

Early Detection of Parkinson's Disease using CNN And RNN

KishoreKumarKalli

Department of ArtificialIntelligence and DataScience
Koneru Lakshmaiah Education Foundation
Vaddeswaram
2100080178ai.ds@gmail.com

GadugoluRoshitha

Department of ArtificialIntelligence and DataScience
Koneru Lakshmaiah Education Foundation
Vaddeswaram
2100080198ai.ds@gmail.com

Dr.A.SrinivasaRao

Department of ArtificialIntelligence and DataScience
Koneru Lakshmaiah Education Foundation
Vaddeswaram
asrinivasarao1@kluniversity.in

Parkinson's Disease (PD) presents a significant challenge due to its progressive nature and the array of motor and non-motor symptoms it entails. One of the early manifestations of PD often includes vocal impairments, making diagnosis systems that leverage vocal features crucial in early detection efforts. Our study proposes two distinct frameworks based on Convolutional Neural Networks (CNNs) aimed at classifying PD using sets of vocal features. The distinguishing factor between these frameworks lies in their approach to combining these feature sets. In the first framework, feature sets are amalgamated before inputting them into a 9-layered CNN. Conversely, the second framework adopts a parallel input layer strategy, wherein feature sets are fed directly into separate convolution layers, allowing for simultaneous extraction of deep features before merging them in a subsequent layer.

The effectiveness of our proposed models was evaluated using a dataset sourced from the UCI Machine Learning repository, with performance validation conducted through Leave-One-Person-Out Cross Validation (LOPO CV). Given the imbalanced class distribution within our data, we employed F-Measure and Matthews Correlation

Coefficient metrics alongside accuracy for comprehensive assessment. Experimental results showcased promising outcomes, particularly with the second framework, which demonstrated superior performance. By leveraging parallel convolution layers, this framework efficiently captured deep features from each feature set, enhancing the classifiers' discriminative power. Notably, the extracted deep features not only effectively differentiated PD patients from healthy individuals but also significantly bolstered the overall classification efficacy.

The comparison between the two proposed frameworks highlights the efficacy of the parallel input layer strategy, particularly in capturing and leveraging deep features from distinct feature sets. These findings not only contribute to the advancement of PD detection systems but also emphasize the importance of innovative approaches in mitigating the challenges posed by imbalanced data distributions and heterogeneous symptom profiles within neurodegenerative diseases like PD. Experimental results showcased promising outcomes, particularly with the second framework, which demonstrated superior performance. By leveraging parallel convolution layers, this framework efficiently

captured deep features from each feature set, enhancing the classifiers' discriminative power. Notably, the extracted deep features not only effectively differentiated PD patients from healthy individuals but also significantly bolstered the overall classification efficacy.

Keywords— convolutional neural networks ,deep learning ,health informatics, Parkinson's Disease

classification, vocal features.

I. INTRODUCTION

In recent years, the utilization of health informatics systems has become increasingly prevalent in both the detection and ongoing monitoring of significant diseases. Specifically, within the realm of Parkinson's Disease (PD), which commonly afflicts individuals over the age of 60, there has been a notable reliance on information systems driven by artificial intelligence (AI). PD presents as a progressive neurodegenerative condition characterized by a myriad of motor and non-motor symptoms. Given the potential for prolonged patient lifespans with early diagnosis, the imperative for highly accurate and dependable health informatics systems in PD detection cannot be overstated. Moreover, such systems are designed not only to facilitate early detection but also to alleviate the burden on clinicians by streamlining the diagnostic process .

The focus of PD detection systems often centers on the identification and assessment of symptom severity using various assessment instruments. Among the most prevalent symptoms in early-stage PD is vocal impairment, with a significant proportion of patients experiencing vocal defects [3]–[7]. Consequently, recent studies in PD detection have placed considerable emphasis on health systems designed to detect vocal disorders. These studies leverage diverse speech signal processing techniques to extract clinically relevant features from speech data. These features are then inputted into a range of artificial learning algorithms, including but not limited to Artificial Neural Networks (ANN) and Support Vector Machines (SVM).

In the first framework, termed feature-level combination, our approach employs Convolutional Neural Networks (CNNs) to directly extract feature representations from the concatenation of various feature sets. This involves designing a Deep Neural Network (DNN) with multiple convolution layers, each utilizing different kernel sizes. Through convolution and pooling operations within these

layers, hierarchical abstract feature representations are generated. These representations are subsequently passed through fully connected layers to perform the classification task.

The second framework, termed model-level combination, diverges from the first in terms of how feature sets are combined. Here, the network comprises multiple parallel convolution layers, each dedicated to a distinct type of feature. During training, convolution operations are simultaneously conducted across these parallel layers, and the resulting representations are concatenated. Similar to the first framework, these concatenated feature representations then proceed through consecutive convolution and fully connected layers to execute the classification process. Rather than analyzing individual feature types' effects on PD classification, both frameworks aim to capture the contributions of each feature type through feature-level and model-level integration approaches.

II. LITERATURE SURVEY

In the realm of Parkinson's Disease (PD) classification, the efficacy of classification studies hinges directly upon the careful selection of relevant feature extraction techniques and appropriate machine learning methods. In the literature, numerous studies have made use of publicly available datasets, such as the one comprising 31 instances (23 PD patients and 8 healthy individuals) with 195 sound recordings [17].

Another notable PD dataset [4] encompasses 40 examples, with 20 PD patients and 20 healthy individuals, each having multiple speech recordings. These datasets commonly feature extracted attributes including vocal fundamental frequency, measures of variation in fundamental frequency, and amplitude variation metrics. As these datasets serve as the foundation for many PD detection investigations, the features extracted from them are often regarded as baseline features

Patients with Parkinson's Disease (PD) commonly experience vocal impairments, which manifest as alterations in vocal loudness, stability, and frequency patterns. Additional symptoms include voice breaks and diminished vocal quality. Speech processing techniques are frequently employed to identify abnormalities in speech patterns, making them preferred tools for automated extraction of PD-related vocal features. Over the past decade, numerous machine learning-based studies have focused on PD detection using vocal features. For

instance, Tsanas et al. [18] introduced a novel PD detection model leveraging vocal features. They employed various feature selection techniques, including Least Absolute Shrinkage and Selection Operator (LASSO), Minimum Redundancy Maximum Relevance (mRmR), Relief, and Local Learning-Based Feature Selection (LLBFS), to identify the top 10 most relevant features for their model. These features primarily originated from shimmer, Harmonic-to-Noise Ratio (HNR), and vocal fold excitation measurements. Performance evaluation utilizing Random Forest (RF) and Support Vector Machines (SVM) classifiers yielded precision rates of up to 98.6%. Notably, their study determined that the feature set yielding the lowest classification error was derived from the Relief selection method.

As highlighted in the study by [16], the presence of multiple voice recordings per individual within available datasets poses a challenge, potentially leading to biased results during performance evaluation, particularly when employing Cross-Validation (CV). Furthermore, utilizing the same individual's recordings for both training and testing steps in CV can exacerbate this bias. Given that our dataset comprises multiple voice recordings per healthy individuals and PD patients, we have opted to employ the Leave-One-Person-Out Cross Validation (LOPO CV) procedure for performance evaluation of the proposed frameworks. In each iteration of LOPO CV, instances corresponding to one individual are withheld as a test set, while the remaining instances from other individuals are utilized as the training set. This approach ensures a rigorous evaluation by simulating real-world scenarios where the model encounters data from unseen individuals during testing.

III. METHODOLOGY

Convolutional Neural Networks (CNNs) are structured with multiple layers where convolution operations are performed. A fundamental distinction between Artificial Neural Networks (ANNs) and CNNs lies in the connectivity pattern within successive layers. In CNNs, each local segment, known as a receptive field, of the input is connected to only one neuron, whereas in ANNs, inputs are fully connected to neurons in the subsequent layer. Within each layer of a CNN, convolution operations are executed by applying filters of various sizes to inputs. Following convolution, the outputs are passed through an activation function, and then pooling layers are employed for sub-sampling from the activated outputs. Through pooling, the dimensions of input data can be automatically reduced by the network.

CNNs exhibit notable attributes including resistance to location variance and compositionality. This resilience stems from the fact that CNNs' trained filters scan across all input data, allowing them to detect patterns irrespective of their precise location. Pooling plays a key role in achieving location invariance by addressing rotation and scaling in the input data. Consequently, CNN filters excel at transforming low-level features acquired from receptive fields into high-level feature representations in deeper layers, thereby preserving the network's compositionality property.

A fundamental distinction between Artificial Neural Networks (ANNs) and CNNs lies in the connectivity pattern within successive layers. In CNNs, each local segment, known as a receptive field, of the input is connected to only one neuron, whereas in ANNs, inputs are fully connected to neurons in the subsequent layer. Within each layer of a CNN, convolution operations are executed by applying filters of various sizes to inputs. Following convolution, the outputs are passed through an activation function, and then pooling layers are employed for sub-sampling from the activated outputs. Through pooling, the dimensions of input data can be automatically reduced by the network.

Additionally, implement clustering algorithms like k-means to group similar cases together. Adapt retrieved cases to align with the specific context and requirements of the current investment scenario through a case adaptation mechanism..

CNNs COME WITH SEVERAL HYPERPARAMETERS, INCLUDING FILTER SIZE, STRIDE, AND POOLING TYPE. THE FILTER SIZE DETERMINES THE LENGTH OF THE SLIDING WINDOW USED IN CONVOLUTION OPERATIONS, WHICH CAN BE APPLIED EITHER TO EACH ELEMENT OF THE INPUT DATA OR TO SPECIFIC REGIONS. STRIDE INDICATES THE NUMBER OF STEPS TAKEN DURING EACH WINDOW SLIDING. POOLING TYPE SPECIFIES WHETHER THE POOLING PROCESS WILL BE APPLIED TO EACH FILTER MAP INDIVIDUALLY (OUTPUTS OF THE FILTERING PROCESS) OR GLOBALLY TO THE FEATURE MAPS RESULTING FROM MULTIPLE FILTERS.

IN OUR STUDY, CNN SERVES AS THE CLASSIFICATION FRAMEWORK [29], [32]. THIS FRAMEWORK OPERATES AS AN END-TO-END NEURAL NETWORK, WHERE THE INPUT DATA CONSISTS OF THE PD DATA TO BE PREDICTED, AND THE OUTPUT PROVIDES ITS CORRESPONDING LABEL. AS DISCUSSED IN SECTION 3, OUR DATASET COMPRISES VARIOUS TYPES OF FEATURE SETS. TO EXPLORE THE IMPACT OF EACH FEATURE TYPE ON THE CLASSIFICATION

PROCESS, WE COMBINE THESE FEATURE SETS TO SERVE AS INPUT DATA FOR THE PROPOSED CLASSIFIERS. BASED ON DIFFERENT FEATURE COMBINATION SCHEMES, WE DESIGN DISTINCT CLASSIFICATION FRAMEWORKS WITH VARYING INPUT LAYERS.

OUR FIRST FRAMEWORK ENTAILS A 9-LAYERED CNN ARCHITECTURE, FEATURING 1 INPUT LAYER, 6 CONVOLUTION LAYERS (EACH COMPRISING TWO CONVOLUTION LAYERS FOLLOWED BY MAX POOLING OPERATION), 1 FULLY-CONNECTED (DENSE) LAYER, AND 1 OUTPUT LAYER. THIS FRAMEWORK, DEPICTED IN FIG. 1, IS TERMED AS FEATURE-LEVEL COMBINATION SINCE DIFFERENT TYPES OF FEATURE SETS ARE CONCATENATED BEFORE BEING FED INTO THE INPUT LAYER.

V. CONCLUSION

In conclusion, deep learning methodologies offer a promising avenue for the classification of Parkinson's Disease (PD) using vocal feature sets. Leveraging Convolutional Neural Networks (CNNs) in particular, researchers have made significant strides in this domain by exploiting the unique characteristics of PD-related vocal impairments. The utilization of vocal features such as fundamental frequency, amplitude variations, and vocal stability has proven instrumental in developing accurate and reliable classification models. By harnessing the power of deep learning, these models can discern subtle patterns and nuances within vocal data, thereby facilitating early detection and monitoring of PD.

One notable aspect of deep learning-based PD classification is its ability to overcome the challenges posed by unbalanced class distributions within datasets. Through innovative techniques such as Leave-One-Person-Out Cross Validation (LOPO CV), researchers can mitigate bias and ensure robust performance evaluation. Moreover, the integration of various feature sets, including those derived from signal processing techniques like Mel-Frequency Cepstral Coefficients (MFCC) and Tunable Q-factor Wavelet Transform (TQWT), enhances the discriminative power of classification models, enabling more accurate predictions.

The development of end-to-end neural network architectures tailored for PD classification represents a significant advancement in the field. These frameworks, comprising multiple layers of convolution, pooling, and

fully connected layers, effectively capture complex relationships within vocal data. By exploring different feature combination schemes, researchers can gain insights into the contributions of individual feature types to the classification process. This not only enhances our understanding of PD-related vocal impairments but also facilitates the refinement of classification models for improved performance.

In summary, deep learning-based approaches hold immense potential for advancing the field of PD classification using vocal feature sets. By leveraging CNNs and innovative feature combination strategies, researchers can develop highly accurate and robust classification models capable of early detection and monitoring of PD. With further refinement and validation, these models have the potential to revolutionize clinical practices by offering efficient and reliable tools for PD diagnosis and management.

VI. REFERENCES

- [1] Xie, H., et al. (2019). Application of Case-Based Reasoning in Renewable Energy Investment Decision Making. *IEEE Access*, 7, 11408-11418.
- [2] Jha, S. K., et al. (2020). Comparative Analysis of Decision Support Systems for Renewable Energy Investment. *International Journal of Renewable Energy Research and Applications*, 10(3), 197-210.
- [3] Foxon, T. J., et al. (2018). Economic Principles and Methodologies for Renewable Energy Investment Analysis. *Renewable and Sustainable Energy Reviews*, 82, 2157-2167.
- [4] Zhang, L., et al. (2021). Financial Modeling and Risk Assessment in Renewable Energy Investment. *Energy Economics*, 98, 105424.
- [5] de la Vega, M. A., et al. (2017). Decision Support System for Renewable Energy Project Selection: A Multi-Criteria Approach. *Sustainability*, 9(9), 1546.
- [6] Chicco, F., et al. (2019). Decision Support Systems for Renewable Energy Integration: A Review of State-of-the-Art. *Renewable and Sustainable Energy Reviews*, 109, 76-87.
- [7] Sovacool, J., et al. (2020). Ethical Implications of Renewable Energy Transitions: A Focus on Social Justice and Equity. *Energy Research & Social Science*, 66, 101480.

[8] J amasb, A., et al. (2018). Regulatory Frameworks and Policy Instruments for Ethical Renewable Energy Investments. *Energy Policy*, 118, 419-426.

[9] Lin, J., et al. (2019). Artificial Intelligence Techniques for Renewable Energy Forecasting and Optimization. *IEEE Transactions on Sustainable Energy*, 10(2), 621-628.

