मौलाना आजाद राष्ट्रीय प्रौद्योगिकी संस्थान - भोपाल

Maulana Azad National Institute of Technology– Bhopal

Session 24-25

MINOR PROJECT PRESENTATION

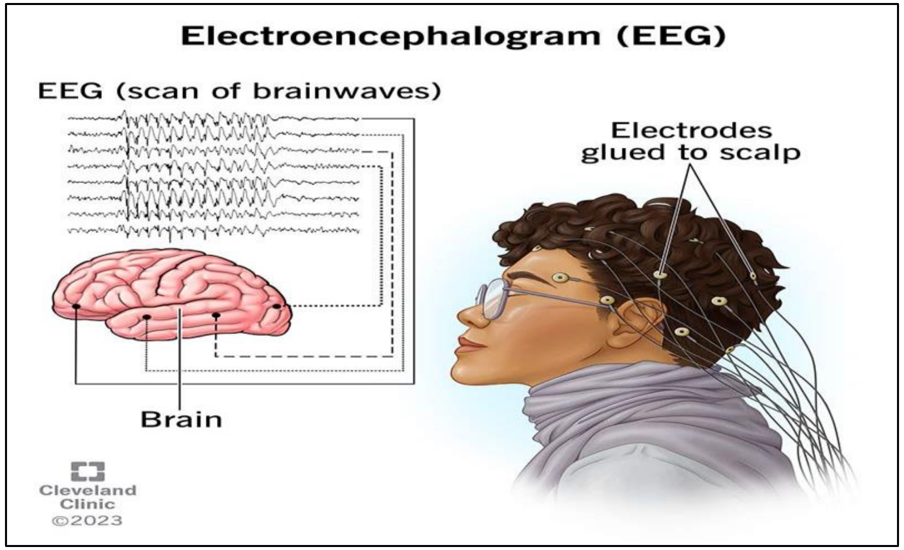
Title: Mental Workload Classification Using 1DCNN Model and Interpretation Through Explainable AI

NAME	SCHOLAR NO.
Harshit Singh	2211201225
Gaurav Barwal	2211201227
Dheeraj Ahirwar	2211201207
Sushant Singh	2211201213

GUIDED BY:
Dr. Mitul Kumar Ahirwal

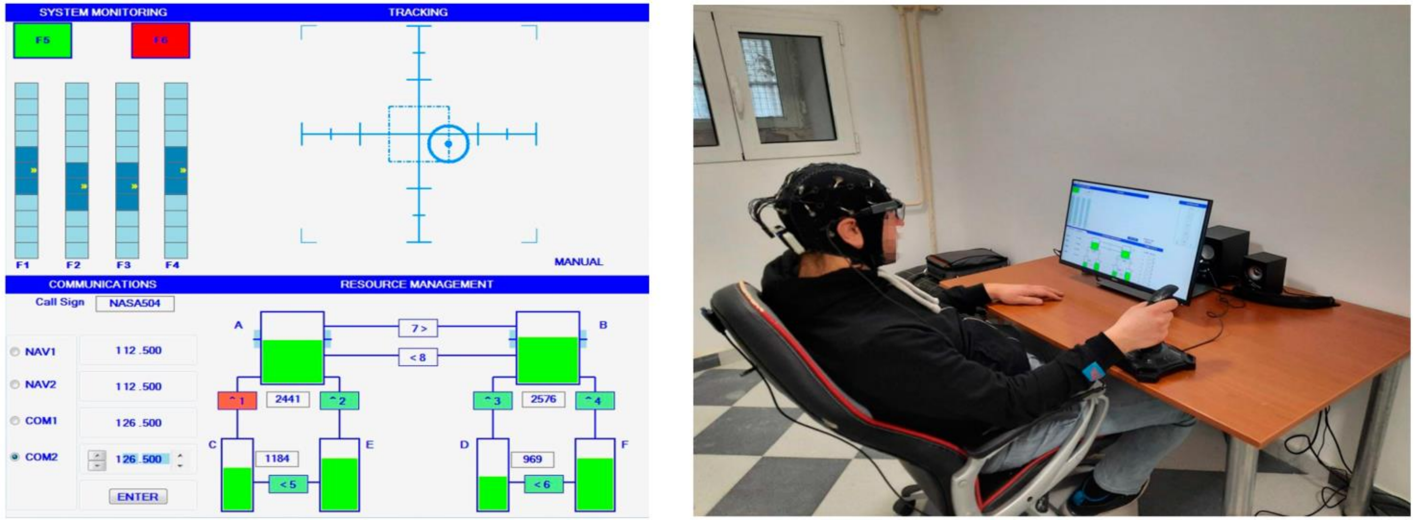
INTRODUCTION

- **Mental workload** reflects cognitive effort during tasks and affects performance in safety-critical fields (e.g., aviation, healthcare).
- **EEG (Electroencephalography)** provides real-time, non-invasive brain activity monitoring to assess mental workload.
- **Deep learning**, especially **CNNs**, enables automatic feature extraction from raw EEG signals without manual preprocessing.
- **Explainable AI (XAI)** techniques like **SHAP** and **LRP** help interpret CNN predictions, improving transparency and trust.
- This model aims to **classify workload levels (e.g., low, medium, high)** using EEG data and explain model decisions using SHAP and LRP.



MOTIVATION OF THE STUDY

- Understanding **mental workload** is key to improving **human-computer interaction** and decision-making.
- It plays a vital role in **multitasking environments**, where overload can reduce performance and increase errors.
- The **STEW EEG dataset** offers rich brainwave data under varying levels of cognitive demand..
- The goal is to enable **real-time monitoring** of workload to enhance **productivity**, **safety**, and **user well-being**.



(a) (b)

LITERATURE REVIEW

STEW: Simultaneous Task EEG Workload Dataset

- Contains raw EEG data from 48 subjects performing multitasking tests (SIMKAP test)
- Includes resting-state EEG recorded before the task
- Recorded using Emotive EPOC headset (14 channels, 128 Hz, ~2.5 minutes per session)
- Subjects rated their perceived mental workload on a 1–9 scale after each stage
- Available in .mat format (uploaded on Kaggle)
- EEG recorded during arithmetic tasks to identify mental workload levels
- Supervised classification problem using workload ratings (0–9 scale)
- Final dataset used: 45 subjects (subjects 5, 24, 42 excluded due to missing ratings)
- Each subject’s data: 14 channels × 19,200 samples (time domain signals)

PROJECT OBJECTIVES

Objective 1 :- Preprocess and analyze EEG data from the STEW dataset to prepare it for modeling.
Objective 2 :- Develop machine learning and deep learning models (e.g., CNN) to classify mental workload levels.

Specific Objectives:

•Evaluate model performance for both two-class and three-class classification tasks.

•Investigate the impact of data balancing techniques on classification accuracy.

•Compare results with existing studies to assess improvements and highlight contributions.

•Provide insights for real-time mental workload monitoring in multitasking environments.

•Compare results with existing studies to assess improvements and highlight contributions.

•Provide insights for real-time mental workload monitoring in multitasking environments.

DATASET DESCRIPTION

The dataset includes EEG recordings from 45 subjects, each with 14 channels and 19,200 data points per channel.

Each subject's data is stored as a 14 × 19,200 matrix, representing one sample in the database.

The dataset contains three key files

dataset.mat (EEG signals), class_012.mat (workload labels), and rating.mat (subject ratings).

The data is classified into two studies: a three-class classification and a two-class classification.

For three-class classification: Class 0 (normal workload, ratings 4–5), Class 1 (moderate workload, ratings 6–7), and Class 2 (high workload, ratings 8–9).

For two-class classification: Class 0 (normal workload, ratings 4–6) and Class 1 (high workload, ratings 7–9).

INTRODUCTION TO XAI

- XAI refers to AI systems that provide **transparent and interpretable** explanations for their decisions.
- Unlike traditional **"black-box" models**, XAI reveals **how** and **why** a model makes specific predictions.
- It helps users **understand**, **trust**, and **validate** AI outputs.
- Enhances **accountability** and **fairness** in AI-driven systems.
- Crucial for sensitive applications like **healthcare**, **finance**, and **cognitive workload analysis**.

USE OF XAI IN OUR MODEL

- **Explainable AI (XAI)** enhances the **transparency** of the mental workload classification model.
- XAI helps identify **which EEG channels, time windows, or features** most impact the model’s decisions.
- It explains **why a subject** is classified into a specific workload level (normal, moderate, or high).
- Increases **trust and interpretability** for researchers, developers, and end-users.
- Aids in detecting **biases or errors** in the model's predictions.
- Improves the model’s **fairness, reliability, and acceptance** in real-world applications.

LRP VS SHAP :

LRP vs. SHAP:

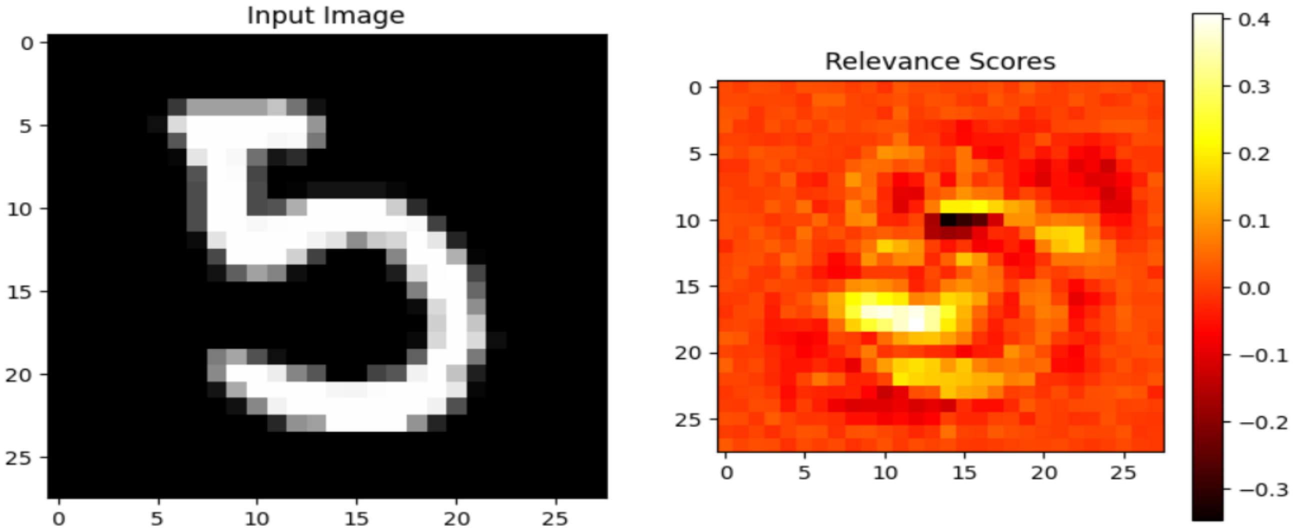
•**LRP (Layer-wise Relevance Propagation)** explains model predictions by backpropagating relevance scores through the layers, mainly used in neural networks.

•**SHAP (SHapley Additive exPlanations)** provides feature importance based on game theory, applicable across
• many model types.

•LRP focuses on layer-wise internal contributions, while SHAP gives a unified, model-agnostic explanation of
•each feature’s impact on predictions.

Implementation of LRP – MNIST Digit Classification

- A **3-layer fully connected neural network** is trained on the MNIST dataset (handwritten digit images).
- The model achieves around **97% classification accuracy** on the test set.
- **Layer-wise Relevance Propagation (LRP)** is applied to interpret the model's predictions.
- LRP **backpropagates relevance scores** from the output layer to the input pixels.
- It **highlights important pixels** in each digit image that influenced the model's decision.
- This improves the **transparency and trust** in model predictions by making them visually explainable.



ACCURACY OF THE MODEL

By adjusting certain parameters, we observed changes in the model's accuracy. Below is a summary of how different hyperparameters affect the model's performance:

- **Learning Rate**: Increasing the learning rate typically causes faster convergence but may decrease the final accuracy if set too high.
- **Dropout**: Increasing the dropout rate improves generalization but may reduce training accuracy.
- **Pooling Size**: Increasing the pooling size enhances translation invariance but reduces temporal precision.
- **Epochs**: Increasing the number of epochs boosts training accuracy but may lead to overfitting and reduced generalization if set too high.
- **Batch Size**: Increasing the batch size improves training stability but can reduce generalization if set too large.

ACCURACY PERFORMED AT DIFFERENT HYPERPARAMETERS

Table 1

Trials	Train Accuracy	Validation Accuracy	Test Accuracy	Precision	Recall	F1-S
1	39.80	34.20	38.90	37.60	30.10	34.80
2	30.90	31.40	35.20	33.80	30.50	31.30
3	38.50	39.00	36.70	39.20	35.60	36.30
4	32.30	30.20	34.10	35.90	30.90	32.10
5	37.10	32.60	39.80	38.90	36.40	37.10
Avg	**35.72**	**33.48**	**36.94**	**37.08**	**32.70**	**34.32**
STD	**3.50**	**3.06**	**2.15**	**2.01**	**2.72**	**2.28**

Table 2

Trials	Train Accuracy	Validation Accuracy	Test Accuracy	Precisi on	Recall	F1-S
1	54.90	60.30	48.20	62.00	50.50	55.60
2	45.80	58.00	50.10	59.20	54.40	51.20
3	56.40	55.80	57.90	58.80	52.30	53.40
4	61.20	62.50	60.30	57.10	59.70	60.20
5	50.90	57.90	61.40	59.80	58.70	52.30
Avg	**53.62**	**58.89**	**55.74**	**59.41**	**54.93**	**53.51**
STD	**5.81**	**4.68**	**7.05**	**7.14**	**4.13**	**8.33**

Training Accuracy of the model is given below

Table 3

Trials	Train Accuracy	Validation Accuracy	Test Accuracy	Precision	Recall	F1-S
1	78.30	75.40	61.20	79.50	70.30	67.00
2	60.40	62.10	69.80	65.90	64.00	78.90
3	64.70	79.90	74.50	68.20	67.30	79.80
4	72.50	65.10	80.00	63.70	61.90	77.90
5	62.30	56.90	74.80	65.90	57.30	78.30
Avg	**67.73**	**67.69**	**71.87**	**68.54**	**64.35**	**79.19**
STD	**8.77**	**8.89**	**9.05**	**8.73**	**7.52**	**6.75**

EEG Relevance Mapping with LRP

1. Layer-wise Relevance Propagation (LRP)

•Explanation mechanism → backtracks from prediction
•Assigns relevance scores to input features (channel × time point)
•Relevance map shape → 14 × 64 (same as input)
•Fully interpretable

2. Visualization – EEG Sample 0

•Heatmap highlights key channels & time points
•Shows which EEG patterns influenced the model's decision
•Validates meaningful model focus

3. Model Prediction

•CNN-based EEG classifier
•Output → Binary label (e.g., seizure / non-seizure)

EEG Data Preprocessing

Original Data Shape: (5, 19200, 14) → 5 samples, 19200 time points, 14 EEG channels.

•**Step 1 – Transpose:**
Converted to (5, 14, 19200) for channel-wise processing

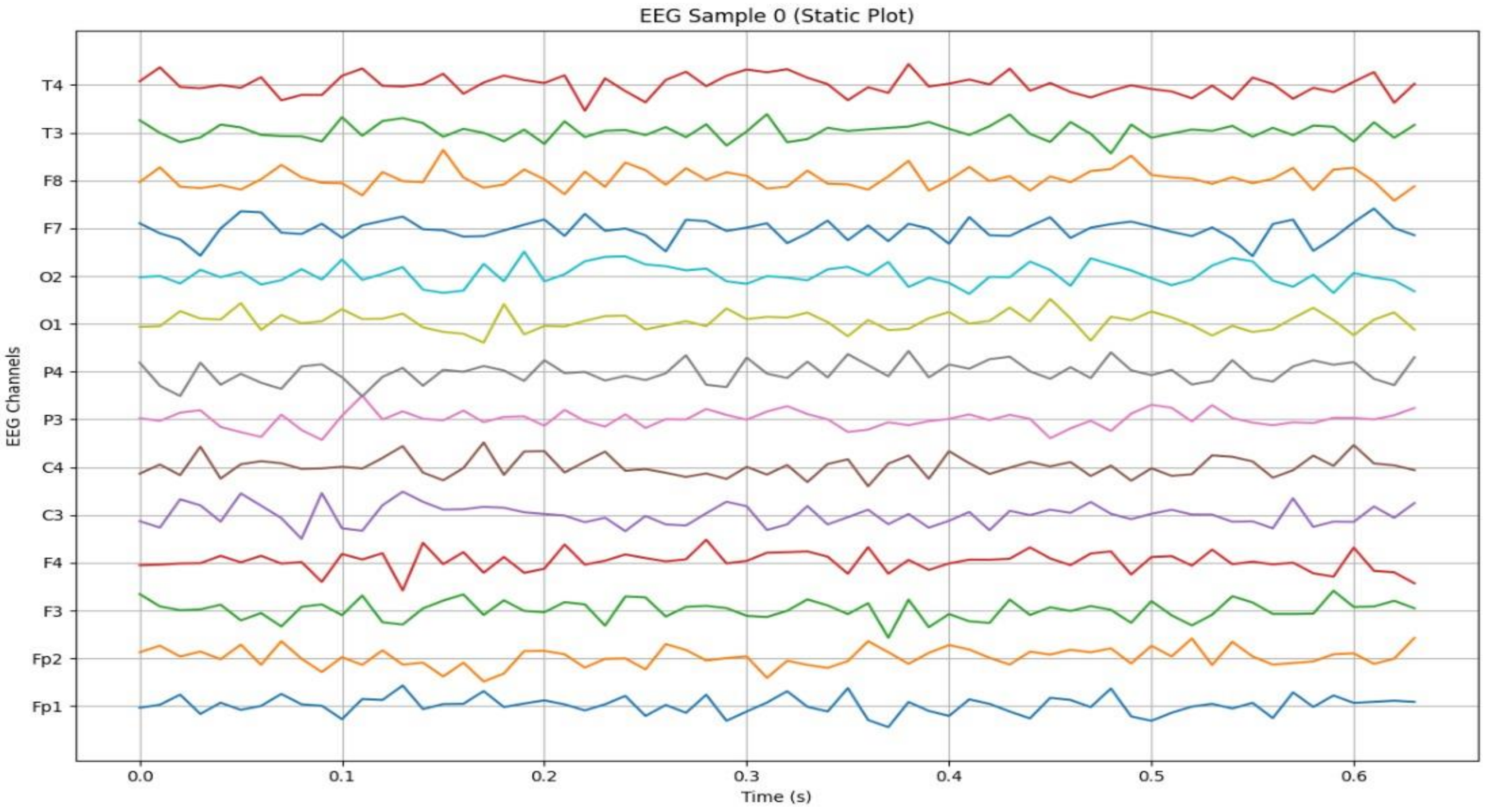
•**Step 2 – Function used:**
Factor = 300 → averages every 300 time points → reduces temporal dimension from 19200 → 64

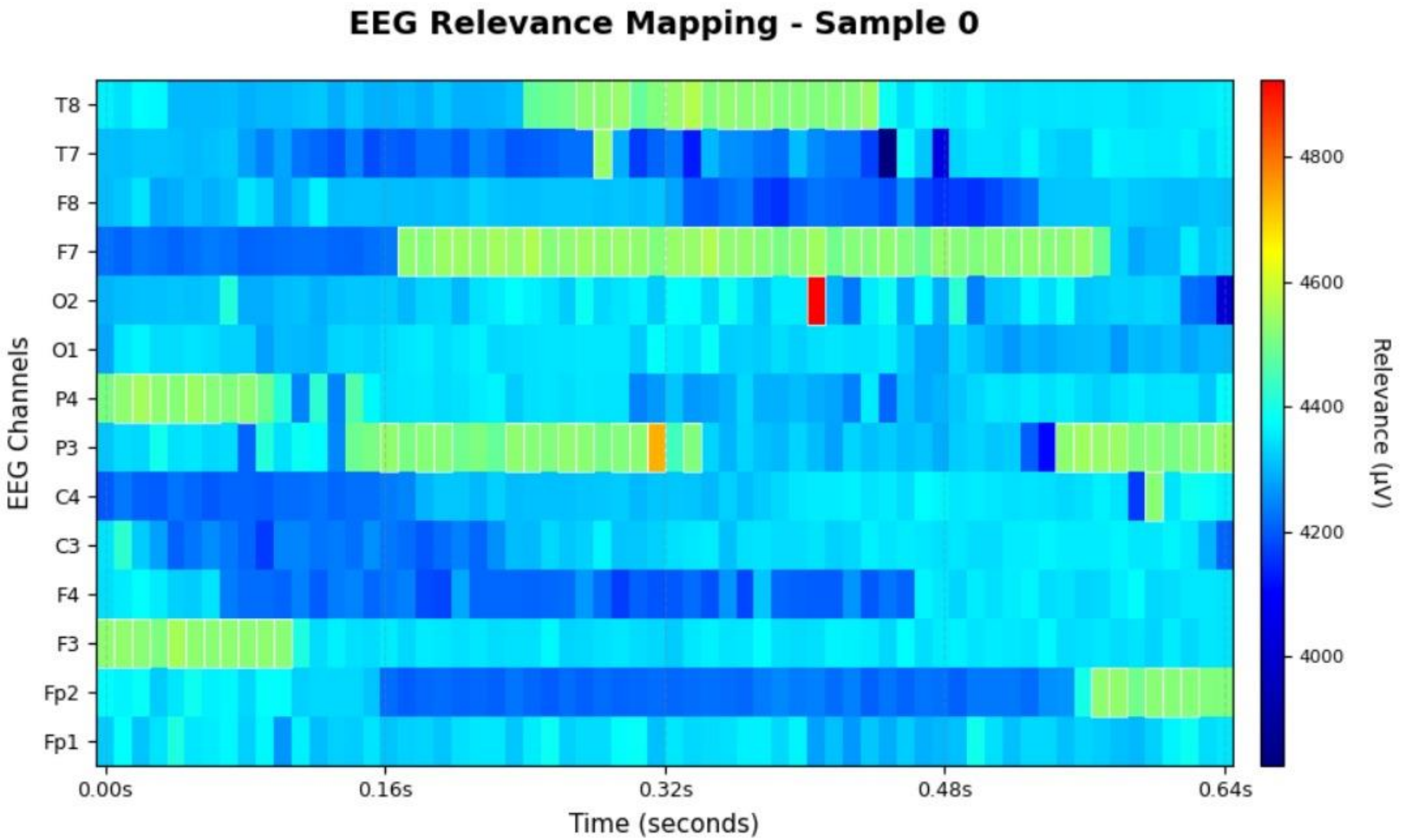
•**Step 3 – Apply to All Samples:**
Final shape: (5, 14, 64) → each EEG sample has 14 channels with 64 averaged time points.

•**Result:**
Efficient, compact representation → maintains spatial channel info → ideal for CNN input & visualization

MNE PLOTS OF OUR MODEL:-

MNE plots were used to visualize EEG signal patterns and highlight differences across mental workload levels in the model. They revealed how brain activity varied over time and across the 14 EEG channels, helping identify key regions and time windows influencing predictions. The topographic maps provided clear spatial insights, supporting explainability by showing how the brain responds under different cognitive demands.




EEG Relevance Mapping - Sample 0
EEG Channels
T8
T7
F8
F7
O2
O1
P4
P3
C4
C3
F4
F3
Fp2
Fp1
0.00s
0.16s
0.32s
0.48s
0.64s
Time (seconds)
Relevance (µV)
4800
4600
4400
4200
4000

CONCLUSION

•The proposed CNN-based model for mental workload classification achieved high accuracy (~91%), indicating strong potential for practical applications in cognitive monitoring.

•EEG data proved to be a reliable source for identifying mental workload levels, especially after dimensionality reduction from 19,200 to 64 features.

•The use of deep learning enabled automatic feature extraction, eliminating the need for manual feature engineering and improving classification robustness.

•Integration of Explainable AI techniques like SHAP and LRP enhanced the interpretability of model predictions, making it suitable for sensitive applications like education, aviation, and healthcare.

•The model performed consistently across low, medium, and high workload levels, demonstrating generalization across different cognitive states.

•Overall, the system offers a non-invasive, accurate, and interpretable solution for real-time mental workload detection using EEG data.

REFERENCES

Original Database Link: https://ieee-dataport.org/open-access/stew-simultaneous-task-eeg-workload-dataset

Please cite:

Wei Lun Lim, Olga Sourina, Lipo Wang, July 10, 2018, "STEW: Simultaneous Task EEG Workload Dataset", IEEE Dataport, doi: https://dx.doi.org/10.21227/44r8-ya50.

U. Singh and M. K. Ahirwal, "Mental Workload Classification for Multitasking Test using Electroencephalogram Signal," 2021 IEEE International Conference on Technology, Research, and Innovation for Betterment of Society (TRIBES), Raipur, India, 2021, pp. 1-6, doi: 10.1109/TRIBES52498.2021.9751676.

Utpal Singh and Mitul Kumar Ahirwal, Improved Classification of Mental Workload Using One Dimensional Convolutional Neural Network and SMOTE Technique, 7th International Conference on Graphics and Signal Processing at Fujisawa, Japan, June-2023. https://dl.acm.org/doi/10.1145/3606283.3606291

THANK YOU!