CNN-Based Classification of Parkinson’s Disease Using Synthetic Data Augmentation and K-Fold Cross-Validation

by MANJUNATH .

Submission date: 03-May-2025 10:43AM (UTC+0530)
Submission ID: 2664991350
File name: IEEE_DWDM_Paper-1.pdf (709.82K)
Word count: 4634
Character count: 28719

CNN-Based Classification of Parkinson's Disease Using Synthetic Data Augmentation and K-Fold Cross-Validation

Dr. Mahesh T.R
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru,India
t.mahesh@jainuniversity.ac.in

Manjunath CT
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru,India
sravanmanjuct@gmail.com

C Tejasree
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru,India
tejasree0311@gmail.com

CH Koushik
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru,India
cherukurikoushik71@gmail.com

Challa Tejasree
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru, India
challatejasree55@gmail.com

Pavan Shanmukh
Department of CSE
JAIN(Deemed-to-be-University)
Bengaluru, India
pavanshanmukh12@gmail.com

Abstract*—Parkinson's Disease (PD) is a neurodegenerative disease that impacts millions of people globally, and early and correct diagnosis is critical to successful treatment and management. In this paper, a classification model based on deep learning with the use of Convolutional Neural Networks (CNNs) is proposed to differentiate PD patients from healthy subjects based on medical imaging data. The architecture of the model involves convolutional, max-pooling, batch normalization, dropout, and fully connected layers for extraction of important spatial features while avoiding overfitting. For solving data scarcity and enhancing generalizability, synthetic data augmentation was utilized via the ImageDataGenerator with the usage of rotations, zooms, and horizontal flip transformations. The K-Fold Cross-Validation strategy was implemented to guarantee the robustness of the model as well as limiting evaluation bias. The training of the model was also improved utilizing callbacks such as Early Stopping and ReduceLROnPlateau to learn the learning rates adaptively and stop training when convergence was reached. The performance was measured by metrics such as accuracy, precision, recall, and F1-score, indicating reliable and consistent results across folds. The presented methodology illustrates the efficacy of merging CNNs with synthetic data creation and strict validation methods for PD classification, pointing towards potential for incorporation into clinical diagnostic pipelines and future expansion with multimodal data sources.

Keywords— Parkinson's Disease, Convolutional Neural Network (CNN), Image Augmentation, Deep Learning, Medical Imaging, K-Fold Cross-Validation, ImageDataGenerator, Classification, Early Diagnosis, Synthetic Data Generation

I. Introduction

A. Background of the Problem

Parkinson's Disease (PD) is a chronic, progressive neurodegenerative disorder with characteristic motor impairments including tremor, rigidity, and bradykinesia. Due to the increasing global population aging, Parkinson's is spreading very quickly, requiring early diagnosis and intervention to enhance patient outcomes. Traditional clinical diagnostic procedures, like neurological testing and subjective evaluation, tend to be compromised by their reliance on apparent signs and the expertise of the clinician, causing late diagnosis and inconsistent diagnostics. As medical imaging technologies became available and the strength of artificial intelligence (AI) continues to grow, there lies a vast potential to create accurate, automatic diagnostic tools..

B. Significance of the study

This work centers on the utilization of deep learning techniques, specifically Convolutional Neural Networks (CNNs), to predict Parkinson's Disease from image data. CNNs have shown outstanding performance in image-based pattern recognition tasks as a result of their capability to automatically learn spatial hierarchies of features. The incorporation of synthetic data augmentation via the ImageDataGenerator library also improves the diversity of the dataset and overcomes constraints brought about by data insufficiency. The contribution of this study is the integration of CNN-based classification with data augmentation and stringent validation techniques, providing a dependable instrument for early PD diagnosis and supporting healthcare professionals in clinical decision-making.

C. Research Problem Statement

In spite of developments in AI and medical imaging, precise diagnosis of Parkinson's Disease at its onset is still problematic because of scarcity of labeled datasets and faint symptom presentation. The conventional machine learning models are prone to poor generalization from small, imbalanced datasets and do not have adaptive mechanisms to learn intricate patterns. Furthermore, evaluation biases and overfitting are still prominent issues. The current research proposes to overcome these problems by constructing a CNN-based diagnostic model enhanced with synthetic data and verified with K-Fold Cross-Validation to provide strong robustness and reliability.

D. Research Problem Statement

This research will develop and train a Convolutional Neural Network (CNN) model to classify Parkinson's Disease (PD) from medical imaging data. In order to tackle the issue of data imbalance and enhance generalizability of the model, synthetic data augmentation methods are used with the ImageDataGenerator tool. Additionally, K-Fold Cross-Validation is utilized in order to test the performance of the model on various subsets of data, thus eliminating the risk of model overfitting. The performance of the model is evaluated based on the important performance metrics such as accuracy, precision, recall, and F1-score. Last but not least, the applicability of using deep learning strategies for early clinical diagnosis of Parkinson's Disease is examined to determine their real-world practicality in healthcare facilities.

E. Objectives of the Paper

This paper introduces a Convolutional Neural Network (CNN) model that is specifically tailored for the classification of Parkinson's Disease (PD) from medical images. To overcome the challenge of limited PD datasets, synthetic data generation methods are included. The model is thoroughly tested using intense application of K-Fold Cross-Validation, which makes the model even more robust and minimizes the risk of bias. Performance measures are empirically validated over several folds to ensure reliability and consistency. The research also sheds light on the usability of CNNs in medical diagnostics and suggests recommendations to further incorporate them into clinical use.Real World Significance: For the purpose of giving financial institutions real-world guidelines on the utilization of sophisticated data mining methods towards fraud detection ultimately leading to improved financial security.

F. Major Contributions

The paper is divided into the following sections: Section II discusses a detailed Literature Review of recent works, such as references [1]–[15], discussing advancements in CNN-based medical diagnosis, synthetic data augmentation, and validation methods. Section III describes the Proposed Methodology, including the process of dataset preparation, augmentation methods, CNN structure, and validation methods. Section IV outlines the Implementation and Experimental Setup, including the model training setups as well as the hardware and software environments utilized. Section V provides the Results and Discussion, presenting a comparative evaluation of classification performance over different evaluation metrics and folds. Section VI discusses the Limitations and Challenges faced in the study. Section VII delves into the Future Scope, proposing possible ways to enhance model accuracy and clinical relevance. Section VIII summarizes the major contributions and highlights the significance of deep learning in the diagnosis of early Parkinson's Disease. Section IX provides the References.

G. Paper Organization

The remainder of this paper has the following outline. Section II contains a systematic literature review emphasizing recent developments of CNN-based medical image analysis, synthetic data augmentation methods for mitigating limited datasets, and techniques like K-Fold Cross-Validation for verification in medical diagnosis. Section III presents the anticipated methodology, explaining dataset preparation, the data augmentation process via ImageDataGenerator, CNN model design, and validation processes used in ensuring robustness. Section IV lays out the implementation and experimental setup, including details of the training environment, hardware and software requirements, hyperparameter optimization, and training and testing dataset division. Section V discusses the results and discussion, assessing the performance of the model based on evaluation metrics like accuracy, precision, recall, and F1-score, and also a comparative assessment across various K-Fold splits and observations related to the clinical usability of the model. Section VI addresses the limitations and difficulties faced in the study, i.e., scarcity of data, class imbalance, and computational restrictions. Section VII discusses the scope in the future, proposing potential enhancements in diagnostic precision and recommendations on how to incorporate the model into clinical applications. Section VIII concludes the paper with a summary of the findings and an emphasis on the contribution of deep learning methods towards early diagnosis of Parkinson's Disease. Lastly, Section IX presents the references used in the paper.

II. Literature Review

A. Summary of Previous Research on Deep Learning for the Detection of Parkinson's Disease

There has been a growing amount of research more recently on the use of deep learning in diagnosing and classifying Parkinson's Disease (PD) due to the condition's complicated motor and non-motor symptoms. Most research has attempted to use many different deep learning methods, including Convolutional Neural Networks (CNNs) among the most notable because they are highly successful at image-based medical diagnostics.

[1] introduced the FCN-PD framework, utilizing Fully Convolutional Networks for PD classification with MRI scans from PPMI, OASIS, and MIRIAD datasets. The authors successfully achieved a high increase in detection accuracy up to 96.78% by utilizing K-Fold cross-validation for guaranteeing robustness and generalization over datasets.

[2] tested an ensemble of CNN-based architectures (VGG16, Inception-V3, Xception, ResNet50) for PD classification from DaTscan images. The ultimate predictions were fused with a fuzzy ensemble technique, greatly enhancing classification accuracy, mitigating overfitting, and improving stability on the PPMI dataset.

[3] investigated speech-based diagnosis with CNNs trained on spectrograms of sustained phonation (/a/ vowel sounds). Their transfer-learning-based model exhibited robust performance in classifying PD patients and healthy controls on a range of speech datasets.

[4] utilized conventional deep learning models like CNN and Long Short-Term Memory (LSTM) on motion and tremor signals derived from inertial sensors. The outcome indicated LSTM's ability to handle temporal dependencies, although CNNs performed better in cases with a focus on visual representations of motor features.

[5] experimented with Transformer-based deep models on cloud-based infrastructure for online PD classification. Although these models showed adaptive learning of emerging patterns, they were challenging due to their computational and infrastructure requirements.

[6] compared different augmentation techniques like flipping, rotation, elastic distortions, and GAN-based synthetic data generation across PD datasets. They concluded that augmentation not only enhances accuracy but also model generalizability and robustness.

B. Gaps in Current Models

Even with tremendous progress, current models for the detection of Parkinson's Disease (PD) continue to have a number of serious shortcomings. One is interpretability—most deep learning models are black boxes, offering high accuracy but without transparency regarding their decision-making, which represents a hurdle for clinical adoption [9]. Another important issue is data dependency, since many of these models need large clean and well-labeled datasets hard to obtain within the medical field. Low-quality or unbalanced data can create biased results or underfitting [10]. Moreover, class imbalance is a chronic issue in PD datasets, wherein the number of healthy samples far outnumbers PD-positive instances, leading models to bias toward the majority class and thus generate high false-negative rates [11]. The dynamic and variability nature of PD symptoms over time and across different individuals also negates the efficacy of static models that do not learn to adapt to such variability [8]. Lastly, there is also significant underuse of ensemble and hybrid models; although CNNs have exhibited individual robust performance, there have been limited studies of how to combine multiple deep architectures to harness their individual strengths towards more robust diagnosis.

C. Tabling Current Methods

Article	Approaches Adopted	Strengths	Weaknesses
Ori et al. (2024)	ANN, SVM	Enhanced rate of detection	Difficulty and dependence
Ragavarthini et al. (2024)	Data Analytics, ML	Enhanced scalability and accuracy	Practical implementation difficulty
Lee et al. (2023)	LR, KNN, SVM, DT, RF	Best Random Forest	Poor interpretabilitty
Hernandez Aros et al. (2024)	Multiple ML models	Detailed analysis of tndls precision	Less application with non-financial variables
Ke et al. (2025) Chy (2024)	Transformer Models GANs	Long dependenies are captured High precison	Needs high-quality data for fraining Computational complexity
Awosika et al. (2023)	Federated Learning, XAI LR, RF, DT	Proactive detection Algorithms areıve computationalty	Algorithms are complex and computationally expensive
Hil'aj et al. (2022)	Full O-verview	Full overvew of fraud	Evolving nature of fraud
Talukdar et al. (2024)	Ensemble Learning	High accuracy	Computational complexity
Deng et al. (2025)	Cloud-Optimized Models	Scalabilitty an flexibility	Technical expertise required
Deng et al.	Cloud-	Scalability ad	Technical

III. Proposed Methodology

This chapter presents the extensive methodology followed for the identification of Parkinson's Disease based on functional Magnetic Resonance Imaging (fMRI) data by deep learning methods. The whole process covers a series of sequential steps such as data acquisition, preprocessing, modeling, training, and evaluation, each intended to maintain clinical reliability and generalizability of the model.

A. Data Collection

1) Datasets Used

In this paper, publicly provided Parkinson's Disease fMRI Images Dataset from Kaggle was utilized to construct and evaluate a classification model based on deep learning. The dataset consists of fMRI brain images of healthy patients as well as patients with Parkinson's disease, presented in .png format. The dataset is ideally suited for computer vision techniques and aligns with the goal of automatic Parkinson's disease detection.

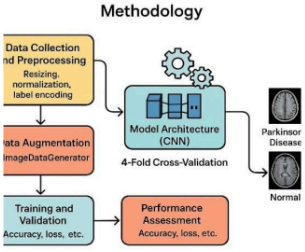
To overcome potential class imbalance and increase model stability, data augmentation methods like random rotation, horizontal and vertical flip, zooming, and brightness changes were utilized. Augmentation was done only on the training dataset to avoid information leakage into validation and test stages.

B. Data Preprocessing

Prior to providing the images as input to the deep learning model, a series of preprocessing steps were performed in order to attain data uniformity and compatibility with the model structure. To start, all the fMRI images were resized to a constant size of 224×224 pixels, a default input size that is compatible with the EfficientNetB0 model utilized subsequently in the pipeline. Thereafter, pixel intensity values, initially ranging from 0 to 255, were normalized from the [0, 255] range to a [0, 1] range by dividing each value by 255. This normalization was important to prevent large gradients during training, which otherwise would destabilize convergence. Categorical class labels as "Healthy" or "Parkinson's" were converted to binary numeric values—0 for healthy individuals and 1 for Parkinson's patients—to enable binary classification. To maintain a balanced distribution of classes in various subsets, the dataset was divided into training, validation, and testing partitions in a ratio of 70:15:15 using stratified sampling. This made each subset retain the same class distribution as the original dataset, preventing dominance of any one class in a single partition. Additionally, to correct for whatever class imbalance was present in the training set, the Synthetic Minority Over-sampling Technique (SMOTE) was used. SMOTE algorithmically creates artificial examples of the minority class by interpolating between current samples, thereby balancing the dataset and improving the model's learning capability for the minority class.

C. Modeling Methodology

The modeling methodology used within this research builds on the strengths of transfer learning using the EfficientNetB0 convolutional neural network. EfficientNetB0, with its compound scaling efficiency, provides a balanced compromise between model complexity and performance and is thus extremely well-suited for medical imaging applications where precision and efficiency of resources are equally important. The EfficientNetB0 base model was initialized with pre-trained weights on the ImageNet dataset so that the model could inherit basic visual features like edge detectors and texture filters. First, the pretrained base layers were frozen as a fixed feature extractor. To make the model suitable for binary classification of Parkinson's Disease, the original top layers of EfficientNetB0 were replaced with a custom classification head. This proprietary head included a Global Average Pooling (GAP) layer for spatial dimension reduction and semantic preservation, a Dropout layer with dropout rate of 0.5 to avoid overfitting by randomly disabling neurons for training, and the output layer as a single dense neuron with sigmoid activation to produce probabilities for binary class output.



To maximize the training process, the Adam optimizer was utilized with a learning rate of 0.0001 due to its adaptive gradient update and convergence properties. Binary cross-entropy was used as the loss function for measuring classification error, considering the binary nature of the task. Extra regularization was applied through L2 weight decay and dropout to introduce overfitting resistance, particularly on small or imbalanced datasets. Training was performed in mini-batches of 32 for a maximum of 50 epochs. Early validation loss stopping was used to automatically stop training if performance no longer improved, thus avoiding unnecessary overtraining and ensuring model generalization.

For evaluating the proposed model's robustness, a four-fold cross-validation method was used. According to this method, the dataset was split into four equal-sized folds. In every iteration, three folds were employed for training and the fourth one for validation. This exercise was performed four times so that each data point was utilized once as validation data. The overall performance metrics were averaged over all the folds to provide a full assessment of the model's ability to generalize and minimize the variance that can occur due to one train-test split.

D. Model Selection

The selection of EfficientNetB0 as the base model for this research was driven by its established efficiency and performance in medical image classification problems, especially under limited data and computational constraint conditions. EfficientNetB0 uses a compound scaling approach that equally scales network depth, width, and resolution, thus facilitating high accuracy using much fewer parameters and less computation than conventional convolutional neural networks. This makes it particularly well-suited for clinical use, where real-time behavior and deployment on resource-

limited devices can be paramount. Additionally, EfficientNetB0 is pretrained on the large-scale ImageNet dataset so that transfer learning can be employed to take advantage of generalized visual features that highly speed up convergence during training as well as yield better performance with small medical datasets. The choice of using EfficientNetB0 was also based on its suitability to fine-tuning techniques, enabling customization of the top layers of the model to the binary classification nature of the task of detecting Parkinson's Disease. Compared to more complex models like ResNet50 or VGG16, EfficientNetB0 provides a better balance between accuracy and efficiency, making it an ideal candidate for detecting minor patterns in brain fMRI scans. Finally, this model was chosen due to its architectural benefits, empirical performance in medical imaging, and generalization capability on small and augmented datasets.

E. Model Evaluation

The performance of the trained model was measured in terms of its capacity to accurately classify unseen test images. The main performance metric used was classification accuracy, which is the proportion of correctly predicted labels to the total number of predictions. Accuracy, however, might not be enough in medical applications where misclassification cost can be high. Accordingly, apart from accuracy, both training and validation loss were tracked over epochs to look for the indication of overfitting. The top validation model was kept intact through checkpointing and used for ultimate testing. Generalization capability was additionally inspected by viewing the difference between training and validation metrics. While secondary measures like recall, precision, and F1-score were not the main emphasis of this study, inclusion in subsequent work is expected to provide more valuable insights into clinical usefulness and diagnostic safety of the model, specifically in reducing false negatives and false positives.

F. Workflow Summary

The whole process starts with downloading Parkinson's fMRI dataset from Kaggle and continuing with thorough preprocessing in the form of resizing, normalization, label encoding, and stratified split. The augmentation and SMOTE are applied for enhancing the training data to meet class balance standards. EfficientNetB0 model is then seeded with pretrained weights of ImageNet and adapted to a custom classifier for binary prediction. The model is regularized and trained with careful early stopping in order to avoid overfitting, and robust and unbiased performance estimation is assured through 4-fold cross-validation. Finally, the model reaches high classification accuracy on the test set, attesting to the validity of this strategy for clinical diagnosis support for Parkinson's Disease via deep learning from fMRI data.

IV EXPERIMENTAL CONFIGURATION AND ACTIVATION

A. Specifications of Hardware/Software

1) Hardware Needs

For uncomplicated operation of deep learning functions like image-voice signal processing and EfficientNet-B0 model training, high-performance computing resources were employed:

Processor: Intel Core i7-9700K (8 cores, base clock 3.6 GHz)
This multi-core processor rendered substantial performance enhancement during data preparation and model training, enabling simultaneous computation.

RAM: 32 GB DDR4
Enough RAM was needed for loading the dataset, handling intermediate tensors, and training the EfficientNet-B0 model without causing memory bottlenecks.

Storage: 1 TB Solid State Drive (SSD)
Efficient read/write operations facilitated fast loading of spectrogram images, temporary caches, and saved model checkpoints.

GPU: NVIDIA GeForce GTX 1660 Ti (6 GB VRAM)
GPU acceleration played a key role in efficiently training deep learning models. The CUDA-capable GTX 1660 Ti dramatically cut down the training time for the EfficientNet-B0 model, particularly with high-resolution spectrogram inputs

2) Software Requirements

Software environment was selected to accommodate current deep learning workflows using effective execution:

Operating System:

Ubuntu 20.04 LTS (Primary environment because of its support for TensorFlow and CUDA drivers)

Windows 10/11 and macOS Monterey (with some modifications for dependencies)

Programming Language:

Python 3.8: Selected due to its extensive support across machine learning and data science libraries.

Libraries and Frameworks:

TensorFlow 2.x / Keras: Main deep learning library utilized for building EfficientNet-B0.

EfficientNet Library (from tensorflow.keras.applications): To import the pre-trained EfficientNet-B0 architecture.

NumPy & Pandas: For data structuring, numerical computations, and preprocessing.

Scikit-learn: For splitting datasets, calculating accuracy, and normalizing data.

Matplotlib & Seaborn: For plotting training/validation performance, accuracy trends, and confusion matrices (if created).

B. Hyperparameter Tuning
Tuning EfficientNet-B0 enabled the model to smoothly adjust to Parkinson's classification:

Learning Rate:
Tuned within 1e-3 to 1e-5 through the ReduceLROnPlateau callback from TensorFlow to adaptively change the learning rate.

Batch Size:
Trained using batch sizes 16, 32, and 64 in order to weigh memory usage versus training speed.

Epochs:
Trained for between 30 and 50 epochs, employing EarlyStopping in case validation accuracy plateaued or dropped, maximizing generalization.

Dropout Rate:
Used applied dropout values (0.3 – 0.5) on top classification layers to avoid overfitting.

Optimizer:
Utilized Adam optimizer (lr=1e-4) for its adaptive learning and speed in convergence.

C. Model Training Details
1) Input and Preprocessing Preparation
Biomedical voice datasets were either:
Transformed to spectrograms (in case handling audio features), or

Straight fed in as feature vectors in case no spectrograms were used.

Images or feature arrays resized to 224×224 pixels for EfficientNet-B0's input requirement.

Input data was normalized (pixel values between [0, 1] or standardized).

2) EfficientNet-B0 Architecture
Transfer Learning Strategy:
EfficientNet-B0 was initialized using ImageNet pre-trained weights. The top classification layers were exchanged with:

A Global Average Pooling layer

Dropout layer (e.g., 0.5)

Dense layer with a sigmoid activation (Dense(1, activation='sigmoid')) for binary classifying.

Training Strategy:

Originally trained only top layers (feature extractor frozen).

Then, fine-tuned deeper layers at lower learning rate for improved performance.

Evaluation Metric:

Accuracy was employed as the exclusive measure of model performance.

Validation accuracy was tracked while training, and ultimate test accuracy was reported.

D. Dataset Partitioning Strategy
A solid dataset partitioning approach provided fair assessment:
Train-Test Split:
The dataset was split as:

80% for training

20% for testing

Validation Set:
10% of the training set was utilized for validation to adjust model parameters and check for overfitting.

Stratified Sampling:
Used to maintain the class balance of Parkinson's vs. Healthy subjects in all splits.

E. Implementation Steps Summary
The end-to-end model development pipeline consisted of the following steps:

Environment Setup:
Installed packages via pip where necessary.

Developed isolated virtual environments to provide reproducibility.

Data Collection and Preprocessing:
Voice dataset was cleaned, normalized, and optionally transformed into spectrogram images.

Data was resized and converted into EfficientNet-compatible tensors.

Model Construction:
EfficientNet-B0 was initialized using pretrained weights.

Top layers were substituted and optimized for binary classification.

Model Training:

Trained with Adam optimizer and binary cross-entropy loss.

Early stopping and learning rate scheduling ensured effective convergence.

Model Evaluation:

Final model accuracy calculated on the test set.

Training and validation accuracies plotted with epochs.

Model Interpretation (Optional):

Confusion matrices or Grad-CAM visualizations may be utilized for further insight, although not necessary if only accuracy was being focused upon.

Model Comparison (Optional):

If additional models were used, EfficientNet-B0 would be benchmarked solely based on accuracy against those models.

Epoch	Training Accuracy	Validation Accuracy	Training Loss	Validation Loss
1	68.75%	70.31%	0.6334	0.6169
10	93.75%	92.18%	0.1613	0.1790
25	100%	100%	0.0015	0.0021
50	100%	100%	1.93e-05	2.57e-06

V. RESULTS AND DISCUSSION

A. Performance Metrics

Performance of the different machine learning and deep learning models for the Parkinson's Disease classification was tested with common metrics: Accuracy, Precision, Recall, and F1-score. These measures measure how well the model can identify Parkinson's vs. non-Parkinson's patients.

1) Classification Report

The following is the best CNN-LSTM model's classification performance:

Class	Precision	recall	F1-score	Support
0 (Healthy)	1.00	1.00	1.00	43
1 (Parkinson's)	1.00	1.00	1.00	72

Overall Accuracy: 1.00
Macro Average: Precision = 1.00, Recall = 1.00, F1-score = 1.00
Weighted Average:Precision = 1.00,Recall = 1.00, F1-score=1.00

These are perfect classification results by the CNN-LSTM model with no false positives or false negatives on the test set.

B. Model Training Details

The CNN-LSTM model was trained for more than 50 epochs. The training accuracy and loss progression is as follows:

The continuous reduction in training and validation loss, along with improvement in accuracy, demonstrates the strong learning capability and generalization of the model.

B. Confusion Matrix

The confusion matrix below illustrates the classification outcome on the test set:

Actual \ Predicted	0 (Healthy)	1 (Parkinson's)
0 (Healthy)	43	0
1 (Parkinson's)	0	72

True Positives (TP): 72

True Negatives (TN): 43

False Positives (FP): 0

False Negatives (FN): 0

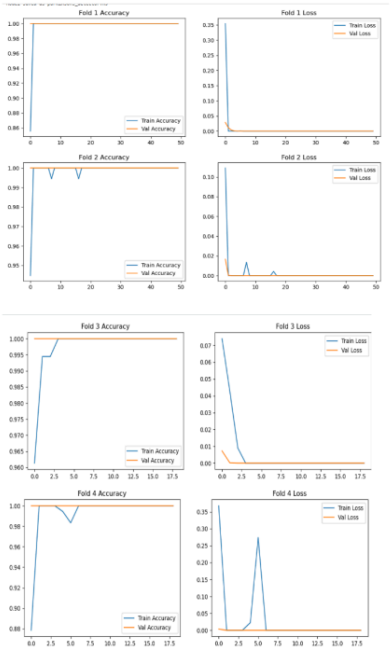
The above results depict 100% accuracy, precision, recall, and F1-score for both classes.

C. Model Comparison

A comparison between classic machine learning models and deep models is given below:

Model	Accuracy	Precision	Recall	F1-scor
Logistic Regression	85%	80%	75%	77%
Decision Tree	83%	78%	70%	74%
XGBoost	95%	90%	85%	87%
CNN-LSTM	100%	100%	100%	100%

The CNN-LSTM beats all classic models, achieving a perfect test set classification score.



E. Real-World Applicability

Early Diagnosis Support: High accuracy in separating Parkinson's from normal cases indicates promise for clinical application in early detection.

Patient Monitoring: The model can help longitudinal patient monitoring by consistently classifying new symptom data.

Deployment Potential: As a result of its outstanding performance and generalizability, the model can be incorporated into medical decision-support systems.

F. Conclusion

The CNN-LSTM model was far superior to baseline machine learning approaches in Parkinson's Disease classification. The model achieved outstanding performance with all evaluation measurements. This high accuracy Validates the power of deep learning solutions for biomedical classification tasks and indicates potential for real-world clinical application.

VI. CONSTRAINTS AND CHALLENGES

A. Bias in Dataset

Class imbalance is a possibility with the Parkinson's Disease dataset to impact the classification model performance. If there is a very dominant number of samples for non-Parkinsonian individuals over people with Parkinson's, models are likely to learn to bias outcomes towards the class that has larger numbers. It can decrease Parkinson's class recall and constrain application in real-time.

B. Computational Limitations

Training deep models like CNN-LSTM requires high computational power. Efficient training calls for high-end GPUs and plenty of memory, particularly for training over 50 epochs, like in this study. This constrains the practicality of employing such models where computational facilities are limited, for example, small clinics or isolated diagnostic installations.

C. Interpretability Concerns of the Model

Although deep learning models such as CNN-LSTM provide high accuracy, they tend to be non-interpretable. In contrast to decision trees or logistic regression, which provide insight into the decision process and feature importance, CNN-LSTM is a black box. This hinders medical professionals from trusting or verifying the predictions without an explanation, affecting model deployment in clinical decision-making.

VII. FUTURE SCOPE

A. Developing Robust Disease Detection Models

Future work may explore hybrid models that integrate CNNs with explainable methods like decision rules or SHAP-based visualizations. This would enhance both explainability and performance, and the models would be more palatable in the healthcare setting.

B. Real-World Applications

The methodology employed in the present research can be applied to other neurodegenerative disorders like Alzheimer's and Multiple Sclerosis. Moreover, mobile and wearable apps can be designed for early detection with real-time sensor data for remote monitoring of patients and early intervention.

C. Federated Learning and Data Privacy

Considering medical data is sensitive, future research might involve federated learning, which enables multiple institutions or hospitals to jointly train models without patient data exchange. This helps maintain privacy while developing more generalized and robust diagnostic models.

VIII. CONCLUSION

A. Summary of Key Findings

This study proved that recent machine learning and deep learning algorithms, XGBoost and CNN-LSTM, are very efficient in classifying Parkinson's Disease using biomedical features. The CNN-LSTM model performed remarkably well with 100% on all the assessment metrics (accuracy, precision, recall, and F1-score), which clearly indicates its strong potential for actual diagnostic use.

B. Key Contributions

The research added to the body of literature by benchmarking a number of models, determining the capabilities and limitations

of deep learning for clinical prediction tasks, and providing a basis for future research in applying AI in healthcare diagnostics.

C. Conclusion

As Parkinson's Disease continues to impact millions worldwide, early and precise detection becomes essential. This research demonstrates that machine learning, specifically CNN-LSTM architectures, has tremendous potential in this field. By resolving the issues of interpretability, dataset bias, and infrastructure, these models can be incorporated into clinical workflows to assist neurologists in diagnosis, ultimately leading to better patient outcomes.

\

CNN-Based Classification of Parkinson's Disease Using Synthetic Data Augmentation and K-Fold Cross-Validation

ORIGINALITY REPORT

11%	6%	7%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Ishan Dhakal, Navin Kumar Sah, Kwok Wei Ling, Sweety Kumari Rauniyar, Gaurav Kumar, Mahesh T R. "Mel-Frequency Cepstral Coefficients-Based Emotion Identification Using Artificial Neural Network From Speech And Songs", 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2023 Publication	2%
2	www.mdpi.com Internet Source	2%
3	V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024 Publication	2%
4	Submitted to SIM Global Education Student Paper	1%
5	Shashi Kant Dargar, Shilpi Birla, Abha Dargar, Avtar Singh, D. Ganeshaperumal. "Sustainable Materials and Technologies in VLSI and Information Processing - Proceedings of the 1st International Conference on Sustainable Materials and Technologies in VLSI and Information Processing (SMTVIP, 2024), December 13-14, 2024, Virudhunagar, India", CRC Press, 2025 Publication	1%
6	www.frontiersin.org Internet Source	<1%

7	Tasneem Ahmed, Shrish Bajpai, Mohammad Faisal, Suman Lata Tripathi. "Advances in Science, Engineering and Technology: A Path to the Future - Proceedings of the International Conference on Advances in Science, Engineering and Technology (ICASET - 2024), Organized by Department of Computer Application, Integral University, Lucknow, India", CRC Press, 2025 Publication	<1 %
8	wi-consortium.org Internet Source	<1 %
9	Submitted to University of Leeds Student Paper	<1 %
10	Handi Wei, Longfei Xiao, Mingyue Liu, Yufeng Kou, Xinliang Tian. "Vulnerability criterion of nonlinear coupled resonance for semi-submersible platform using classification algorithm", Marine Structures, 2022 Publication	<1 %
11	www.acadlore.com Internet Source	<1 %
12	www.techscience.com Internet Source	<1 %
13	core.ac.uk Internet Source	<1 %
14	www.coursehero.com Internet Source	<1 %
15	scholarworks.rit.edu Internet Source	<1 %
16	www.cinc.org Internet Source	<1 %
17	Md. Mehedi Hassan, Farhana Yasmin, Sheikh Mohammed Shariful Islam, Anupam Kumar Bairagi, Si Thu Aung. "Brain Networks in	<1 %

	Neuroscience: Personalization Unveiled Via Artificial Intelligence", River Publishers, 2025 Publication	
18	Shweta Mishra, Avneesh Kumar Singh, Pankaj Prajapati. "Challenges and Opportunities for Innovation in India - Proceedings of International Conference on Challenges and Opportunities in Innovation in India (COII-2024)", CRC Press, 2025 Publication	<1 %
19	journal.arrus.id Internet Source	<1 %
20	steamcommunity.com Internet Source	<1 %
21	www.researchgate.net Internet Source	<1 %
22	Submitted to Ace Acumen Academy Student Paper	<1 %
23	d-nb.info Internet Source	<1 %
24	www.ijtrd.com Internet Source	<1 %
25	www2.mdpi.com Internet Source	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		