing the dataset to find correlations between independent variables such as iron content, pulse rate, DFA, RPDE, and PPE.(13) Following this, an ANN is fed regressed data to calculate a person's PD probability. By using the advantages of both approaches, the hybrid RA-ANN technique provides precise probability estimates from RA and refines them at the neurone level via threshold comparisons using ANN. With the use of SCILAB software and C programming, the concept was put into practice, and it produced an accuracy of 93,46 %, outperforming more conventional techniques like SVM & k-NN classifiers. The suggested approach shows improved performance by efficiently processing unprocessed data and using DL technologies to identify PD early on and the working is shown in figure 2.

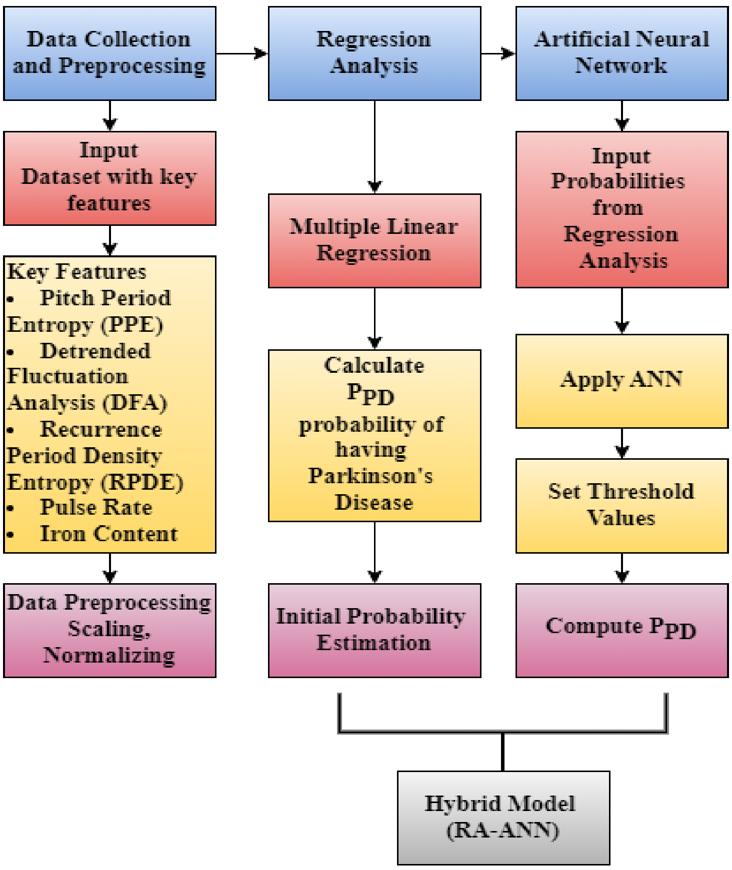

Figure 2. Hybrid ANN Model with Probability Rules

https://doi.org/10.56294/sctconf20251353

To use deep learning algorithms to identify PD from handwriting samples. Two main designs are CNN and CNN-BLSTM networks. Handwriting signals are transformed into 2D images, such as spectrograms and time series concatenations, by the CNN approach, which extracts features for classification. Conversely, the CNN-BLSTM model learns both temporal and spatial trends by directly processing raw time-series data, therefore enabling the network.[14] Two approaches were used to handle overfitting resulting from a small dataset: data augmentation methods include jittering, scaling, and time-warping; transfer learning, wherein models are pre-trained on comparable handwriting datasets and refined for the PD data. Emphasizing the need of temporal data management and successful augmentation methods, this approach shows the possibilities of DL in PD identification. The comparisons of the models are shown in figure 3.

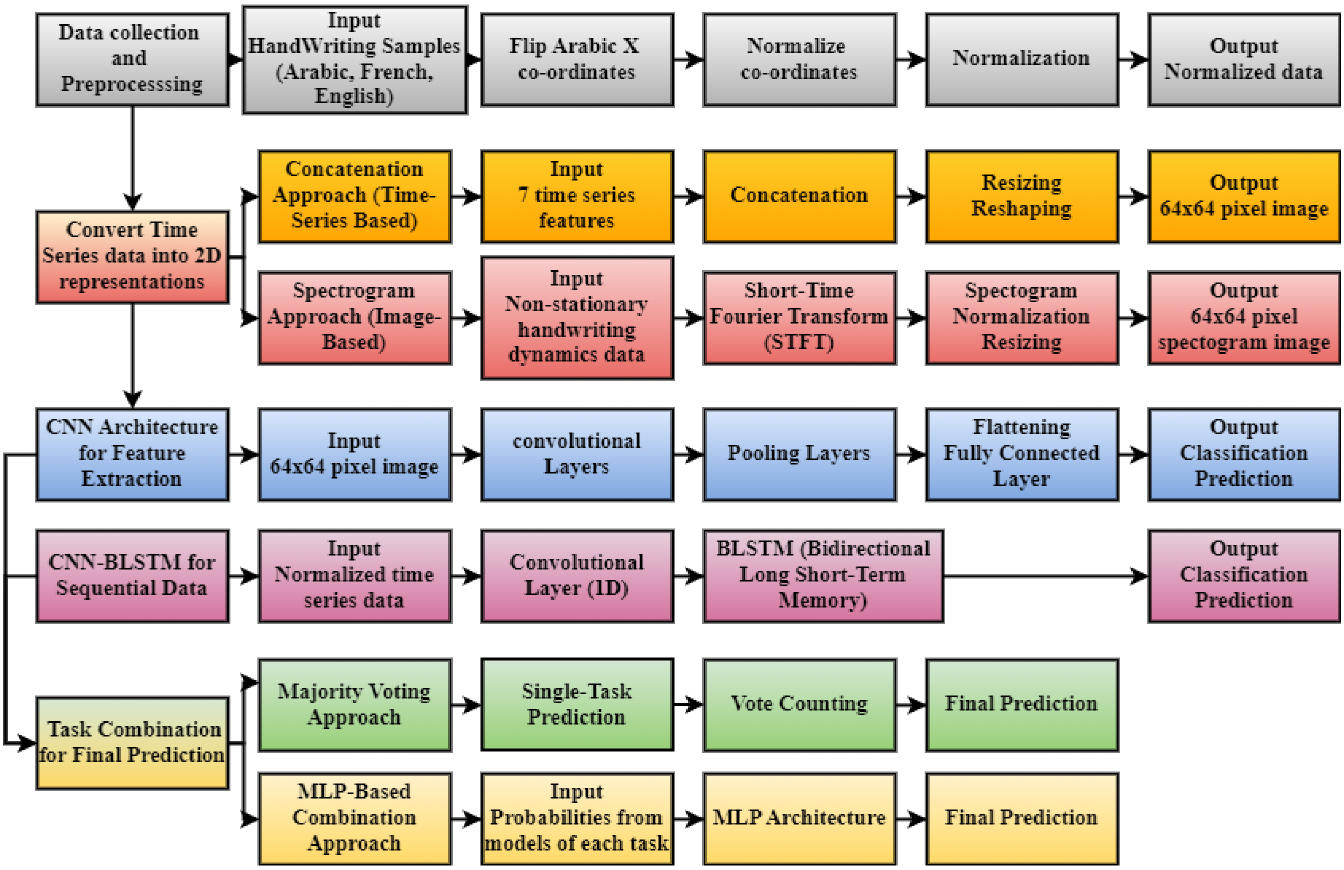

Figure 3. Comparative Analysis on Different CNN models

Detection using Vocal Features

DL-based systems for voice feature-based PD categorization. Two models grounded on CNN are presented. The first method, the feature-level mix, concatenates many voice feature sets then passes them through a 9-layer CNN for classifications and it is shown in figure 4. Layers of convolution and pooling as well as a fully linked layer for last decision-making comprise this. Using parallel convolutional layers to analyze many feature sets concurrently, the second method known as the model-level combination allows the framework to extract deep features from every input type concurrently.[15] Before being further handled for categorization, these retrieved features are consolidated in a single layer. A UCI ML repository dataset is used to test these frameworks' performance using LOPO CV. With several vocal characteristics including MFCC, TQWT, & wavelet-based features utilized for classification, the dataset comprises voice recordings from PD people and healthy persons. Furthermore used with accuracy for performance assessment are measures such as F-Measure & MCC because to the uneven character of the dataset.

Using speech recordings, detecting Parkinson's illness takes use of a deep learning method. Specifically, a modified ResNet architecture which is shown in figure 5 is originally intended for picture classification is supplied with the spectrogram images created by the approach when sustained vowel sounds are converted into them. Before fine-tuning with a PC-GITA, the network is first trained on the large ImageNet dataset. It is then trained again on a smaller speech dataset (SVD). Data augmentation methods include random rolling and band-pass filtering is used before creating the spectrograms to help to reduce overfitting resulting from the tiny sample size.[16] The last model combines dropout for regularizing a three-layer dense network. Emphasizing the efficiency of transfer learning from natural pictures to spectrogram-based speech features, this technique exhibits encouraging results with over 90 % accuracy in diagnosing PD from voice recordings.

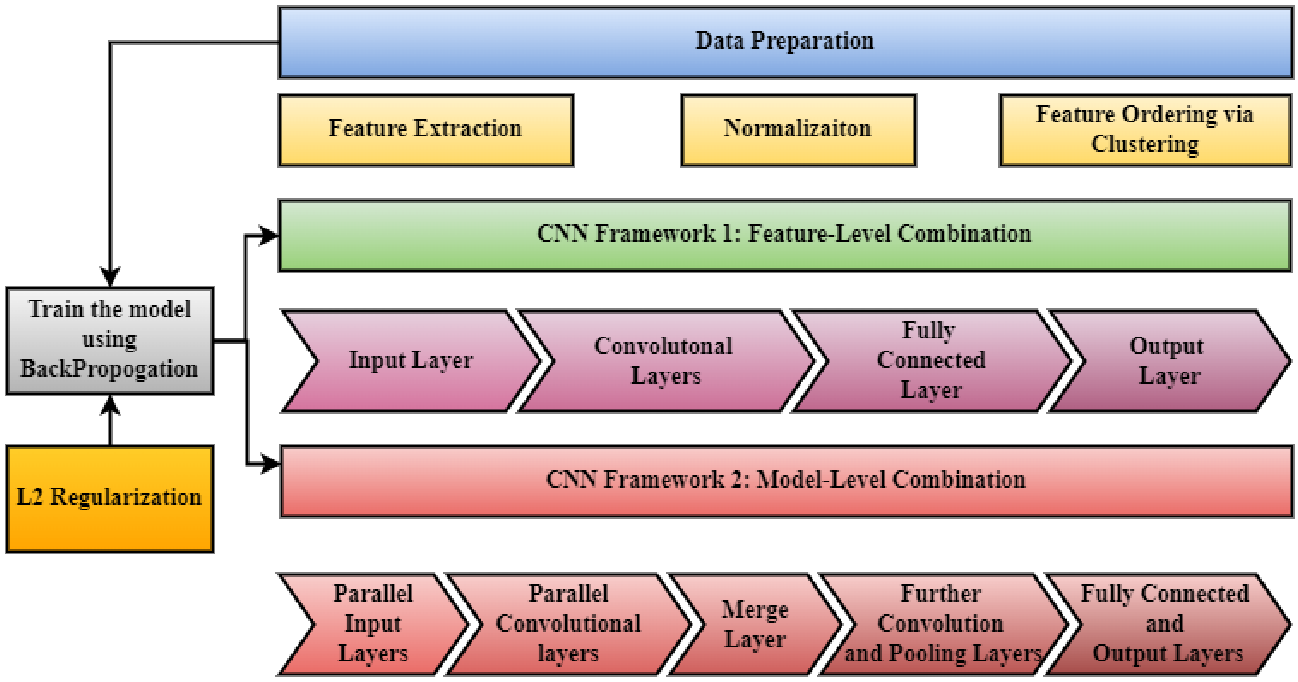

Figure 4. Layered CNN Framework

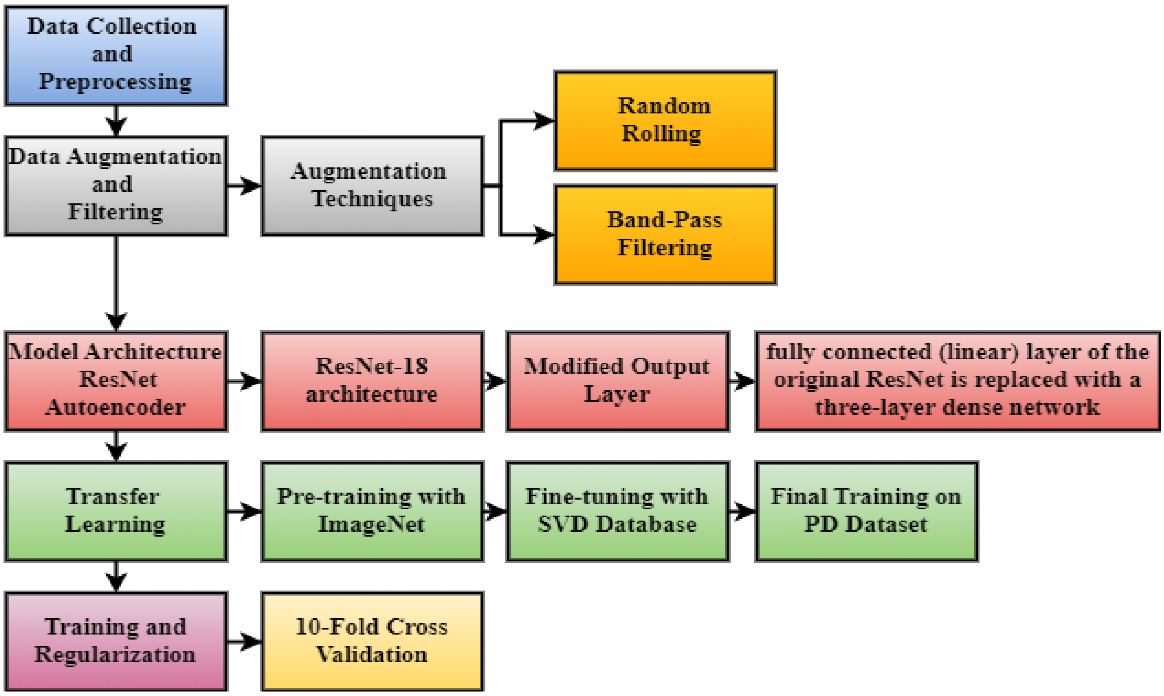

Figure 5. Modified Architecture of ResNet

Literature survey

A study by Hakan et al.[1] concentrated on classifying voice characteristics linked to a particular neurological disorder. The authors of the study put forward two deep learning-based frameworks. A feature-level combination is used in the initial framework, whereas a model-level combination is used in the second. To improve classification accuracy, a variety of speech recording-derived features are combined. The dataset for this study included recordings from 252 individuals (188 affected and 64 healthy), incorporating three voice recordings and four different feature types: Teager-Kaiser Energy Operator Wavelet Transform (TQWT), Wavelet, Mel-frequency Cepstral Coefficients (MFCC), and a concatenated feature set that includes time-frequency, baseline, and vocal fold characteristics. In the feature-level framework, different features are merged before being input into a nine-layered deep learning model. The model-level framework allows for the feeding of feature sets into parallel convolutional layers, enabling simultaneous extraction of deep features from each feature set.

Initial experiments using individual feature sets indicated that TQWT features yielded the best performance across various metrics, and combining TQWT with either MFCC or Wavelet features further improved accuracy and Matthews Correlation Coefficient (MCC). Binary combinations of features reached an accuracy of up to 0,849 using the model-level approach, while triple feature combinations achieved an accuracy of 0,869. This research demonstrates a notable enhancement in classification accuracy by leveraging deep learning frameworks, particularly through the feature-level and model-level combinations. The innovative introduction of parallel convolutional layers in the model-level framework allows for the simultaneous extraction of deep features from multiple feature sets, leading to improved classification outcomes. Nevertheless, the study's dependence on a rather small dataset (252 individuals) restricts the findings' generalizability to larger and more varied populations, casting doubt on the model's suitability in practical contexts. Additionally, while the study focuses solely on vocal data, the complexity of the neurological condition involves both motor and non-motor functions, which may restrict the model's overall efficacy.

Caliskan et al.[2] explored the diagnosis of a specific neurological condition through an examination of speech deficits, proposing a classifier based on deep learning techniques. The study utilized two datasets for training the model: one focused on various types of sound recordings and the other specifically aimed at detecting the condition. A 70/30 the training and evaluation partition and 10-fold cross-validation were used to evaluate the classifier's performance and compare it to cutting-edge methods like Support Vector Machine (SVM), Decision Tree (DT), and Naïve Bayes (NB) classifiers. Notably, especially within one of the datasets, the deep learning model performed better than the other classifiers in terms of accuracy and specificity. The study demonstrates the model's ability to uncover hidden patterns in the data, which is particularly useful for speech analysis-based early diagnosis and remote condition monitoring. However, despite its high accuracy and specificity, the deep learning model exhibited lower sensitivity compared to some alternative classifiers, particularly in the dataset focused on various sound recordings. This suggests that the model may overlook true positive cases, potentially impacting early detection efforts. Additionally, the study identified a lack of analytical parameter tuning in the methodology, which limits the generalizability of the model's performance. This oversight may hinder the model's adaptability to other datasets without extensive retraining and tuning, raising concerns about its applicability in diverse scenarios.

Sivaranjini et al.[3] proposed a classification method for a specific neurological condition using a transfer-learned deep learning architecture, specifically AlexNet, in conjunction with Magnetic Resonance Imaging (MRI). The study utilized a well-known database containing MRI scans from 182 participants, including 82 healthy individuals and 100 diagnosed with the condition. To enhance image quality and reduce noise, the authors applied Gaussian filtering and normalization processes, facilitating the identification of critical features for classification. The approach leverages a pre-trained model, traditionally applied to color photographs, to categorize MRI images into two classes: those affected by the condition and healthy controls. The final fully connected layer of the model was fine-tuned for this classification task. Through transfer learning, the early-layer weights were retrained using natural images, optimizing the later layers with the MRI data. 20 % of the dataset was put aside for testing, and the remaining 80 % was set aside for training. As evidenced by the ROC curve's Area Under the Curve (AUC) of 0,9618, the results showed a good discriminative capacity with an accuracy of 88,90 % and a specificity of 89,30 %. This method proved more robust than traditional machine learning approaches as it analyzes the entire brain, avoiding errors associated with region-specific segmentation. However, the study identified limitations, including the relatively small dataset, which restricts the model's generalizability. A more extensive and diverse dataset could potentially enhance the model's performance. Additionally, deep learning architectures like AlexNet are complex and demand significant computational resources, especially when applied to medical imaging tasks.

Zehra et al.[4] demonstrated a novel approach that applies deep learning algorithms to a voice dataset in order to predict the severity of a certain neurological disorder. The study placed a strong emphasis on voice data analysis, which can reveal information about how severe a patient's motor symptoms are. To do this, the researchers used TensorFlow to create a deep neural network that classified patients into "severe" and "non-severe" groups based on voice data. A standardized rating scale that gauges the severity of both motor symptoms and the overall state of the ailment served as the basis for determining the patient's severity. The deep neural network was designed with three layers that are hidden, a layer for input, and a layer for output with two neurons representing the output classes. The dataset included 16 biomedical voice characteristics and included 5,875 voice recordings from 42 patients. After normalization, the data was divided into training (80 %) and testing (20 %) datasets. Previous methods received accuracies around 47,2 % testing accuracy for total UPDRS and 44,3 % testing accuracy for motor UPDRS while the proposed model achieves 62,73 % testing and 94,44 % training accuracies for motor UPDRS and 81,67 % testing and 83,37 % training accuracies for motor UPDRS. By focusing on voice data, the study introduces a unique method of utilizing non-invasive indicators of PD severity, which could enhance patient monitoring and evaluation. The limiting dataset size limits the proposed model's generalizability which can be confirmed with less testing accuracy as the model may struggle with unseen data and concerning this approaches reliability.

M. Wodzinski et al.[5] explored a methodology to detect Parkinson's disease (PD) by analyzing sustained vowel sounds using a modified ResNet architecture. The ResNet is initially designed for image classification. The researchers transformed audio recordings into spectrograms(visual representations of the frequency spectrum over time) and passed these images as input to ResNet model. The dataset comprised recordings from 100 participants in which 50 were healthy people and 50 patients who were diagnosed with PD and each patient was recorded thrice. The approach proposed in this paper achieved a 91,7 % accuracy on the 10-fold validation set, highlighting the potential of using frequency-based features for PD detection. The results imply that natural image-based features can be successfully transferred to voice signal spectrogram representations. The model's accuracy is on par with the most advanced techniques for detecting Parkinson's disease. We can derive a conclusion that using frequency-based features alone is a promising diagnostic tool. The authors successfully applied data augmentation techniques. and helped the model to generalize better despite the small dataset size. But still, such promising approaches need a larger dataset and diverse examples to further increase the generalization. The model relies solely on frequency-based features, which, while effective, may overlook other important vocal characteristics that could improve diagnostic accuracy. A more thorough evaluation of PD might be possible with a wider range of features.

J. C. Vásquez-Correa et al.[6] introduced a multimodal approach of deep learning for identifying patients with a specific neurological condition by utilizing various biosignals, including speech, handwriting, and gait. The study investigated the effectiveness of these signals, both separately and in tandem, to improve diagnostic accuracy for the condition. The research emphasized the use of advanced neural network architectures, particularly focusing on the strengths of deep learning techniques in analyzing these diverse types of data. This innovative approach highlights the potential for a more comprehensive assessment by integrating multiple biosignals to enhance the detection capabilities for the neurological condition compared to healthy individuals. The authors explored both unimodal and multimodal combinations of biosignals and the best classification results were achieved by fusing all three modalities, achieving an accuracy of 97,6 %. The authors suggested to use of multimodal data and also described multimodal data collection (speech recordings, handwriting tasks, and gait analysis using accelerometers and gyroscopes). The features from these bio-signals were extracted using CNNs and focused on both onsets and offsets of data while extracting the features. Combining these modalities provided better insights into motor deficits caused by PD. Results show that combining all three bio-signals provides the best performance. The fusion method outperforms using each signal individually and exceeds baseline methods using traditional features. Speech and gait data contribute significantly to the improved classification. While the paper demonstrates high classification accuracy with multimodal data, to generalize this approach we must focus on the different languages for speech and different populations. Additionally, the model ought to be able to categorize the disease's severity.

Sahu et al.[7] investigated the use of deep learning methods to identify a certain neurological illness early, emphasizing the difficulties associated with symptom overlap with other disorders. Traditional diagnostic methods, such as MRI and SPECT imaging, while useful, often lack the necessary sensitivity for early identification of the condition. Previous research has utilized various machine learning algorithms, reporting classification accuracies ranging from 74 % to 90 %. However, these approaches have encountered challenges related to feature extraction and data preprocessing. The authors propose a hybrid method that integrates Regression Analysis (RA) with Artificial Neural Networks (ANN), aiming to overcome the limitations of individual models and improve overall classification accuracy. It utilizes a dataset with 741 values across multiple physiological parameters, achieving an impressive accuracy of 93,46 %. This underscores the potential of integrating diverse data types, including speech patterns and physiological metrics, to improve diagnostic capabilities. Despite its contributions, the paper has some limitations. First, the reliance on a single dataset may limit generalizability, as performance could vary with different populations or datasets. Second, the hybrid model's intricacy might make it difficult to apply in actual clinical settings, necessitating a large investment of time and knowledge.

Oh et al.[8] created a machine learning method that uses electroencephalogram (EEG) data and a thirteen-layer deep learning architecture to diagnose a particular neurological illness. Because of its mild early symptoms, this disorder, which is defined by impairment of motor function due to neuronal death, sometimes goes undiagnosed until later stages. To find irregularities in brain activity, the researchers collected EEG data from twenty individuals who had the condition and twenty healthy controls. With a sensitivity of 84,71 %, a specificity of 91,77 %, and an overall accuracy of 88,25 %, the deep learning model demonstrated impressive performance characteristics. This approach eliminates the need for the extraction of features manually, increasing the diagnostic process's speed and efficiency and offering a viable route to more accurate and timely condition identification. The model's dependability is increased by the study's use of a tenfold cross-validation technique. There are some significant restrictions, though. The study's small sample size may not be representative of the general population, to start. Second, the complex CNN structure requires significant computational resources, making it less accessible for widespread clinical use. Future research aims to expand the dataset and explore applications for other neurological disorders.

In their study, Catherine Taleb et al.[9] use CNN-BLSTM model, a deep learning approachto diagnose Parkinson's disease (PD) by examining handwriting. Their research expands on earlier studies that used common machine learning techniques, like Support Vector Machines (SVM), and demonstrated accuracy ranging from 89,09 % to 96,87 %. By employing both majority voting and a Multi-Layer Perceptron (MLP) for model output combination, the authors achieve a notable accuracy of 97,62 %. They emphasize the importance of data augmentation techniques, like jittering and synthetic data creation, in enhancing model performance. Additionally, the inclusion of diverse features, such as pressure and altitude, significantly improves classification outcomes, highlighting their relevance to neuro-motor dysfunction in PD patients. However, the study faces limitations; the reliance on specific datasets raises concerns about the findings' applicability to larger populations and clinical contexts. Moreover, the complexity of the CNN-BLSTM architecture may hinder its real-time applicability and increase computational costs, potentially restricting its deployment in resource-constrained environments. These factors underline the need for further research to validate the model across various datasets and optimize its performance for practical use.

Aşuroğlu et al.[10] proposed an innovative hybrid deep learning framework that combines convolutional neural networks with local weighted random forest for analyzing gait signals to predict the severity of a specific neurological condition. The study utilized ground reaction force measurements obtained from both healthy individuals and patients diagnosed with the condition, sourced from a well-known gait dataset. The feature extraction process involved analyzing the ground reaction force signals, incorporating both time-domain and frequency-domain characteristics. These features are essential for distinguishing walking patterns that correspond to varying levels of motor symptom severity, thereby enhancing the accuracy of the predictions made by the framework. The authors emphasize the importance of constructing a feature vector from raw GRF data, as direct input poses challenges due to varying signal lengths and temporal characteristics. Their hybrid model effectively captures local relationships between extracted features, demonstrating improved prediction accuracy over traditional models. Results show that the proposed framework achieved a Correlation Coefficient (CC) of 0,897, with a Mean Absolute Error (MAE) of 3,009, outpacing baseline models like Linear Regression and Support Vector Regression. While the study contributes significantly to gait analysis in PD, it has some limitations. First off, the results' applicability to larger populations may be limited by their dependence on a particular dataset. Second, the hybrid model's complexity might make it difficult to scale and use in real-time, which could prevent doctors from using it. Overall, Aşuroğlu et al.'s work represents a meaningful advancement in leveraging deep learning for PD monitoring, though further validation across diverse datasets is recommended.

Materials used

Depending on the kind of data being investigated, many datasets are accessible for PD study and detection; each has a special use. The speech patterns of healthy individuals and PD patients are recorded in the PDSD, which is one of the most frequently employed datasets. This dataset is useful for examining slurred speech, voice tremors, and other PD-related vocal abnormalities.[17] It has 23 speech measurements from 31 people, 23 of whom have Parkinson's disease. The UCI Parkinson's collection is another one that is often used. The PPMI dataset provides clinical evaluations, imaging data, and biospecimens for motor function data. One of the most complete datasets available, PPMI tracks the development of PD in identified individuals across time. Other datasets include wearable sensor datasets tracking motor activities in real time and gait datasets capturing patient movement patterns. Important in PD diagnosis, brain pictures found in medical imaging databases such as the Parkinson's Disease DaTscan database help to investigate dopamine transporter levels.

Table 1. Description about the Datasets

S. No.	Dataset Name	No. of Attributes	Data Type	Class Labels
1	Parkinson's Disease Speech Dataset	26 voice features	Speech Data	PD, Non-PD
2	UCI Parkinson Dataset	23 voice features	Speech Data	PD, Non-PD
3	Parkinson Progression Markers	Varies	Imaging, Clinical	Progressive PD, Stable PD, control
4	Gait Dataset	Various sensor attributes	Gait Data	PD, Non-PD
5	Wearable Sensor Dataset	6 axis motion data	Sensor Data	PD, Non-PD
6	Parkinson's Disease DaTscan	2D or 3D brain images	Medical Images	PD, Non-PD

There are different kinds of speech clips in the PDSD, such as continuous phonation, words, and unscripted speech. Its main purpose is to study vocal signs of Parkinson's disease, since people with the disease often have problems with pronunciation, speaking in a monotone, and making sounds.[18] There are 195 speech samples from 31 people in the UCI Parkinson's collection. These recordings have traits like jitter, shimmer, and HNR that can show vocal fold irregularities connected to Parkinson's disease. One of the biggest datasets out there is the PPMI dataset, which has a lot of different types of data, such as MRI and DaTscan pictures, study of cerebral fluid, and clinical reports. The main goal of PPMI is to find signs for Parkinson's disease development, which makes it very useful for CNN uses. The gait datasets use sensor-based devices to record how patients walk. This information is useful for finding motor signs like bradykinesia, stiffness, and tremors. Approximator and gyroscope data from smartwatches may be used to monitor PD patients' motor symptoms. Different datasets each have their own use and offer different traits that researchers can use in various PD detection models.

DISCUSSION

The following research gaps are identified while analyzing the existing systems:

1. Despite the complexity of Parkinson's disease, which affects both motor and non-motor processes, the study exclusively used vocal (speech) data for categorization. This method might miss significant illness characteristics that are usually represented by motor-related signs, such tremors, rigidity, or problems with walking.
2. Scalability and real-time application issues may arise due to the hybrid deep learning framework's relative complexity, which combines CNN and LWRF. When using the model in clinical settings where prompt, real-time predictions are required, its computational requirements may prove to be an obstacle.
3. A more comprehensive model for PD severity prediction could be developed in the future by integrating multimodal data (such as speech analysis, handwriting samples, tremor data, and brain imaging). Comprehensive and reliable assessments of illness development may be possible using multimodal deep learning frameworks that integrate data from different sources.

Table 2. Existing System Analysis

Author(s)	Algorithm	Merits	Demerits	Accuracy
Hakan et al	CNN (Feature-level & Model-level combination)	Innovative use of parallel convolution layers for simultaneous feature extraction, improved accuracy with TQWT	Small dataset (252 individuals) limits generalization.	Up to 0,869 with triple feature combination
Caliskan et al.	DNN	Superior to SVM, DT, and NB classifiers in accuracy and specificity, effective for remote PD monitoring.	Lower sensitivity, lack of analytical parameter tuning reduces adaptability.	High accuracy, specific to OPD dataset.
Sahu L et al.	Hybrid ANN and Regression Analysis	Transfer learning avoids region-specific segmentation errors, high AUC (0,9618)	Small MRI dataset (182 subjects), complex model requires more computational power.	88,90 % accuracy, 89,30 % specificity.
Taleb et al.	CNN-BLSTM with handwriting analysis	Non-invasive voice analysis for PD severity, higher accuracies compared to prior methods	Limited dataset size (42 patients), overfitting indicated by lower testing accuracy.	94,44 % training, 62,73 % testing accuracy (Motor UPDRS).
Sivaranjini et al.	Transfer-learned CNN (AlexNet)	Effective use of spectrograms, high accuracy with frequency-based features, data augmentation improves generalization	Small sample size (100 participants), overlooks other vocal features.	91,7 % accuracy.
Zehra et al.	DNN with Tensorflow	Overcomes individual model limitations, high accuracy across multiple physiological parameters	Limited generalization due to single dataset, high computational complexity.	93,46 % accuracy.
Skalski et al.	Modified ResNet	High accuracy (97,62 %) with MLP and majority voting, effective data augmentation improves model performance	Dataset reliance limits generalization, complexity of model affects real-time applicability.	97,62 % accuracy.
Aşuroğlu et al.	Hybrid CNN and Local Weighted Random Forest (LWRF)	Effective for gait analysis, captures local relationships in GRF data, outperforms baseline models.	Limited dataset may reduce applicability, high computational complexity affects real-time deployment.	Correlation Coefficient of 0,897, MAE of 3,009.

CONCLUSIONS

One new area to improve early identification and tracking of PD is the use of machine learning methods for spotting. These models learn from trends in past data and then use that information to put new data into one of two groups: PD data and non-PD data. Wearable tech with sensors can collect data on patients all the time, so machine learning models can see how the disease is getting worse in real time. There is also speech analysis that uses ML. Algorithms can find small changes in voice rhythms that are often early signs of PD. The good thing about machine learning for finding PD is that it can work with big, varied datasets and, once learnt, can do tests more quickly and correctly than old-fashioned ways. Due to their capacity to get around data restrictions, shorten training times, and enhance performance on tiny datasets, pre-trained models are very helpful for detecting Parkinson's disease. They present a more expedient and effective path toward the development of precise, real-time detection systems that may be implemented in everyday and clinical environments. Researchers and physicians can expedite the development of AI-based Parkinson's disease (PD) detection systems by utilizing transfer learning and well-proven architectures. This makes these models an invaluable asset in the battle against neurodegenerative disorders.

BIBLIOGRAPHIC REFERENCES

1. Gunduz, Hakan. "Deep learning-based Parkinson's disease classification using vocal feature sets." Ieee access 7 (2019): 115540-115551.

2. Caliskan, Abdullah, et al. "Diagnosis of the Parkinson disease by using deep neural network classifier." IU-Journal of Electrical & Electronics Engineering 17.2 (2017): 3311-3318.

3. Sivaranjini, S., Sujatha, C.M. Deep learning based diagnosis of Parkinson's disease using convolutional neural network. Multimed Tools Appl 79, 15467–15479 (2020). https://doi.org/10.1007/s11042-019-7469-8

4. Senturk, Zehra Karapinar. "Early diagnosis of Parkinson's disease using machine learning algorithms." Medical hypotheses 138 (2020): 109603.

5. M. Wodzinski, A. Skalski, D. Hemmerling, J. R. Orozco-Arroyave and E. Nöth, "Deep Learning Approach to Parkinson's Disease Detection Using Voice Recordings and Convolutional Neural Network Dedicated to Image Classification," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 2019, pp. 717-720, doi: 10.1109/EMBC.2019.8856972.

6. J. C. Vásquez-Correa, T. Arias-Vergara, J. R. Orozco-Arroyave, B. Eskofier, J. Klucken and E. Nöth, "Multimodal Assessment of Parkinson's Disease: A Deep Learning Approach," in IEEE Journal of Biomedical and Health Informatics, vol. 23, no. 4, pp. 1618-1630, July 2019, doi: 10.1109/JBHI.2018.2866873.

7. Sahu, L., Sharma, R., Sahu, I., Das, M., Sahu, B., & Kumar, R. (2022). Efficient detection of Parkinson's disease using deep learning techniques over medical data. Expert Systems, 39(3), e12787. https://doi.org/10.1111/exsy.12787

8. Oh, S.L., Hagiwara, Y., Raghavendra, U. et al. A deep learning approach for Parkinson's disease diagnosis from EEG signals. Neural Comput & Applic 32, 10927–10933 (2020). https://doi.org/10.1007/s00521-018-3689-5

9. 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).

10. Aşuroğlu, T., Oğul, H. A deep learning approach for parkinson's disease severity assessment. Health Technol. 12, 943–953 (2022)

11. Yadav, S., Singh, M.K. & Pal, S. Artificial Intelligence Model for Parkinson Disease Detection Using Machine Learning Algorithms. Biomedical Materials & Devices 1, 899–911 (2023). https://doi.org/10.1007/s44174-023-00068-x

12. Agrawal, S., Sahu, S.P. Image-based Parkinson disease detection using deep transfer learning and optimization algorithm. Int. j. inf. tecnol. 16, 871–879 (2024). https://doi.org/10.1007/s41870-023-01601-3

13. Islam, Md. A., Hasan Majumder, Md. Z., Hussein, Md. A., Hossain, K. M., & Miah, Md. S. (2024). A review of machine learning and deep learning algorithms for Parkinson's disease detection using handwriting and voice datasets. In Heliyon (Vol. 10, Issue 3, p. e25469). Elsevier BV. https://doi.org/10.1016/j.heliyon.2024.e25469

14. Erdaş ÇB, Sümer E. 2023. A fully automated approach involving neuroimaging and deep learning for Parkinson's disease detection and severity prediction. PeerJ Computer Science 9:e1485 https://doi.org/10.7717/peerj-cs.1485

15. Khaskhoussy, R., & Ayed, Y. B. (2023). Improving Parkinson's disease recognition through voice analysis using deep learning. In Pattern Recognition Letters (Vol. 168, pp. 64–70). Elsevier BV. https://doi.org/10.1016/j.patrec.2023.03.011

16. Nilashi, M., Abumalloh, R. A., Yusuf, S. Y. M., Thi, H. H., Alsulami, M., Abosaq, H., Alyami, S., & Alghamdi, A. (2023). Early diagnosis of Parkinson's disease: A combined method using deep learning and neuro-fuzzy techniques. In Computational Biology and Chemistry (Vol. 102, p. 107788). Elsevier BV. https://doi.org/10.1016/j.compbiolchem.2022.107788

17. Md Abu Sayed, Duc Minh Cao, Islam, M. T., Tayaba, M., Md Eyasin Ul Islam Pavel, Md Tuhin Mia, Eftekhar Hossain Ayon, Nur Nobe, Bishnu Padh Ghosh, & Sarkar, M. (2023). Parkinson's Disease Detection through Vocal Biomarkers and Advanced Machine Learning Algorithms. Journal of Computer Science and Technology Studies, 5(4), 142–149. https://doi.org/10.32996/jcsts.2023.5.4.14

18. Balnarsaiah, B., Nayak, B. A., Sujeetha, G. S., Babu, B. S., & Vallabhaneni, R. B. (2023). RETRACTED ARTICLE: Parkinson's disease detection using modified ResNeXt deep learning model from brain MRI images. In Soft Computing (Vol. 27, Issue 16, pp. 11905–11914). Springer Science and Business Media LLC. https://doi.org/10.1007/s00500-023-08535-9

19. Saleh, S., Cherradi, B., El Gannour, O. et al. Predicting patients with Parkinson's disease using Machine Learning and ensemble voting technique. Multimed Tools Appl 83, 33207–33234 (2024). https://doi.org/10.1007/s11042-023-16881-x

20. Rahman, S., Hasan, M., Sarkar, A.K. and Khan, F. 2023. Classification of Parkinson's Disease using Speech Signal with Machine Learning and Deep Learning Approaches. European Journal of Electrical Engineering and Computer Science. 7, 2 (Mar. 2023), 20–27. DOI: https://doi.org/10.24018/ejece.2023.7.2.488.

FINANCING
No financing.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest.

AUTHORSHIP CONTRIBUTION
Data curation: Sowjanya Vuddanti, A. Srinagesh.
Methodology: Sowjanya Vuddanti, A. Srinagesh.
Software: Sowjanya Vuddanti, A. Srinagesh.
Drafting - original draft: Sowjanya Vuddanti, A. Srinagesh.
Writing - proofreading and editing: Sowjanya Vuddanti, A. Srinagesh.