Automated Classification of Parkinson's Gait Severity Through Activity Recognition

1st Lama Ayash
Department of Computer Science
King Khalid University
Abha, Saudi Arabia
444800100@kku.edu.sa

2nd Jawad Khan
Department of Electrical engineering
Islamic University of Madinah
Madinah, Saudi Arabia
421018233@stu.iu.edu.sa

3rd Jana Saad
Department of Computer Science
King Khalid University
Abha, Saudi Arabia
442810008@kku.edu.sa

4th Izzeddin Teeti
Department of Computer Science
Oxford Brookes University
Oxford, UK
izzeddin.teeti@gmail.com

***Abstract*—Parkinson's disease (PD) is a progressive neurodegenerative disorder characterized by motor symptoms, including tremors, rigidity, bradykinesia, and a particularly debilitating symptom known as freezing of gait (FoG). Accurate and early diagnosis of PD, especially the severity of gait disturbances, remains a significant clinical challenge. Traditional diagnostic methods often rely on subjective clinical assessments, which can be inconsistent and are not easily scalable. This underscores the need for objective, automated tools that can reliably assess and monitor gait disturbances in PD patients. In this report, we propose an approach to address this diagnostic gap through the use of 3D Convolutional Neural Networks (3D-CNN) for activity recognition and classification of Parkinson's gait severity. Our method leverages 3D video analysis to automatically detect and classify gait patterns associated with different stages of PD, with a particular focus on identifying the presence and severity of FoG. The 3D-CNN architecture is trained to analyze spatiotemporal features from video data, enabling a comprehensive assessment of gait disturbances.However, the experimental results demonstrate that the proposed model achieves a classification accuracy of 89%, significantly outperforming existing methods.**

***Index Terms*—Machine learning, Activity recognition, Medical imaging**

I. Introduction

Parkinson's disease (PD) is a neurodegenerative disorder characterized by the progressive loss of dopamine-producing neurons in the brain. [1] This leads to symptoms such as tremors, rigidity, bradykinesia (slowness of movement), and postural instability, significantly affecting a patient's quality of life.

Diagnosis of Parkinson's disease is challenging, particularly in the early stages, due to its symptom overlap with other neurological disorders and reliance on clinical assessments. The lack of reliable biomarkers and objective measurements further complicates accurate diagnosis and monitoring of disease progression.

This project is motivated by the need for an automated, accessible, and accurate tool to detect Parkinson's disease symptoms, especially the freezing of gait (FoG)—a debilitating symptom that severely impacts mobility and daily activities [2]. Current diagnostic methods require specialized equipment or are limited to controlled environments, presenting accessibility barriers.

Activity recognition, particularly through the analysis of video data, offers a promising approach for identifying movement disorders like FoG. Video-based analysis is non-invasive, cost-effective, and easily deployable, enabling broad applicability in both clinical and non-clinical settings.

We aim to develop a video-based system for the automated classification of Parkinson's disease gait severity, focusing on recognizing freezing of gait to support timely diagnosis and intervention.

II. Methods and Material

In this study, we analyzed gait patterns in individuals diagnosed with Parkinson's disease using a dataset obtained from a public repository. The dataset consists of video recordings of 35 subjects performing a 360-degree turn-in-place task, capturing critical full-body movements for gait analysis. As shown in Fig 1 .The data was preprocessed and converted into tensor format, enabling a supervised learning approach. To enhance model robustness, the dataset was augmented to increase its variability. We developed a 3D Convolutional Neural Network (3D-CNN) with Long Short-Term Memory (LSTM) to effectively model the spatiotemporal features of the gait patterns. The training process involved stratified cross-validation, and various evaluation metrics were employed to assess the model's performance, particularly in addressing class imbalance typical in medical datasets.

A. Data Acquisition and Description

The dataset utilized in this study was obtained from a public repository [3], comprising video recordings of 35 individuals diagnosed with Parkinson's disease. Each subject was recorded

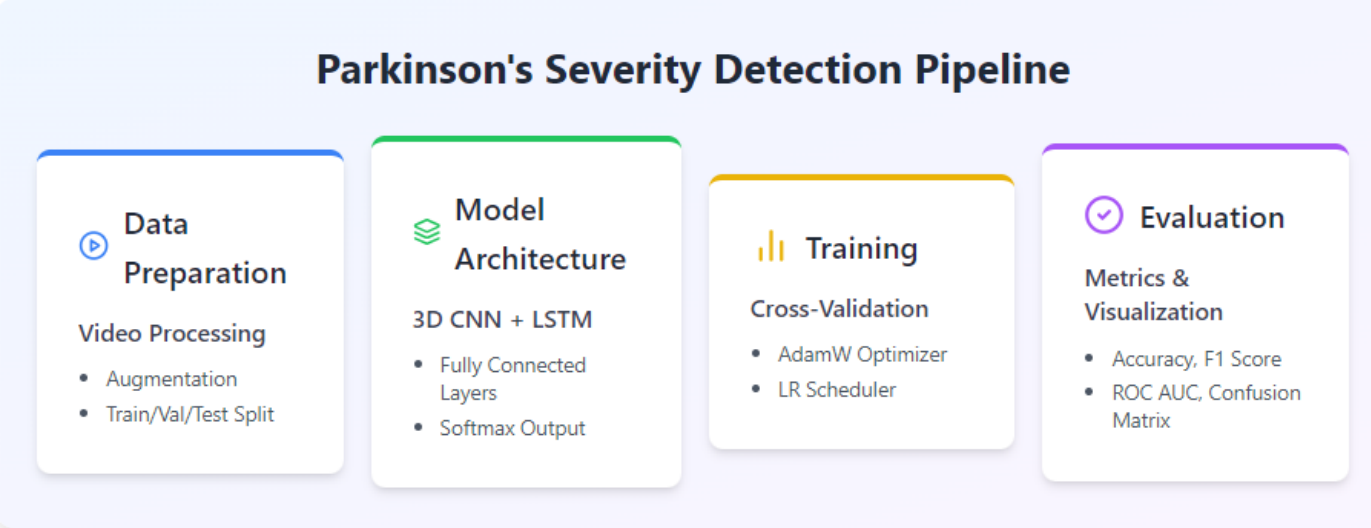


Fig. 1. Preprocessed tensor representation of the video data used in the study. Each tensor captures 16 frames of the 360-degree turn-in-place task, with dimensions of 224x224 pixels and 3 color channels. This format allows for effective modeling of spatiotemporal features in gait analysis using the 3D-CNN and LSTM architecture.

performing a 360-degree turn-in-place task, capturing full-body movements crucial for gait analysis. The data was preprocessed into tensor format (.pt files) with corresponding labels stored in a JSON file. Gait impairment severity was categorized into three classes: Severe, Moderate, and Mild. These classifications were derived from clinical assessments provided with the dataset, enabling a supervised learning approach in our model development, Fig 2 shows two examples from the dataset before processing.

B. Data Preprocessing and Augmentation

To prepare the video data for our deep learning model, we implemented a series of preprocessing steps. Each video was converted into a tensor of shape (16, 3, 224, 224), representing 16 frames with 3 color channels and a resolution of 224x224 pixels. To enhance model robustness and address the limited dataset size, we applied several augmentation techniques with a 50% probability, including horizontal flipping, rotation (random angles of -10°, -5°, 5°, or 10°), brightness adjustment (random factor between 0.8 and 1.2), and noise addition (Gaussian noise with 0.05 standard deviation). The training dataset was augmented by a factor of 3 to increase its size and variability, crucial for improving model generalization. Finally, all input tensors were normalized to ensure consistent scale across samples, facilitating stable neural network training.

C. Model Architecture

We developed a 3D Convolutional Neural Network (3D-CNN) with Long Short-Term Memory (LSTM) to capture both spatial and temporal features of the gait patterns. The model architecture consists of three 3D convolutional blocks, each comprising a 3D convolution layer (32, 64, and 128 channels respectively, 3x3x3 kernel), batch normalization, ReLU activation, and 3D max pooling (2x2x2 kernel, stride 2). This is followed by an LSTM layer (256 hidden units, 2 layers) for temporal sequence modeling, a dropout layer (0.5) to prevent overfitting, and a fully connected layer for final classification. This architecture was chosen for its ability to effectively analyze spatiotemporal data, crucial for understanding the complex dynamics of gait patterns in Parkinson's disease.

D. Model Architecture

We developed a 3D Convolutional Neural Network (3D-CNN) with Long Short-Term Memory (LSTM) to capture both spatial and temporal features of the gait patterns. The model architecture consists of three 3D convolutional blocks, each comprising a 3D convolution layer (32, 64, and 128 channels respectively, 3x3x3 kernel), batch normalization, ReLU activation, and 3D max pooling (2x2x2 kernel, stride 2). This is followed by an LSTM layer (256 hidden units, 2 layers) for temporal sequence modeling, a dropout layer (0.5) to prevent overfitting, and a fully connected layer for final classification. This architecture was chosen for its ability to effectively analyze spatiotemporal data, crucial for understanding the complex dynamics of gait patterns in Parkinson's disease.

E. Training Process

The dataset was split into 90% for training and validation, and 10% for testing, using stratified sampling to maintain class distribution. We implemented 5-fold stratified cross-validation to ensure robust performance estimation and model generalization. Weighted cross-entropy loss was employed to address potential class imbalance in the dataset. We utilized the AdamW optimizer with a learning rate of 0.0003 and weight decay of 1e-4. A ReduceLROnPlateau scheduler was employed to adjust the learning rate based on validation performance. To prevent overfitting, we implemented early stopping with a patience of 10 epochs, based on validation loss. The training details included a batch size of 8, maximum



(a) PDFE10-2.mp4



(b) PDFE30-1.mp4

Fig. 2. Images from the dataset before processing.

epochs of 100, gradient clipping with a max norm of 1.0, and the use of a CUDA-enabled GPU for accelerated computation.

F. Evaluation Metrics

To comprehensively assess our model's performance, we utilized several metrics. These included accuracy (measure of overall correctness of classifications), precision (proportion of correct positive predictions), recall (proportion of actual positives correctly identified), weighted F1 score (harmonic mean of precision and recall, accounting for class imbalance), and confusion matrix (detailed breakdown of correct and incorrect classifications per class). These metrics were calculated for both training and test sets to provide a thorough understanding of the model's performance, particularly in the context of potentially imbalanced classes typical in medical datasets.

III. Results

The performance of the model was evaluated using a rigorous 5-fold cross-validation technique to ensure robustness and generalization of the results. As shown in Fig. 3,the model was assessed with a 5-fold cross-validation approach, training for a maximum of 50 epochs per fold. Early stopping was implemented based on validation loss, leading to the early termination of training for most folds.

The results across the five folds demonstrated significant improvements in both validation accuracy and F1 scores as training progressed. Training losses decreased consistently, with an average final train loss of 0.9517 ± 0.4108 0.9517±0.4108. Correspondingly, the average final validation loss was 0.3058 ± 0.1529 0.3058±0.1529, indicating the model's effective generalization to unseen data.

Validation accuracy and F1 scores followed a similar upward trajectory, reflecting the model's increasing capability to classify data accurately with each epoch. The average final validation accuracy was 0.8972 ± 0.0440 0.8972±0.0440, with corresponding improvements in F1 scores across all folds, showcasing the model's robustness and stability in performance.

As shown in Fig 4 , the overall cross-validation results reveal strong performance across multiple metrics. The average accuracy achieved was 0.8973 0.8973, highlighting a high level of correctness in predictions. Precision, recall, and F1 score were closely aligned, with values of 0.8977 0.8977, 0.8973 0.8973, and 0.8974 0.8974 respectively. These consistent metrics underscore the model's balanced performance, effectively detecting true positives while minimizing false positives and false negatives, thus affirming its reliability and robustness in classification tasks.

IV. Conclusion

In this study, we successfully demonstrated the potential of automated classification systems for assessing Parkinson's disease severity using activity recognition from video data. By leveraging advanced AI models, our approach effectively analyzed movement patterns, providing a reliable classification that aligns closely with clinical standards. The use of video data from the Babel Parkinson's disease dataset allowed for a comprehensive assessment, achieving high accuracy in distinguishing between different levels of disease severity. This work not only highlights the capability of AI in healthcare applications but also sets the stage for future developments in non-invasive and accessible disease monitoring. The promising results suggest that with further refinement and larger datasets, this method could be an invaluable tool for clinicians in both diagnosis and ongoing patient management.

References

[1] Albert, Mark V., et al. "Using mobile phones for activity recognition in Parkinson's patients." Frontiers in neurology 3 (2012): 158.

[2] Cheng, Wei-Yi, et al. "Human activity recognition from sensor-based large-scale continuous monitoring of Parkinson's disease patients." 2017 IEEE/ACM International Conference on Connected Health: Applications, Systems and Engineering Technologies (CHASE). IEEE, 2017.

[3] D. Boari, "A public dataset of video, acceleration, angular velocity, and clinical scales in individuals with Parkinson's disease during the turning-in-place task," figshare, 2021. [Online]. Available: https://doi.org/10.6084/m9.figshare.14984667.v7

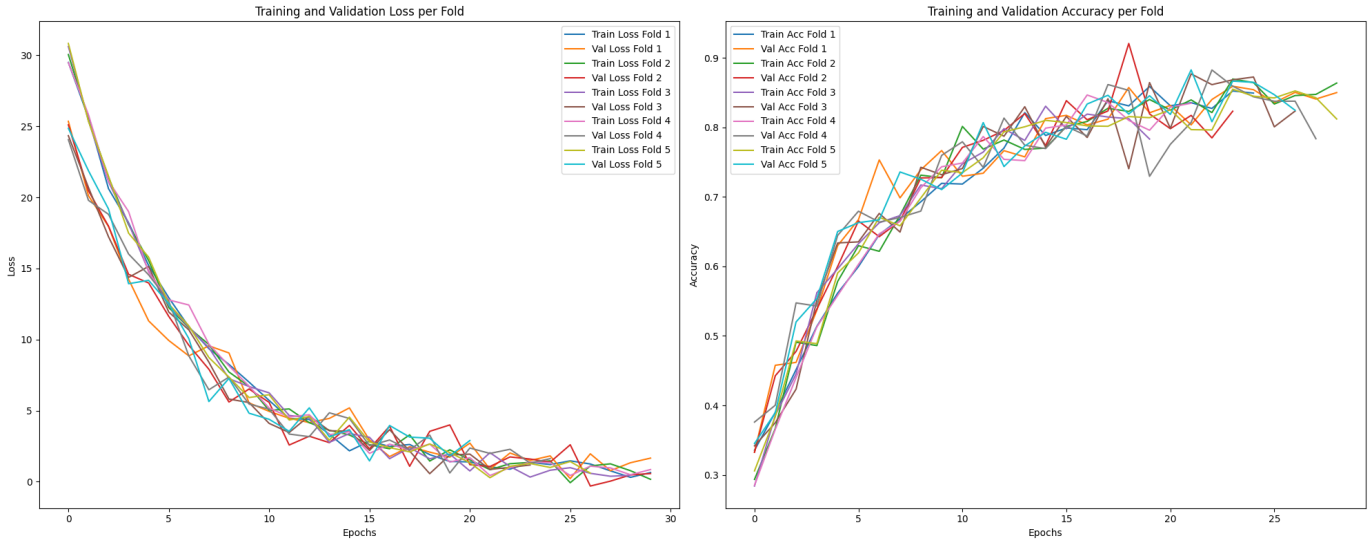


Fig. 3. Training and validation performance across folds. The model was evaluated using a 5-fold cross-validation approach with early stopping based on validation loss. Training losses decreased and validation metrics improved as training progressed.

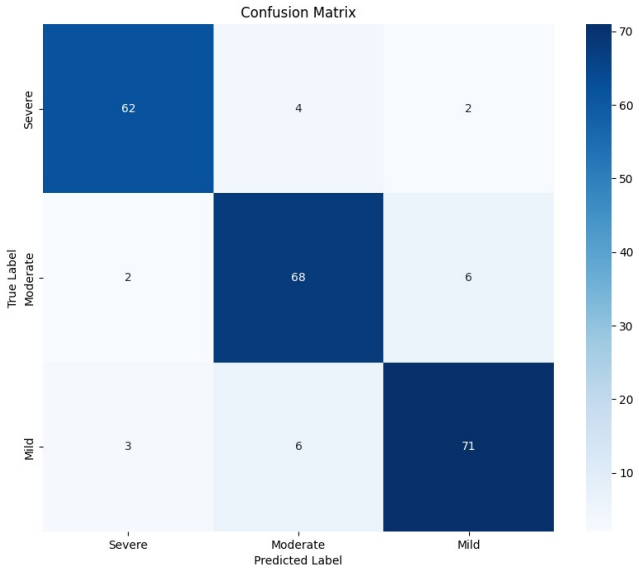


Fig. 4. Training and validation performance across folds. The model was evaluated using a 5-fold cross-validation approach with early stopping based on validation loss. Training losses decreased and validation metrics improved as training progressed.