A Project Report

On

GENERATION OF BRAIN GRAPH USING EEG DATA

BY

DEEPESH BHARDWAJ

SE21UCSE049

&

GNAN REDDY B.

SE21UCSE064

Under the supervision of

SHABNAM SAMIMA

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR

DEGREE OF BACHELOR OF TECHNOLOGY

PR 4101: PROJECT TYPE COURSE

ECOLE CENTRALE SCHOOL OF ENGINEERING

MAHINDRA UNIVERSITY

HYDERABAD

(DEC 2024)

ACKNOWLEDGMENTS

We would like to express our sincere gratitude to our project supervisor, Prof. SHABNAM SAMIMA, for her invaluable guidance, continuous support, and insightful suggestions throughout this project. Her expertise in the field of brain-computer interfaces and signal processing has been instrumental in shaping our research approach and methodology.

We also thank the faculty and staff of Ecole Centrale School of Engineering, Mahindra University, for providing the necessary resources and infrastructure that made this research possible.

Ecole Centrale School of Engineering

Mahindra University

Hyderabad

Certificate

This is to certify that the project report entitled "**GENERATION OF BRAIN GRAPH USING EEG DATA**" submitted by Mr. DEEPESH BHARDWAJ (Roll No. SE21UCSE049) and Mr. GNAN REDDY B. (Roll No. SE21UCSE064) in partial fulfillment of the requirements of the course PR 4101, Project Course, embodies the work done by them under my supervision and guidance.

SHABNAM SAMIMA

Ecole Centrale School of Engineering, Hyderabad.

Date: 23/12/2024

ABSTRACT

This research presents the development and implementation of an automated system for detecting driver fatigue using electroencephalography (EEG) data. The study analyzed a comprehensive dataset comprising 62 EEG recording sessions from 27 participants engaged in sustained-attention driving tasks. The system processes multi-channel EEG signals and reaction time data to classify driver fatigue levels in real-time. Using advanced signal processing techniques and machine learning algorithms, we achieved a classification accuracy of 95%, demonstrating the system's reliability in distinguishing between low, medium, and high fatigue states. The implemented solution offers potential applications in enhancing road safety through early detection of driver fatigue.

CONTENTS

Title page……………………………………………………….…….1

Acknowledgements……………………………………………………..2

Certificate…………………………………………………………......3

Abstract………………………………………………………….....4

1. Introduction…………………………………………….…………6

2. Problem Definition.……………………………………………….7

3. Background and Related
Work………..…………………………………………………….....8

4. Implementation…………………………………………………….9

5. Results……………………………………………………………..11

Conclusion…………………………………………………………13

References…………………………………………………………14

1. INTRODUCTION

1.1 Background

Driver fatigue is a critical safety concern that significantly contributes to road accidents worldwide. According to recent statistics, fatigue-related accidents account for approximately 20% of road fatalities. The development of reliable fatigue detection systems is therefore crucial for improving road safety.

1.2 Research Objectives

The primary objectives of this research are:

1. To develop an automated system for processing and analyzing EEG data for fatigue detection

2. To implement a real-time pipeline for continuous monitoring of driver fatigue levels

3. To validate the system's effectiveness through comprehensive testing and analysis

4. To create a practical solution that can be integrated into existing vehicle safety systems

1.3 Scope

This project focuses on:

- Analysis of 32-channel EEG data

- Processing of event-related potentials

- Implementation of machine learning algorithms for fatigue classification

- Development of a real-time monitoring system

2. PROBLEM DEFINITION

2.1 Research Questions

The project addresses the following key questions:

1. How can EEG data be effectively processed to detect different levels of driver fatigue?

2. What features from EEG signals are most indicative of fatigue states?

3. How can reaction time data be integrated with EEG analysis for improved detection accuracy?

4. What is the optimal approach for implementing real-time fatigue monitoring?

2.2 Mathematical Representation

The fatigue detection problem can be formulated as a multi-class classification problem:

- Input: EEG signals $X \in R^{(32\times T)}$, where T is the time duration

- Output: Fatigue level $y \in \{0,1,2\}$ representing low, medium, and high fatigue states

- Feature space: Combination of EEG band powers and reaction time metrics

3. BACKGROUND AND RELATED WORK

3.1 Literature Review

Recent research in EEG-based fatigue detection has shown promising results in early identification of driver drowsiness. Key works include:

1. Cao et al. (2019) demonstrated the feasibility of using multi-channel EEG recordings during sustained-attention driving tasks

2. Lin et al. (2014) developed wireless and wearable EEG systems for evaluating driver vigilance

3. Chuang et al. (2018) investigated brain electrodynamic signatures against fatigue during driving

3.2 Theoretical Framework

The research builds upon established theories in:

- EEG signal processing and analysis

- Machine learning for biosignal classification

- Human factors in driving performance

- Cognitive load assessment

4. IMPLEMENTATION

4.1 Dataset Description

The implementation utilized:

- 62 EEG recording sessions

- 27 participants

- 32 EEG channels

- Total of 81,576 events

- Left and right lane departure events

- Response onset times

4.2 Data Processing Pipeline

The implementation followed these key steps:

A. Dataset Splitting and Preprocessing:

This Python function, `calculate_baselines_and_normalize`, performs baseline calculation and normalization on reaction times from multiple participants. Here's a summary of its functionality:

1. **Input**: A list of dictionaries (`reaction_times`), where each dictionary represents a participant and contains:
    - `'left_reaction_times'`: List of reaction times for left-hand responses.
    - `'right_reaction_times'`: List of reaction times for right-hand responses.
    - `'file'`: Identifier for the participant.
2. **Processing**:
    - For each participant:
        - Compute the **baseline** reaction time for the left and right hands using the mean of the first 10 values of the respective reaction time arrays.
        - **Normalize** the reaction times by dividing all values by their respective baseline (left and right).

3. **Output**:
    - A list of **baselines** for all participants, containing the computed left and right baselines.
    - A list of **normalized data** for all participants, containing the normalized reaction times for left and right responses.
4. **Return**:
    - The function returns two lists: one for baselines and another for normalized data. Each list item corresponds to a participant's data.

This function is useful for standardizing reaction time measurements across participants by normalizing them relative to their initial baseline performance.

B. Feature Extraction:

This function, `extract_eeg_features`, processes EEG data to extract power features for specified frequency bands:

1. **Input**:
    - `file_path`: Path to the EEG data file in EEGLAB format.
    - `ebands`: A dictionary defining frequency bands (e.g., alpha, beta) with their min and max frequencies.
    - `sfreq`: Sampling frequency (not used explicitly in the code).
2. **Steps**:
    - Loads the EEG data using MNE and preloads it for processing.
    - Applies a bandpass filter to retain frequencies between 0.5 and 30 Hz.
    - Computes the power spectral density (PSD) using Welch's method.
    - For each frequency band in `ebands`, calculates the band power using the `compute_band_power` function and stores it in a dictionary.
3. **Output**:
    - Returns a dictionary of extracted band power features for each specified frequency band.

4.3 Model Development

The fatigue classification model uses:

- Gradient Boosting Classifier

- Input features: EEG band powers and reaction times

- Three-level classification scheme

- Cross-validation for model evaluation

5. RESULTS

5.1 Model Performance

The system achieved:

- Overall accuracy: 95%

- Precision: 97% (macro average)

- Recall: 93% (macro average)

- F1-Score: 95% (weighted average)

Confusion Matrix:

```
[[ 4 0 0]
 [ 0 4 1]
 [ 0 0 10]]
```

5.2 Feature Importance

Analysis revealed the following feature importance:

- Alpha band: 26%

- Theta band: 22%

- Beta band: 20%

- Delta band: 15%

- Reaction times: ~8% each

5.3 Real-time Performance

The system demonstrated:

- Continuous monitoring capability

- Quick response to fatigue onset

- Reliable state transitions

- Minimal processing delay

CONCLUSION

Summary

The implemented system successfully demonstrates:

- Accurate fatigue detection (95% accuracy)
- Real-time monitoring capability
- Robust feature extraction
- Practical applicability

Future Work

Potential areas for future development include:

- Integration with vehicle systems
- Mobile application development
- Enhanced feature extraction
- Multi-modal analysis

REFERENCES

1. Cao, Z., Chuang, M., King, J. T., & Lin, C. T. (2019). Multi-channel EEG recordings during a sustained-attention driving task. Scientific Data, 6(1), 19.

2. Lin, C. T., et al. (2014). Wireless and wearable EEG system for evaluating driver vigilance. IEEE Transactions on Biomedical Circuits and Systems, 8(2), 165-176.

3. Chuang, C. H., et al. (2018). Brain electrodynamic and hemodynamic signatures against fatigue during driving. Frontiers in Neuroscience, 12, 181.

4. Huang, R. S., Jung, T. P., & Makeig, S. (2009). Tonic changes in EEG power spectra during simulated driving. Lecture Notes in Computer Science, 5638, 394-403.

5. Liu, Y. T., et al. (2016). Brain dynamics in predicting driving fatigue using a recurrent self-evolving fuzzy neural network. IEEE Transactions on Neural Networks and Learning Systems, 27(2), 347-360.