EEG Epoch Classification with Fourier Transform and LLM Transformer: Insights into Parkinson disease

Caleb Cooper[1], Raven Lee[2], Manliang Feng[3], Jinghe Mao[1]

[1]Department of Biology, Tougaloo College, Tougaloo, MS, USA
[2]Department of Mathematics, Tougaloo College, Tougaloo, MS, USA
[3]Department of Chemistry, Tougaloo College, Tougaloo, MS, USA

Electroencephalography (EEG) biomarkers hold promise for advancing the diagnosis and understanding of neurological conditions, including Parkinson's disease. This study integrates advanced signal processing techniques, including Fast Fourier Transform (FFT), Short-Time Fourier Transform (STFT), and Wavelet Transform, with Transformer Encoder based supervised learning to classify event-related neural activity from EEG data. We hypothesize that combining frequency-domain transformations with time-domain features can enhance biomarker detection, particularly in 1–3 second epoch segments.

To address EEG's susceptibility to noise and artifacts, preprocessing steps such as Independent Component Analysis (ICA) and wavelet decomposition are employed to reduce noise from muscle movements, eye blinks, and external interference. Fixed-length epochs (1–3 seconds) are used, with overlapping windows to minimize information loss and improve the capture of event-related potentials (ERPs). The extracted features include both frequency-domain characteristics from FFT and time-domain metrics such as mean and variance, supplemented by STFT and Wavelet Transform to capture time-frequency dynamics. Labeled EEG data from open-source datasets (e.g., PhysioNet, BCI Competition) and those labeled using Fourier Transform methods provide the ground truth for training a Transformer encoder-based model, fine-tuned for time-series analysis. The model incorporates positional encodings to handle sequential dependencies in EEG data while balancing computational efficiency with pre-trained architectures. Performance is evaluated through cross-subject and cross-event validation using accuracy, precision, recall, and F1-score to ensure generalizability.

Preliminary results demonstrate that FFT effectively identifies frequency-domain biomarkers linked to Parkinson's disease, while STFT and Wavelet Transform offer complementary strengths in detecting transient features. The inclusion of time-domain metrics further improves classification accuracy. This research highlights the synergy between advanced computational methods and neuroscience applications, paving the way for more precise diagnostic tools. Future work will validate findings across diverse datasets, explore semi-supervised learning for labeling efficiency, and expand applications to other neurological conditions, contributing to the broader field of neuroscience diagnostics.

Acknowledgement: This material is based on work supported by the Department of the Air Force under the Air Force Contract No. FA955023D0001. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Department of the Air Force.