Frontiers in Computing and Intelligent Systems
ISSN: 2832-6024 | Vol. 10, No. 3, 2024

Parkinson's Disease Speech Classification Using 1D Convolutional Neural Networks

Zhen Yue [1], Haiyang Wang [2]

[1] School of Information and Control Engineering, Jilin Institute of Chemical Technology, Jilin Jilin, 132021, China
[2] School of Mechanical and Control Engineering, Baicheng Normal University, Baicheng Jilin, 137000, China

Abstract: Parkinson's Disease (PD) is a neurodegenerative disorder caused by a lack of dopamine secretion. Both motor and non-motor activities of Parkinson's patients are affected. This study proposes a method for Parkinson's disease audio feature classification based on Convolutional Neural Networks (CNN). By extracting features from the speech signals of Parkinson's patients, and leveraging the powerful feature extraction and classification capabilities of CNNs, an efficient diagnosis of Parkinson's disease can be achieved. To evaluate the performance of the proposed method, experiments were conducted on two datasets, achieving accuracy rates of 100% and 92.86%, respectively.

Keywords: Parkinson Disease; Speech Features; Convolutional Neural Networks.

1. Introduction

Parkinson's Disease (PD) is a common neurodegenerative disorder characterized by motor impairments, tremors, muscle rigidity, as well as speech and gait problems. Speech features are one of the common symptoms of PD patients, and analyzing the speech signals of patients can provide important auxiliary information for early diagnosis of the disease. Traditional Parkinson's speech diagnosis methods often rely on the experience of clinical experts, and the evaluation process is cumbersome and subjective. In recent years, with the development of machine learning and deep learning technologies, audio signal-based automated diagnostic methods have gradually become a research hotspot. Sakar et al. [1] (2013) used a KNN+SVM approach for classification. Naranjo et al. [2] (2016) proposed a clinical expert system based on Bayesian methods for detecting Parkinson's disease, achieving an accuracy of 75.20%. Ali et al. [3] (2019) applied the chi-square statistical feature selection method to select the best features from Parkinson's disease datasets, achieving an accuracy of 100%. Traditional diagnostic methods mainly rely on the clinical experience of doctors and physical examinations of patients, but these methods often suffer from subjectivity and long diagnostic periods. In contrast, audio signal-based diagnostic methods can quickly and objectively assess the health status of patients by analyzing their speech features, offering great potential. Researchers can effectively use machine learning and related algorithms to perform early diagnosis of Parkinson's disease, monitor disease progression, and provide personalized treatment plans, thus offering new possibilities for early intervention of Parkinson's disease.

2. Dataset and Preprocessing

2.1. Dataset

Dataset1

This dataset was collected in collaboration between Professor Max Little from Oxford University and the National Centre for Voice and Speech (Little et al., 2007). It contains 195 speech recording samples from 31 individuals, 23 of whom are Parkinson's disease (PD) patients. Each individual has approximately 6 recordings, with each recording having 22 speech-related features.

Dataset2

This dataset includes speech recordings from 40 individuals (20 PD patients and 20 healthy individuals) recorded at the Neurology Department of the Cerrahpaşa Medical Faculty, Istanbul University (Sakar et al., 2013). The dataset contains 1040 samples, including sentences, vowels, numbers, and words. It consists of two independent training and testing files. The training set contains multiple speech recordings, while the testing set includes recordings from 28 PD patients, who pronounced two vowels ("a" and "o") three times, resulting in 168 samples. Each individual in the training set has 26 recordings, while each individual in the testing set has only 6 recordings. Each recording contains 26 speech-related features.

2.2. Preprocessing

For Dataset 1, the first 6 samples of each subject are stacked to form a shape of (6, 22). For Dataset 2, since the number of samples per subject differs between the training set (26 samples per individual) and the testing set (6 samples per individual), the first 6 samples of each subject in the training set are selected and stacked to form a shape of (6, 26), ensuring consistency with the number of samples in the testing set. The training and testing sets are split in an 8:2 ratio.

In Dataset 2, the test set provided by the authors contains only one class, PD patients. Therefore, we mix the samples from the training and testing sets and split them again in an 8:2 ratio to ensure balanced representation of both classes.

This translation accurately reflects the structure and content of your original text while maintaining clarity and technical accuracy. Let me know if you need further adjustments!

3. Models and Results

3.1. Model Architecture

Table 1. Model Architecture

Layer (type)	Filters	Kernel Size	Activation Function
Conv1d	32	3	ReLU
Maxpooling1d	-	2	-
flatten	-	-	-
dense	64	-	ReLU
dense	1	-	Sigmoid

3.2. Results

The four-evaluation metrics used as results are Precision, Recall, F1 Score, and Accuracy.

Table 2. Classification Performance

Dataset	Precision	Recall	F1 Score	Accuracy
Dataset1	100%	100%	100%	100%
Dataset2	88.89%	100.00%	94.12%	92.86%

Table 3. Comparison of Dataset 1 with Other Methods

Authors	Classifier	Accuracy
Fayyazifar et al. (2017) [5]	Bagging algorithm	98.28%
Haq et al. (2018) [6]	deep NN	98.00%
Haq et al. (2019) [7]	L1-Norm SVM	99.00%
Sharma et al (2020)[8]	kNN	99.25%
Prposed	1D-CNN	100%

Table 4. Comparison of Dataset 2 with Other Methods

Authors	Classifier	Accuracy
Sakar et al. (2013)[1]	kNN +SVM	68.45%
Li et al. (2017)[9]	Ensemble learning algorithm	86.5%
Ali et al. (2019)[3]	NN	100%
Sharma et al(2020)[8]	kNN	94.54%
Pramanik et al (2023) [9]	SysFor	92.86%
Ali L et al (2024) [10]	DNN	97.5%
Prposed	1D-CNN	92.86%

4. Summary

The results of using a one-dimensional convolutional neural network for Parkinson's speech classification show varying performance across different datasets. In Dataset1, the model achieves 100% in all metrics (Precision, Recall, F1 Score, and Accuracy), demonstrating excellent performance. In Dataset2, the model still performs well, despite a slightly lower precision (88.89%), with a perfect recall of 100%, indicating that it effectively identifies all Parkinson's patients. Overall, the model exhibits high accuracy and stability in the classification task.

Acknowledgments

Natural Science Foundation of Science and Technology Department of Jilin Province, No. 20130101170JC; Jilin Provincial Department of Education "12th Five-Year" Science and Technology Research Fund Project (2014404).

References

[1] Sakar B E, Isenkul M E, Sakar C O, et al. Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings[J]. IEEE journal of biomedical and health informatics, 2013, 17(4): 828-834.

[2] Naranjo L, Perez C J, Campos-Roca Y, et al. Addressing voice recording replications for Parkinson's disease detection[J]. Expert Systems with Applications, 2016, 46: 286-292.

[3] Ali L, Zhu C, Zhou M, et al. Early diagnosis of Parkinson's disease from multiple voice recordings by simultaneous sample and feature selection[J]. Expert Systems with Applications, 2019, 137: 22-28.

[4] Little M, Mcsharry P, Roberts S, et al. Exploiting nonlinear recurrence and fractal scaling properties for voice disorder detection[J]. Nature Precedings, 2007: 1-1.

[5] Fayyazifar, N., & Samadiani, N. (2017). Parkinson's disease detection using ensemble techniques and genetic algorithm. In 2017 artificial intelligence and signal processing (AISP) (pp. 162–165). Shiraz.

[6] Haq A U, Li J, Memon M H, et al. Comparative analysis of the classification performance of machine learning classifiers and deep neural network classifier for prediction of Parkinson disease[C]//2018 15th international computer conference on wavelet active media technology and information processing (ICCWAMTIP). IEEE, 2018: 101-106.

[7] Haq, A. U., Li, J. P., Memon, M. H., Khan, J., Malik, A., Ahmad, T., … Shahid, M. (2019). Feature selection based on L1-norm support vector machine and effective recognition system for Parkinson's disease using voice recordings. IEEE Access, 7, 37718–37734. https://doi.org/10. 1109/ ACCESS. 2019. 2906350

[8] Sharma S R, Singh B, Kaur M. Classification of Parkinson disease using binary Rao optimization algorithms[J]. Expert Systems, 2021, 38(4): e12674.

[9] Li Y, Yang L, Wang P, et al. Classification of Parkinson's disease by decision tree based instance selection and ensemble learning algorithms[J]. Journal of Medical Imaging and Health Informatics, 2017, 7(2): 444-452.

[10] Pramanik M, Pradhan R, Nandy P, et al. The ForEx++ based decision tree ensemble approach for robust detection of Parkinson's disease[J]. Journal of Ambient Intelligence and Humanized Computing, 2023, 14(9): 11429-11453.

[11] Ali L, Javeed A, Noor A, et al. Parkinson's disease detection based on features refinement through L1 regularized SVM and deep neural network[J]. Scientific Reports, 2024, 14(1): 1333.