EEG Person Identification using Deep Learning: Full Project Report

This report summarizes the complete workflow for identifying individuals based on their Electroencephalography (EEG) signals during motor imagery tasks. We leveraged a hybrid deep learning architecture combining Convolutional Neural Networks (CNNs) and Gated Recurrent Units (GRUs) to capture both spatial-frequency and temporal features of EEG data from the PhysioNet EEG Motor Movement/Imagery Dataset.

1. Project Overview

The goal of this project was to build a robust system for person identification using EEG signals. The approach involved a multi-stage process:

1. **Data Acquisition & Preprocessing**: Downloading raw EEG data, cleaning, filtering, segmenting, and normalizing it.
2. **Model Development & Training**: Implementing and training a CNN-GRU deep learning model using PyTorch.
3. **Performance Evaluation**: Rigorously assessing the model's accuracy, F1-score, and other metrics.
4. **Data & Feature Visualization**: Gaining insights into the EEG data and model's learning process through various plots.

2. Data Acquisition & Preprocessing Summary

The data was sourced from the PhysioNet EEG Motor Movement/Imagery Dataset. Key preprocessing steps included:

- **Number of Subjects**: 109 subjects.
- **Sampling Rate**: 160 Hz.
- **Channel Selection**: Only 10 motor cortex channels were selected (`'C3', 'Cz', 'C4', 'CP3', 'CPz', 'CP4', 'FC3', 'FCz', 'FC4', 'Pz'`) after ensuring consistency across files.
- **Filtering**: A band-pass filter was applied (0.5-45.0 Hz) to remove noise and drift.
- **Segmentation**: Continuous EEG data was segmented into 2.0-second epochs with 50% overlap, resulting in 320 samples per segment.
- **Normalization**: Each segment was Z-score normalized independently.
- **Data Split**: The processed data was split into an 80% training set (63,701 samples) and a 20% test set (15,926 samples), ensuring stratification across subjects.

After preprocessing, the EEG data had a shape of `(samples, 10 channels, 320 time points)`.

3. Model Architecture & Training Summary

Model Architecture (CNN_GRU_Model)

The chosen architecture was a hybrid CNN-GRU model designed to extract both spatial-frequency and temporal features:

- **Input Layer**: `(10 channels, 320 samples)`.
- **CNN Feature Extractor**: Consisted of three 1D convolutional blocks. Each block included a `Conv1d` layer, `BatchNorm1d`, `ReLU` activation, and `MaxPool1d`. The filters used were `[32, 64, 128]`. A `Dropout` layer (p=0.3) was applied after the CNN blocks.
- **GRU Temporal Encoder**: A 2-layer Bidirectional GRU with 128 hidden units. This allowed the model to capture temporal dependencies in both forward and backward directions across the CNN-extracted features. A `Dropout` layer (p=0.3) was used within the GRU.
- **Classification Head**: A dense layer with 256 hidden units, followed by `BatchNorm1d`, `ReLU`, `Dropout` (p=0.5), and a final linear layer producing logits for the 109 subject classes.

Total Trainable Parameters: 663,501.

Training Configuration

- **Batch Size**: 64
- **Learning Rate**: 0.001
- **Optimizer**: Adam with `weight_decay=1e-4` (L2 regularization).
- **Loss Function**: Cross-Entropy Loss.
- **Epochs**: Trained for 50 epochs with Early Stopping (patience=10) and Learning Rate Scheduler (ReduceLROnPlateau, patience=5).
- **Device**: Training was performed on CPU.

Training Dynamics

- **Epochs Trained**: 50.
- **Best Validation Accuracy**: Achieved 99.87% at epoch 42.
- **Convergence**: The model converged well, with minimal overfitting, indicated by a final train-validation accuracy gap of 0.0007 and a low standard deviation of validation loss over the last 10 epochs (0.001179).

4. Performance Evaluation

The model demonstrated excellent performance on the test set:

Overall Metrics

- **Accuracy (Top-1): 99.87%**
- **Top-3 Accuracy**: 99.96%

- **Top-5 Accuracy**: **99.97%**
- **Top-10 Accuracy**: 99.99%
- **F1 Score (Weighted)**: 0.9987
- **F1 Score (Macro)**: 0.9988
- **Precision (Weighted)**: 0.9988
- **Recall (Weighted)**: 0.9987

Per-Subject Variability

- **Mean Per-Class Accuracy**: 99.88%
- **Worst Subject Accuracy**: 97.96% (Subject 75)
- **Best Subject Accuracy**: 100.00% (Subject 1, and many others)
- **F1 Score Statistics**: Mean F1 of 0.9988, min of 0.9863 (Subject 75), max of 1.0000.

These results indicate that while the model performs exceptionally well overall, there's slight variability in identification difficulty across subjects.

Error Analysis

- **Error Rate**: Only 0.13% of test samples were misclassified (20 out of 15926).
- **Confidence**: Correct predictions had a mean confidence of 0.9984, while incorrect predictions had a lower mean confidence of 0.7109, suggesting the model is less confident on its mistakes.
- **Most Confused Pairs**: The most common confusions involved pairs like S96 $\to$ S26, S95 $\to$ S84, and S75 $\to$ S84, each occurring twice.

5. Visualizations & Insights

EEG Signal Visualization

- Plots of raw EEG signals for different subjects showed distinct patterns, though complex to interpret directly.

Spectrogram Analysis

- **Individual Spectrograms**: Demonstrated time-frequency characteristics of EEG segments, highlighting power distribution across different frequencies over time.
- **Average Spectrograms per Subject**: Visualizing average spectrograms across selected channels and segments for various subjects revealed subtle but distinct frequency power signatures, supporting the idea of unique EEG

fingerprints" for identification.

t-SNE Feature Embeddings

- **Feature Extraction**: Deep features were extracted from the model's GRU output layer.
- **t-SNE Visualization**: t-Distributed Stochastic Neighbor Embedding (t-SNE) was applied to a subset of these features. The plots clearly showed distinct, well-separated clusters for different subjects, confirming that the model learns discriminative features for person identification. The cluster quality analysis indicated a high separation ratio, confirming strong inter-cluster distances and low intra-cluster spreads.

CNN Feature Maps Visualization

- Intermediate CNN feature maps were visualized, illustrating how the convolutional layers transform raw EEG signals into higher-level, abstract representations. These maps captured localized patterns in time and across channels, which are then fed into the GRU layers.

Power Spectrum Analysis

- Average Power Spectral Density (PSD) plots by subject revealed differences in power distribution across frequency bands (Delta, Theta, Alpha, Beta, Gamma) among individuals. This further supports the presence of unique neurophysiological characteristics that the model exploits for identification.

6. Conclusion & Future Work

The CNN + GRU hybrid model achieved an outstanding **99.87% accuracy** for identifying individuals from EEG motor imagery signals across 109 subjects. This remarkable performance validates the hypothesis that unique neural signatures exist within brainwave patterns, which can be effectively captured by deep learning architectures.

This work has significant implications for:

- **Biometric Authentication**: Developing highly secure identity verification systems.
- **Brain-Computer Interfaces (BCI)**: Enhancing personalization and control mechanisms.
- **Neurological Research**: Aiding in the study of individual differences in brain activity.

Recommendations for Future Improvements

To further enhance model robustness and generalize performance, especially in more challenging real-world scenarios, the following avenues could be explored:

1. **Advanced Preprocessing**: Investigate techniques like Independent Component Analysis (ICA) for artifact removal, Common Spatial Patterns (CSP), or Wavelet decomposition for feature extraction.
2. **Data Augmentation**: Implement more sophisticated data augmentation strategies (e.g., time-warping, channel shuffling, generative models) to increase data variability and reduce overfitting.

3. **Model Architecture Exploration**: Experiment with attention mechanisms for channel and temporal weighting, Transformer-based encoders, or deeper/wider CNN/GRU configurations.
4. **Transfer Learning/Cross-Session Evaluation**: Evaluate model performance across different recording sessions or with smaller subject-specific datasets to assess generalization capabilities.
5. **Interpretability**: Further investigate model interpretability methods to understand which specific EEG features or temporal dynamics are most critical for identification.

This project successfully demonstrated the power of deep learning in discerning individual identities from complex neurophysiological data, paving the way for advanced applications in various fields.