Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 2 (2025)

Parkinson's Disease Detection on Unbalanced Speech Data using Convolutional Neural Networks

Mrs. A. G. Phakatkar [1], Vaishnavi Taware [2]

[1,2] Department of Computer Engineering, SCTR's Pune Institute of Computer Technology

Article History:

Received: *30-07-2024*

Revised: *08-09-2024*

Accepted: *17-09-2024*

Abstract:

Parkinson's disease is a progressive condition impacting movement and communication. Initially, symptoms manifest primarily in speech difficulties, which worsen over time, affecting aspects such as pitch and articulation. Detecting signs of Parkinson's disease often relies on analysing speech patterns. In this study, a Convolutional Neural Network is utilized for Parkinson's speech detection. Convolutional Neural Network excels in capturing subtle spatial structures and local patterns, crucial for discerning the nuanced pitch, rhythm, and phonetic traits of individuals with Parkinson's disease. The research employs acoustic voice measures like jitter and shimmer as speech input parameters, utilizing a dataset sourced from the UCI machine learning repository. The dataset exhibits a class imbalance problem. To address the class imbalance issue Synthetic Minority Over-sampling Technique algorithm is used. The Convolutional Neural Network algorithm significantly enhances Parkinson's disease voice detection exhibiting a remarkable accuracy of 91.52% outperforming traditional machine learning approaches.

Keywords: Parkinson's Disease; Machine Learning; Deep Learning; Convolutional Neural Networks; SMOTE

1. Introduction

Parkinson's disease (PD) is a progressive condition affecting both motor and non-motor functions, with no known cure. It impacts the brain's neurons responsible for dopamine production, crucial for coordination, resulting in a variety of motor symptoms like tremors and balance issues, as well as non-motor symptoms such as sleep disturbances and speech difficulties. PD has complex origins involving genetic and environmental factors.

Speech impairments are common non-motor symptoms of PD, stemming from reduced control over vocal muscles and cognitive decline. Analysis of PD speech reveals shorter phonation time, increased jitter and shimmer, and altered pitch range. The Unified Parkinson's Disease Rating Scale evaluates these symptoms comprehensively.

PD progresses through stages, with vocal cord injuries often appearing early. Vocal impairment is easily measurable and can be assessed remotely through telemedicine, enabling patients to conduct tests at home using their phones. Acoustic measurements provide valuable insights into speech characteristics, aiding diagnosis and research. This study focuses on detecting PD speech disorders using Convolutional Neural Networks (CNN), which excel at recognizing subtle variations in speech patterns. By analyzing pitch, rhythm, and phonics, CNNs offer a non-intrusive method for early PD detection. The paper outlines the use of Synthetic Minority Over-sampling Technique (SMOTE) techniques to handle imbalanced data and the implementation of CNN models for classification.

The paper's structure includes sections on related work, CNN algorithm explanation, SMOTE technique presentation, methodology for PD detection, experimental results, and concluding remarks with future research directions.

2. Literature Survey

Previous research in PD detection has explored various machine learning techniques, with an increasing focus on utilizing voice characteristics due to their early manifestation in PD patients. In [1] K-Nearest Neighbors (KNN) and Support Vector Machines (SVM) classifiers were applied to the UCI speech dataset for the identification of Parkinson’s voice disorder The dataset consisted of 80 subjects, where 40 were diagnosed with PD and 40 healthy individuals. Each person contributed three sound samples, resulting in a total of 240 samples. From these, 44 feature vectors were created, and 177 features were extracted. SVM achieved 91.25% accuracy and KNN achieved 91.23% accuracy. These results demonstrate the effectiveness of both techniques in accurately identifying Parkinson's voice disorder based on voice characteristics.

In [2], a comparative study of different machine learning classifiers and Twin-Support Vector Machine (TSVM) classifiers with and without feature selection methods was carried out. The dataset, sourced from the UCI Machine Learning Repository, consists of 147 samples categorized as PD and 48 samples categorized as HC. To address class imbalance issues, SMOTE oversampling method is implemented. Correlation-based feature subset selection method used. Using a forward selection method with Weka’s Best First searching techniques 13 features were selected. The different machine learning classifiers such as logistic regression (LR), Support Vector Machines (SVM), Naive Bayes (NB), Decision Tree (DT), and K-Nearest Neighbor (KNN) were used for comparative analysis. With Feature selection LR achieved 85.0%, SVM achieved 84.7%, Naïve Bayes achieved 80.6%, DT achieved 86.4%, and KNN achieved 88.8%. Among these algorithms, TSVM achieved the highest accuracy of 93.2% without feature selection and 93.9% with feature selection.

The comparative analysis of ensemble learning classifiers to classify PD is conducted in [3]. The authors investigate Parkinson's disease classification using vocal datasets and compare two different ensemble learning techniques. They introduce the stacking classifier and voting classifier to distinguish between PD patients and non-PD patients. The dataset comprises sustained phonation recordings of the vowel /a/ from 188 PD patients and 64 non-PD patients, sourced from the UCI machine learning repository created by the Department of Neurology at Cerrahpasa Faculty of Medicine, Istanbul University. The voting classifier achieves an accuracy of 88.8%, while the stacking classifier achieves the highest accuracy of 92.2%.

An innovative framework for detecting voice loss in Parkinson's Disease (PD) by employing a two-level feature selection process based on weight was introduced in [4]. Principal Component Analysis (PCA) and the Eigenvector Centrality Feature Selection (ECFS) methods were employed for feature selection. Features selection sets generated from PCA and ECFS were combined. The SVM classifier applied on hybrid weighted features selected set. The dataset was taken from the UCI machine learning repository which includes 756 instances and 753 features where each subject has 3 records. The proposed cubic kernel-SVM model achieved 94% accuracy, while the SVM classifier without the feature selection method achieved 88% accuracy.

Various machine learning techniques for PD speech feature classification were examined and compared using Multilayer Perceptron (MLP), KNN, SVM, Random Forest (RF), and Extreme Gradient Boosting (XGBooster) classifiers in [5]. The authors proposed a Principal Component Analysis-Support Vector Machine (PCA-SVM) and a Sparse Autoencoder- Support Vector Machine (SAE-SVM) two hybrid models for analyzing voice disorders in patients. This study uses a UCI speech dataset containing a total of 756 voice samples which contains 754 speech feature attributes. Out of 252 subjects,188 belonged to PD samples and the remaining 64 belonged to healthy samples. SMOTE techniques applied to the dataset. RF classifiers achieved 83.6% accuracy, MLP classifier achieved 84.5% accuracy, KNN classifier achieved 76.5% accuracy, XGBoost classifier achieved 88.1% accuracy and SVM classifier achieved 85.4% accuracy whereas the proposed model PCA-SVM achieved 88.9% accuracy and the highest accuracy 93.5% achieved by SAE-SVM model.

In [6], a deep learning approach, CNN, is employed to detect PD speech signals. Two-dimensional Convolutional Neural Networks (2D-CNNs) are used to extract dynamic features from speech signals. These features are processed using time-distributed techniques to capture temporal dynamics. Further, a one-dimensional CNN (1D-CNN) captured the dependencies among these dynamic features. The two datasets used in the study were collected from the GYENNO SCIENCE Parkinson Speech data and the PC-PITA database. The GYENNO SCIENCE dataset includes 45 individuals where 30 PD subjects and 15 healthy individuals. This dataset has dynamic voice features. The PC-PITA database consists of 100 subjects, where 50 people belong to healthy controls and 50 people belong to PD patients. This dataset has speech samples in Spanish. The proposed model achieves an accuracy of 81.6% for the sustained vowel /a/ task and 75.3% for reading short sentences on the GYENNO SCIENCE dataset. For the PC-PITA database, the proposed system achieved 92% accuracy for tasks involving reading a sentence in Spanish.

In [7] authors examine the utilization of deep learning and artificial intelligence methodologies for analyzing speech and language patterns in Parkinson's disease. To examine speech and language patterns in Parkinson's patients, they applied 1D and 2D CNN algorithms, as well as Wav2Vec 2.0, Bidirectional Encoder Representations from Transformers (BERT), and BETO models. The dataset contains 165 Colombian Spanish native speakers, where 80 PD cases and 85 healthy cases. The result obtained by the speech model using wav2vec 2.0 is 88% and by using CNN and BETO result achieved by the language model is 77%.

Machine learning classifiers, including SVM, Random Forests, Logistic Regression, and KNN, are applied and compared to analyze subtle changes and acoustic measures in the speech of individuals with Parkinson's disease (PD) [8]. The two datasets were taken from PPMI and UCI. The dataset contains acoustic speech measures such as shimmer, jitter, etc. The dataset contains a total of 31 individuals where 23 have PD .195 record attributes are collected from individuals. Random Forest classifier and SVM achieved the highest 91.83% result compared to other machine learning techniques.

The study [9] focuses on utilizing multi-class classification techniques to distinguish between Parkinson’s disease (PD) and Adductor Spasmodic Dysphonia (ADSD) based on voice disorder characteristics. The authors employ machine learning classifiers such as Naïve Bayes, Random Forest, and Multi-Layer Perceptron for multiclass voice disorder classification. The dataset collected from

PVT contains 51 PD subjects, 60 dysphonic subjects, and 111 healthy subjects. From the sustained vowel and Italian sentence voice dataset, a total of 6373 voice features were extracted. Statistical methods and genetic algorithms were used for feature selection. The combination of the genetic algorithm and MLP achieved 98.39 % accuracy, the RF classifier achieved 96.77 accuracy and the Naive Bayes classifier achieved the highest result 99.46% for multiclass classification.

A novel approach for the automatic and early detection of Parkinson's disease (PD) through the analysis of acoustic signals using classification algorithms proposed in [10]. The study evaluates selected features and performs hyperparameter tuning of machine learning (ML) algorithms. To balance the dataset, SMOTE is employed, while Recursive Feature Elimination (REF) is utilized for feature selection. Additionally, feature extraction is conducted using t-SNE and PCA algorithms. The machine learning classifiers applied in the study include KNN, SVM, Decision Trees (DT), Random Forests (RF), and Multi-Layer Perceptron (MLP). The dataset was taken from the UCI Machine Learning Repository and contains 195 samples, with 147 identified as Parkinson's Disease (PD) cases and 48 as healthy individuals. The Random Forest (RF) classifiers with the t-SNE algorithm, obtained an accuracy of 97%. Meanwhile, the combination of PCA and MLP achieved the highest accuracy of 98%.

3. Convolution Neural Network

In the realm of deep learning, CNNs stand out as specialized neural networks extensively utilized in various domains such as image processing, pattern recognition, video analysis, and speech recognition. Serving as an advanced iteration of Artificial Neural Networks (ANNs), CNNs are tailored to handle structured data like images and sequences efficiently. CNNs feature multiple layers, with convolutional layers being a cornerstone component. These layers apply filters to the input data, extracting essential features such as edges and textures. Subsequently, pooling layers are employed to down-sample the feature maps while preserving critical information. Finally, fully connected layers are responsible for making predictions based on the pooled outputs. The architecture of CNNs is characterized by its layered structure, as depicted in Fig. 1. Each layer plays a distinct role in the overall process, with convolutional layers crucial for extracting informative patterns via their filters. As the data progresses through subsequent layers, these patterns are further refined, ultimately contributing to the network's ability to make accurate diagnoses [11,12].

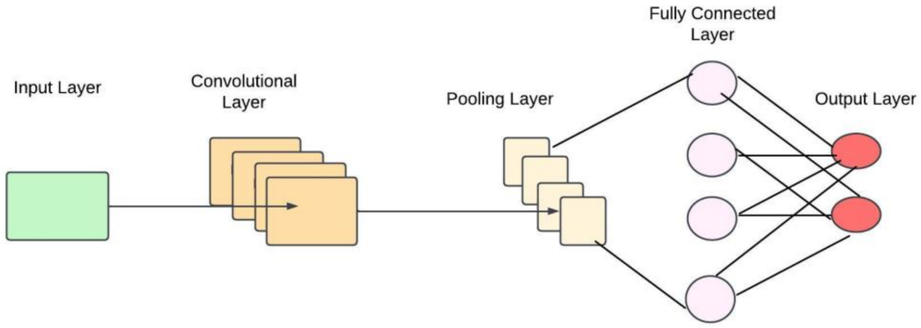

Fig. 1 Architecture of CNN [13]

Key components of CNN architecture

- Input layer: The input layer receives raw input data. Here the raw input data is speech features.
- Convolutional Layers: The convolutional layer uses a kernel which is also called a filter and it captures hierarchical representation and detects spatial patterns of input.
- Activation function: The activation function facilitates the network's capacity to grasp intricate relationships by introducing non-linear elements. Specifically, the rectified linear unit (ReLU) activation function substitutes negative values with zeros.
- Pooling (Subsampling or Down-sampling) Layer: Max pooling gives maximum value from a set of values and reduces the spatial dimensions of the input volume by preserving the most important features.
- Flattening: Flattening converts the high-dimensional feature maps into a one-dimensional vector and for a fully connected layer, it prepares data for processing.
- Fully connected layers: It is also called a dense layer. It connects every one-layer neuron to every next-layer neuron. This is the last stage of the network to perform various machine learning tasks such as regression and classification functions.
- Dropout: In most cases Dropout is optional. With dense layer dropout (0.5) added prevents overfitting by reducing the dependencies on specific neurons and encouraging the network to learn the important features. It is a regularization technique.
- Output layer: The output layer provides the outcomes and performs classification tasks. The SoftMax activation function is applied to transform the output into probabilities. At compilation, Adam Optimizer is used to optimize a neural network model. The obtained data is stored in a .h5 format of file for future use during the testing.

4. Synthetic Minority Over-sampling Technique

SMOTE, which stands for Synthetic Minority Over-sampling Technique, represents a cutting-edge statistical method devised to address the challenge of imbalanced datasets. Imbalance occurs when one class within a dataset possesses significantly fewer instances than another. This issue can distort the performance of machine learning algorithms. SMOTE effectively tackles this problem by generating synthetic samples for the minority class, thus rebalancing the class distribution. The technique achieves this by creating synthetic instances along line segments connecting existing minority class instances. By doing so, SMOTE ensures a more equitable distribution across classes within the dataset. This balanced representation enhances the effectiveness of various machine learning tasks, particularly classification [14,15].

5. Methodolgy

In this work, the speech dataset is sourced from the UCI Machine Learning Repository. Subsequently, preprocessing techniques are applied, followed by the resolution of class imbalance through the utilization of the SMOTE algorithm. Feature extraction and selection processes are then carried out. The model undergoes training using the CNN algorithm. The architectural layout of the methodology is illustrated in Fig. 2.

5.1 Data Pre-processing

During the data pre-processing phase, various checks are conducted, encompassing an assessment of information, identification of null values, handling of missing values, detection of duplicate values, outlier analysis and handling unbalanced speech data. Additionally, categorical data is transformed into numerical format, and Min-Max scaling is employed to standardize the numerical features, ensuring consistency within a range of -1 to 1. The SMOTE algorithm is applied to address class imbalance problems.

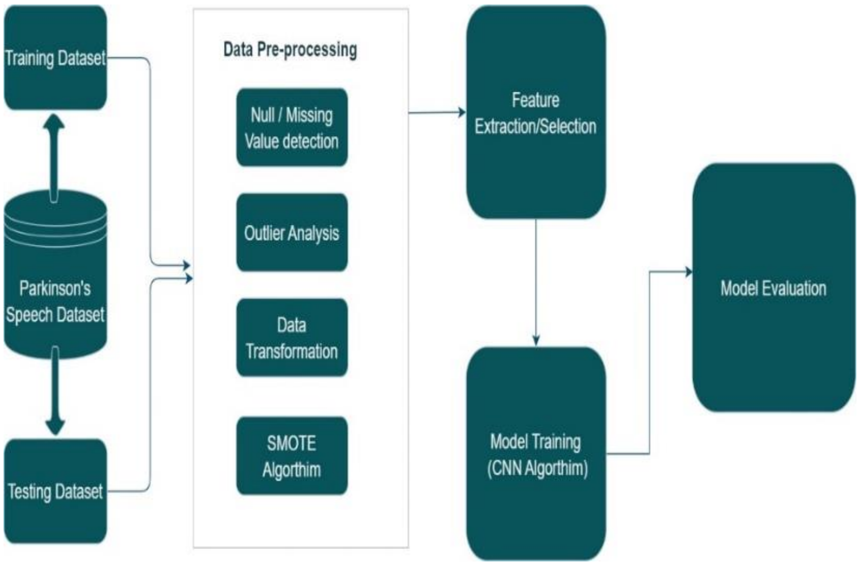

Fig. 2 System Architecture

5.2 Feature Extraction and Selection

Feature extraction plays a crucial role in developing an effective Parkinson’s detection system from speech samples. It involves reducing the dimensionality of data while retaining relevant information. In this study, speech parameters such as Jitter, amplitude, shimmer, MDVP, and pitch, along with 24 features, are considered. Out of the 24 feature attributes, all are chosen except for the attribute containing the names.

5.3 Model Training

To train the CNN classifier for distinguishing Parkinson’s patients from healthy individuals using the speech dataset, the raw input data comprises speech features. The convolutional layer was added as a first layer with filter 32 and kernel size 3 x 3. For the second layer of the convolution neural network,64 Kernels of size 3 x 3 are used. The convolutional neural network added the third layer with filter 128 of size 3 x 3. 1D CNN is used with filter size (32,64,128) and Kernel size (3) along with RELU activation function. Following the convolutional layer and activation function a max pooling 1D layer is added with kernel Size 2 x 2. Flattening is added. With Batch normalization, the first dense layer of size 256 is added with the RELU activation function. The second dense layer is added of size 512. These two layers are added with the RELU activation function. The third dense layer is added of size 10 with SoftMax activation function. With dense layer dropout (0.5) added. The SoftMax activation function is applied in the output layer. At compilation, Adam Optimizer is used to optimize a neural network model with epochs 20 and batch size 32.

Algorithm: CNN Model for Parkinson Diseases Detection

```
Input: Speech dataset from UCI Machine Learning Repository
Output: Trained Model, Evaluation Results
START
Step 1: Load speech dataset
Step 2: Splitting Data and Oversampling Train set with SMOTE sampling technique
 Step 3: Model Definition
3.1 Add Convolutional, Pooling, Dropout, and Dense layers to the model
 Step 4: Model Training
4.1 Compile the model with Adam optimizer
4.2 Define epochs (20), batch_size(32) and early stopping
4.3 Train the model on the oversampled data using early stopping for regularization
Step 5: Evaluate the trained model on the testing data and compute the accuracy
```

6. Experimental Result

This section describes the dataset particulars and performance criteria employed to assess the model's capability in detecting PD speech disorders. The study's experiments illustrate the efficacy of tackling an imbalanced dataset through the utilization of the SMOTE algorithm, resulting in superior PD detection performance. Employing a CNN model, the system achieves PD patient classification. Performance measures including precision, recall, and F1-score are calculated to gauge the model's efficacy. The experimental investigation was conducted utilizing Python programming within a Google Colaboratory (Google Colab) environment.

6.1 Dataset Description

The speech dataset originates from the UCI Machine Learning Repository, developed by Max Little in collaboration with the National Centre for Voice and Speech in Colorado. It comprises 195 voice samples collected from 31 individuals, including 23 with Parkinson's disease and 8 healthy controls. Each subject contributed six voice samples, yielding a total of 24 attributes per sample. The dataset captures the averages of six phonation’s per subject, ranging from 1 to 36 seconds in duration.

6.2 Performance Parameter for Model Evaluation

To classify Parkinson’s disease, we divided the dataset into three sections: the training set, comprising 70% of the dataset; and the testing set, comprising 30%. During the testing phase, performance evaluation metrics such as precision, sensitivity, specificity, and accuracy were employed to assess the model's effectiveness.

$$Precision = \frac{TP}{TP+FP}. \quad (1)$$

$$Recall = \frac{TP}{TP+FN}. \quad (2)$$

$$F1 = 2 * \frac{\text{Precision + Recal}}{\text{Precision}\times\text{Recall}}. \quad (3)$$

$$\text{Accuracy: } \frac{TP+TN}{TP+\text{TN+FP}+FN}. \quad (4)$$

6.3 Class Imbalance Problem

The speech dataset provided presents a class imbalance challenge, with 147 samples attributed to the PD group and only 48 samples representing healthy individuals. To rectify this imbalance, the SMOTE technique is utilized. SMOTE uses only the training data to generate synthetic samples for the minority class. The testing data remains completely separate and untouched during this process. By applying the SMOTE algorithm, the number of samples in the healthy class (minority) is augmented from 48 to 147 through the generation of new synthetic samples. Consequently, the class imbalance issue is effectively addressed. The SMOTE technique showcases promising outcomes in this context. Figure 3 depicts the class distribution before the application of the SMOTE algorithm, highlighting a class imbalance scenario where 147 samples are associated with status 1 (unhealthy), and 48 samples are associated with status 0 (healthy).

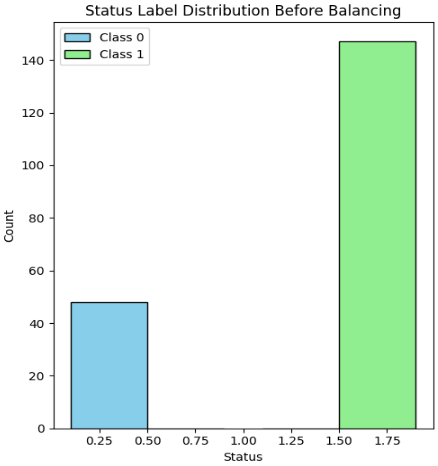

Fig. 3 Class distribution of PD dataset before applying SMOTE algorithm.

After applying the SMOTE algorithm to the PD dataset, the count of the healthy group (status 0) has been increased to match the count of the unhealthy group (status 1), effectively managing the class distribution, as depicted in the following fig. 4.

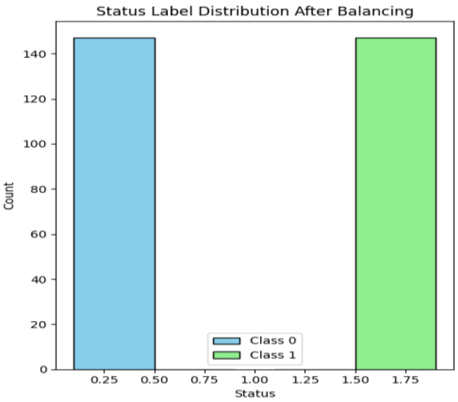

Fig. 4 Class distribution of PD dataset after applying SMOTE algorithm.

6.3 Results and Discussions

The study utilizes the CNN algorithm with the Adam optimizer to classify PD patients, employing a dataset collected from the UCI Machine Learning Repository. Confusion matrices are generated to evaluate the model's performance, alongside precision, recall, and F1-score calculations. Additionally, the study visualizes accuracy and loss graphs of the model. Experimental findings reveal that employing the SMOTE algorithm with CNN, using 20 epochs and a batch size of 32, achieves an accuracy of 91.52% for acoustic speech features. A comparison is conducted before and after applying the SMOTE algorithm to address the class imbalance issue in the PD dataset. Initially, with 48 samples in the healthy group (class 0) and 147 in the unhealthy group (class 1), the dataset exhibits skewness towards PD patients. The CNN model applied to the unbalanced dataset yields an accuracy of 74.35%. However, performance metrics such as precision, recall, and F1-score are relatively low.

To improve model performance, the SMOTE algorithm is utilized to balance the dataset. Subsequently, the CNN model is applied to the balanced dataset, resulting in an accuracy of 91.52%. Notably, precision metrics demonstrate superior performance compared to the pre-SMOTE application. Table 3 depicts the performance evaluation metrics before and after employing the SMOTE algorithm. Prior to SMOTE, the CNN model exhibits lower performance metrics. However, post-SMOTE application, there is a significant enhancement in the CNN model's performance, leading to overall higher performance metrics.

Table 3 Performance Evaluation Metrics Before and After Applying SMOTE

	Before SMOTE	After SMOTE
Precision	60%	92.01%
Recall	62.05%	91.43%
F1-score	60.68%	91.48%
Accuracy	71.54%	91.52%

The confusion matrix provides a tabular summary of the proposed model's performance, detailing the count of accurate and inaccurate predictions of Parkinson’s disease. It serves as a visual aid in evaluating the model's effectiveness. In this representation, individuals with Parkinson's disease are represented by 1, while healthy individuals are denoted by 0.

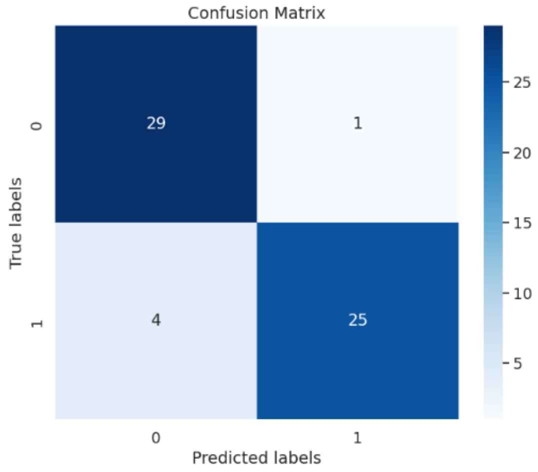

Fig. 5 Confusion Matrix

Figure 5 displays the confusion matrix illustrating the classification results of Parkinson’s disease. Within the confusion matrix, the true label indicates the actual presence or absence of PD in a given sample, reflecting the correct classification of whether PD is present or absent. The predicted label signifies the model's prediction regarding the presence or absence of PD in the sample, based on learned patterns from the training data. The CNN model identifies 29 true negatives (TN), 25 true positives (TP), 1 false positive (FP), and 4 false negatives (FN).

Where:

- TP (True Positive): The system correctly identifies the presence of PD, where both the actual and predicted labels indicate the disease is present.
- TN (True Negative): It accurately identifies the absence of PD, where both the actual and predicted labels indicate the absence of PD.
- FP (False Positive): The system incorrectly suggests the presence of PD when it is absent. The actual label indicates no disease, but the prediction suggests otherwise.
- FN (False Negative): The system fails to detect PD despite its actual presence. The actual label indicates the disease, but the prediction fails to identify it.

The CNN model is trained using the Adam optimizer for 20 epochs with a batch size of 32. Following this training, graphs are generated to visualize the model's accuracy and loss. These graphs offer insights into the accuracy achieved by both the training and validation datasets. Higher accuracy values suggest better model performance. The loss graph illustrates the model's performance after each iteration, with lower loss values indicating improved performance, whereas higher loss values suggest suboptimal performance. Figure 6 displays the Model Accuracy and Model Loss graphs.

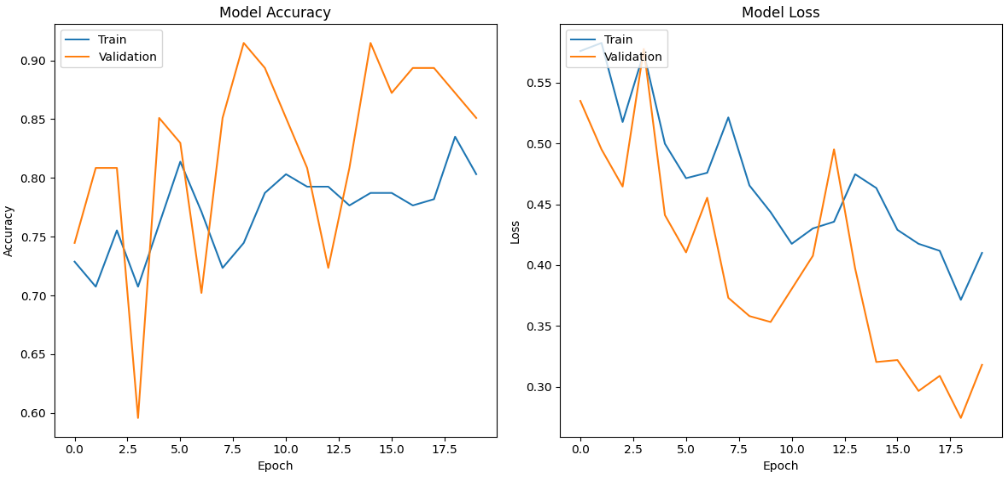

Fig. 6 The graph of Model Accuracy and Model Loss

7. Conclusion and Future Work

Speech presents itself as a promising biomarker for monitoring health conditions, including the detection of Parkinson’s disease. Deep learning methodologies have demonstrated superior efficacy in identifying Parkinson's speech disorders. In this investigation, we utilized a CNN model to discern Parkinson’s disease through speech analysis. The CNN model underwent training using a dataset comprising acoustic speech metrics like shimmer, jitter, and MDVP. To address class imbalance challenges, the SMOTE algorithm was employed. The CNN model showcased exceptional accuracy, reaching 91.52% in distinguishing individuals with Parkinson's disease from those without, based on speech data. Furthermore, the model achieved a precision of 92.01%, recall of 91.43%, and F1-score of 91.48% for speech attributes.

Future research endeavors will delve into refining the CNN model, potentially by augmenting layers and integrating various deep learning techniques. Additionally, the inclusion of novel features alongside existing ones may further enhance the model's efficacy. Furthermore, exploring hybrid algorithmic combinations such as CNN-LSTM and CNN-RNN holds promise in potentially augmenting PD speech detection capabilities.

References

[1] Yaman, Orhan, Fatih Ertam, and Turker Tuncer. "Automated Parkinson’s disease recognition based on statistical pooling method using acoustic features." *Medical hypotheses* 135 (2020): 109483.

[2] Thapa, Surendrabikram, Surabhi Adhikari, Awishkar Ghimire, and Anshuman Aditya. "Feature selection based twin-support vector machine for the diagnosis of Parkinson’s disease." In 2020 IEEE 8th R10 humanitarian technology conference (R10-HTC), pp. 1-6. IEEE, 2020.

[3] Younis Thanoun, Mohammed, and M. O. H. A. M. M. A. D. T. YASEEN. "A comparative study of Parkinson disease diagnosis in machine learning." In Proceedings of the 4th International Conference on Advances in Artificial Intelligence, pp. 23-28. 2020.

[4] Ashour, Amira S., Majid Kamal A. Nour, Kemal Polat, Yanhui Guo, Wafaa Alsaggaf, and Amira El-Attar. "A novel framework of two successive feature selection levels using weight-based procedure for voice-loss detection in Parkinson’s disease." *Ieee Access* 8 (2020): 76193-76203.

[5] Hoq, Muntasir, Mohammed Nazim Uddin, and Seung-Bo Park. "Vocal feature extraction-based artificial intelligent model for Parkinson’s disease detection." *Diagnostics* 11, no. 6 (2021): 1076.

[6] Quan, Changqin, Kang Ren, Zhiwei Luo, Zhonglue Chen, and Yun Ling. "End-to-end deep learning approach for Parkinson’s disease detection from speech signals." *Biocybernetics and Biomedical Engineering* 42, no. 2 (2022): 556-574.

[7] Escobar-Grisales, Daniel, Cristian David Ríos-Urrego, and Juan Rafael Orozco-Arroyave. "Deep learning and artificial intelligence applied to model speech and language in Parkinson’s disease." *Diagnostics* 13, no. 13 (2023): 2163.

[8] Govindu, Aditi, and Sushila Palwe. "Early detection of Parkinson's disease using machine learning." *Procedia Computer Science* 218 (2023): 249-261.

[9] Cesarini, Valerio, Giovanni Saggio, Antonio Suppa, Francesco Asci, Antonio Pisani, Alessandra Calculli, Rayan Fayad, Mohamad Hajj-Hassan, and Giovanni Costantini. "Voice Disorder Multi-Class Classification for the Distinction of Parkinson’s Disease and Adductor Spasmodic Dysphonia." *Applied Sciences* 13, no. 15 (2023): 8562.

[10] Alalayah, Khaled M., Ebrahim Mohammed Senan, Hany F. Atlam, Ibrahim Abdulrab Ahmed, and Hamzeh Salameh Ahmad Shatnawi. "Automatic and early detection of Parkinson’s disease by analyzing acoustic signals using classification algorithms based on recursive feature elimination method." *Diagnostics* 13, no. 11 (2023): 1924.

[11] Montesinos-López, Osval & Montesinos, Abelardo & Crossa, Jose. (2022). Convolutional Neural Networks. 10.1007/978-3-030-89010-0_13.

[12] Vaz, Joel Markus, and S. Balaji. "Convolutional neural networks (CNNs): Concepts and applications in pharmacogenomics." *Molecular diversity* 25, no. 3 (2021): 1569-1584.

[13] Hadi, Muhammad Usman, Rizwan Qureshi, Ayesha Ahmed, and Nadeem Iftikhar. "A lightweight CORONA-NET for COVID-19 detection in X-ray images." *Expert Systems with Applications* 225 (2023): 120023.

[14] Polat, Kemal. "A hybrid approach to Parkinson disease classification using speech signal: the combination of smote and random forests." In *2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT)*, pp. 1-3. Ieee, 2019.

[15] Chawla, Nitesh V., Kevin W. Bowyer, Lawrence O. Hall, and W. Philip Kegelmeyer. "SMOTE: synthetic minority over-sampling technique." *Journal of artificial intelligence research* 16 (2002): 321-357.

[16] Little,Max. (2008). Parkinsons. UCI Machine Learning Repository..

[17] Ouhmida, Asmae, Abdelhadi Raihani, Bouchaib Cherradi, and Sara Sandabad. "Parkinson’s diagnosis hybrid system based on deep learning classification with imbalanced dataset." *International Journal of Electrical and Computer Engineering (IJECE)* 13, no. 3 (2023): 3204-3216.

[18] Reddy, Mittapalle Kiran, and Paavo Alku. "Exemplar-Based Sparse Representations for Detection of Parkinson's Disease From Speech." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 31 (2023): 1386-1396.

[19] Quan, Changqin, Kang Ren, and Zhiwei Luo. "A deep learning based method for Parkinson’s disease detection using dynamic features of speech." *IEEE Access* 9 (2021): 10239-10252.

[20] Ogawa, Mitsuhiro, and Yiran Yang. "Residual-Network-Based Deep Learning for Parkinson's Disease Classification using Vocal Datasets." In *2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech)*, pp. 275-277. IEEE, 2021.

[21] Asmae, Ouhmida, Raihani Abdelhadi, Cherradi Bouchaib, Sandabad Sara, and Khalili Tajeddine. "Parkinson’s disease identification using KNN and ANN Algorithms based on Voice Disorder." In *2020 1st International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)*, pp. 1-6. IEEE, 2020.

[22] Demir, Fatih, Abdulkadir Sengur, Ali Ari, Kamran Siddique, and Mohammed Alswaitti. "Feature mapping and deep long short term memory network-based efficient approach for Parkinson’s disease diagnosis." *IEEE Access* 9 (2021): 149456-149464.

[23] Sharanyaa, S., P. N. Renjith, and K. Ramesh. "Classification of Parkinson's disease using speech attributes with parametric and nonparametric machine learning techniques." In *2020 3rd international conference on intelligent sustainable systems (ICISS)*, pp. 437-442. IEEE, 2020.

[24] Zahid, Laiba, Muazzam Maqsood, Mehr Yahya Durrani, Maheen Bakhtyar, Junaid Baber, Habibullah Jamal, Irfan Mehmood, and Oh-Young Song. "A spectrogram-based deep feature assisted computer-aided diagnostic system for Parkinson’s disease." *IEEE Access* 8 (2020): 35482-35495.

[25] Bocklet, Tobias, Elmar Nöth, Georg Stemmer, Hana Ruzickova, and Jan Rusz. "Detection of persons with Parkinson's disease by acoustic, vocal, and prosodic analysis." In *2011 IEEE workshop on automatic speech recognition & understanding*, pp. 478-483. IEEE, 2011.

[26] Agarwal, Aarushi, Spriha Chandrayan, and Sitanshu S. Sahu. "Prediction of Parkinson's disease using speech signal with Extreme Learning Machine." In *2016 international conference on electrical, electronics, and optimization techniques (ICEEOT)*, pp. 3776-3779. IEEE, 2016..

[27] Rueda, Alice, and Sridhar Krishnan. "Feature analysis of dysphonia speech for monitoring Parkinson's disease." In *2017 39th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*, pp. 2308-2311. IEEE, 2017.

[28] Wan, Shaohua, Yan Liang, Yin Zhang, and Mohsen Guizani. "Deep multi-layer perceptron classifier for behavior analysis to estimate Parkinson’s disease severity using smartphones." *IEEE Access* 6 (2018): 36825-36833.

[29] Moro-Velazquez, Laureano, JaeJin Cho, Shinji Watanabe, Mark A. Hasegawa-Johnson, Odette Scharenborg, Heejin Kim, and Najim Dehak. "Study of the Performance of Automatic Speech Recognition Systems in Speakers with Parkinson's Disease." In *Interspeech*, vol. 9, pp. 3875-3879. 2019.

[30] Ali, Liaqat, Ce Zhu, Zhonghao Zhang, and Yipeng Liu. "Automated detection of Parkinson’s disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network." *IEEE journal of translational engineering in health and medicine* 7 (2019): 1-10.

[31] Gunduz, Hakan. "Deep learning-based Parkinson’s disease classification using vocal feature sets." *Ieee access* 7 (2019): 115540-115551.

[32] Appakaya, Sai Bharadwaj, and Ravi Sankar. "Parkinson’s disease classification using pitch synchronous speech segments and fine Gaussian kernels based SVM." In *2020 42nd annual international conference of the IEEE engineering in medicine & biology society (EMBC)*, pp. 236-239. IEEE, 2020.

[33] Moro-Velazquez, Laureano, Jesus Villalba, and Najim Dehak. "Using x-vectors to automatically detect parkinson’s disease from speech." In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 1155-1159. IEEE, 2020.

[34] Amato, Federica, Luigi Borzì, Gabriella Olmo, Carlo Alberto Artusi, Gabriele Imbalzano, and Leonardo Lopiano. "Speech impairment in Parkinson’s disease: acoustic analysis of unvoiced consonants in Italian native speakers." *IEEE Access* 9 (2021): 166370-166381.

[35] Liu, Yuanyuan, Nelly Penttilä, Tiina Ihalainen, Juulia Lintula, Rachel Convey, and Okko Räsänen. "Language-independent approach for automatic computation of vowel articulation features in dysarthric speech assessment." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 29 (2021): 2228-2243.

[36] Demir, Fatih, Abdulkadir Sengur, Ali Ari, Kamran Siddique, and Mohammed Alswaitti. "Feature mapping and deep long short term memory network-based efficient approach for Parkinson’s disease diagnosis." *IEEE Access* 9 (2021): 149456-149464.

[37] Laganas, Christos, Dimitrios Iakovakis, Stelios Hadjidimitriou, Vasileios Charisis, Sofia B. Dias, Sevasti Bostantzopoulou, Zoe Katsarou et al. "Parkinson’s disease detection based on running speech data from phone calls." *IEEE Transactions on Biomedical Engineering* 69, no. 5 (2021): 1573-1584.

[38] Liu, Yuanyuan, Mittapalle Kiran Reddy, Nelly Penttilä, Tiina Ihalainen, Paavo Alku, and Okko Räsänen. "Automatic Assessment of Parkinson's Disease Using Speech Representations of Phonation and Articulation." *IEEE/ACM Transactions on Audio, Speech, and Language Processing* 31 (2022): 242-255.