EEG Person Identification Project Report

1. Overview

Person identification using EEG Motor Movement/Imagery dataset (109 subjects).

2. Data & Preprocessing

Resampled 128 Hz; bandpass 1-40 Hz; 2s non-overlapping windows.
Mel-spectrogram (64 mel bins, hop length 64) normalized per window.
Channels averaged (future: multi-channel spatial features).

3. Model Architecture

Conv2D blocks -> GRU(128) -> GRU(64) -> Dense(softmax 109).
Captures local time-frequency + temporal evolution.

4. Training Setup

Adam, categorical crossentropy, accuracy metric, EarlyStopping + LR reduction.

5. Results

Test Accuracy: [REPLACE_WITH_VALUE]
Weighted F1 Score: [REPLACE_WITH_VALUE]
Top-10 class confusion subset and classification report computed.

6. Figures

Figure 1: Example Spectrogram (from preprocessing sanity plot).
Figure 2: Confusion Matrix (Top 10 frequent test-set classes).
Figure 3: t-SNE of penultimate GRU embeddings (test subset).
Insert images by regenerating and exporting PNGs: fig_spectrogram.png, fig_confusion.png, fig_tsne.png.

7. Interpretation

Spectrogram patterns differ across subjects; confusion shows harder pairs; t-SNE clusters partly distinct.
Overlap suggests need for spatial features and augmentation.

8. Strengths

Modular pipeline; hybrid architecture; embeddings visualization aids qualitative assessment.

9. Limitations

Single averaged channel; potential class imbalance; no session generalization yet.

10. Improvements

Multi-channel/CSP; augmentation; transformer encoders; session-aware splits; HP tuning.

11. Reproducibility

Run preprocessing -> save X.npy/Y.npy; run training -> save model + test splits; run report/viz.
Save figures: spectrogram/confusion/tsne then embed via LaTeX or Pandoc for final PDF.

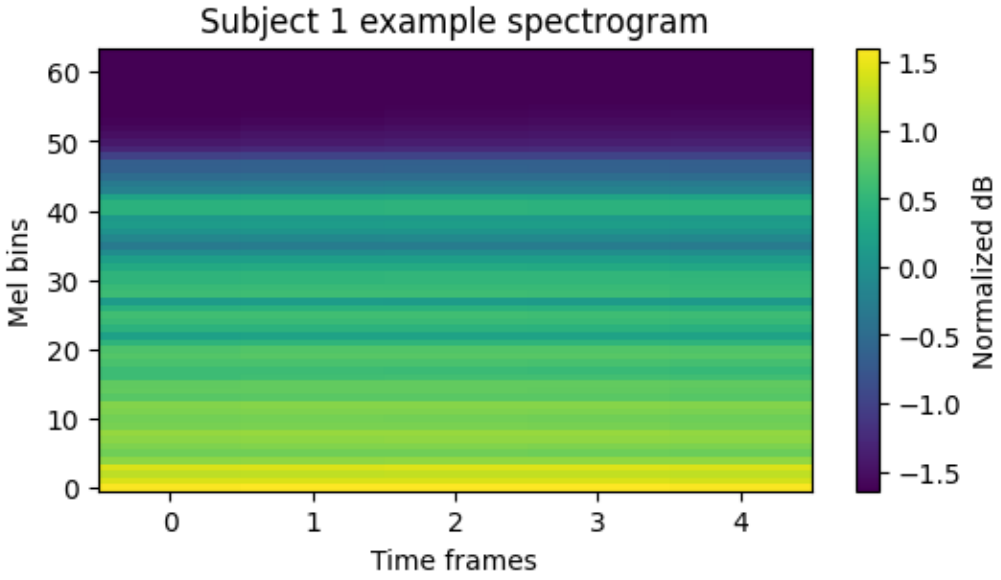
12. Ethical Use

Respect privacy; EEG biometric use requires consent; avoid re-identification misuse.

Metrics placeholders: replace after executing original evaluation cells.

Automated generation script provided in generate_pdf_report.ipynb.

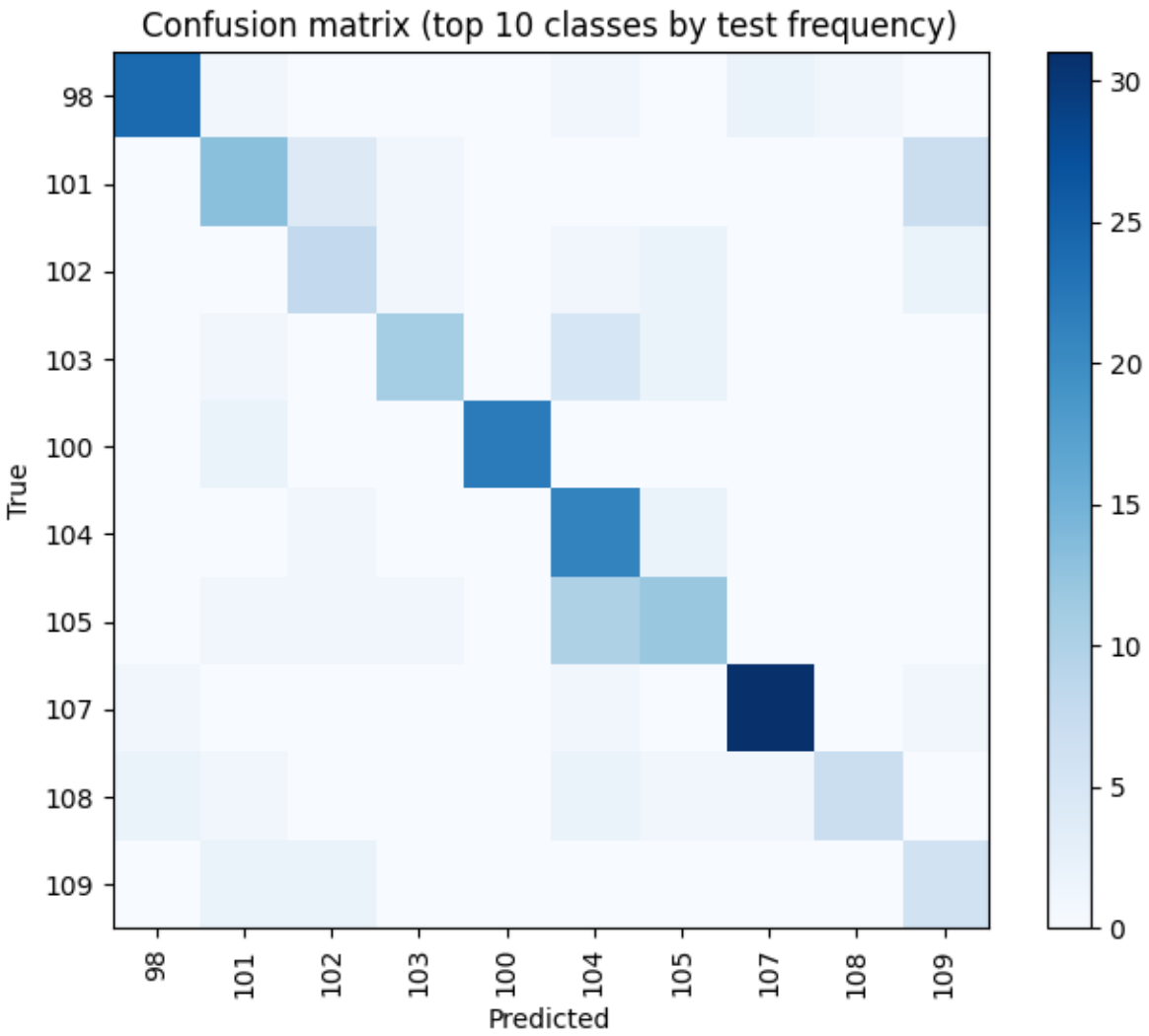
Report updated to include figure slots.

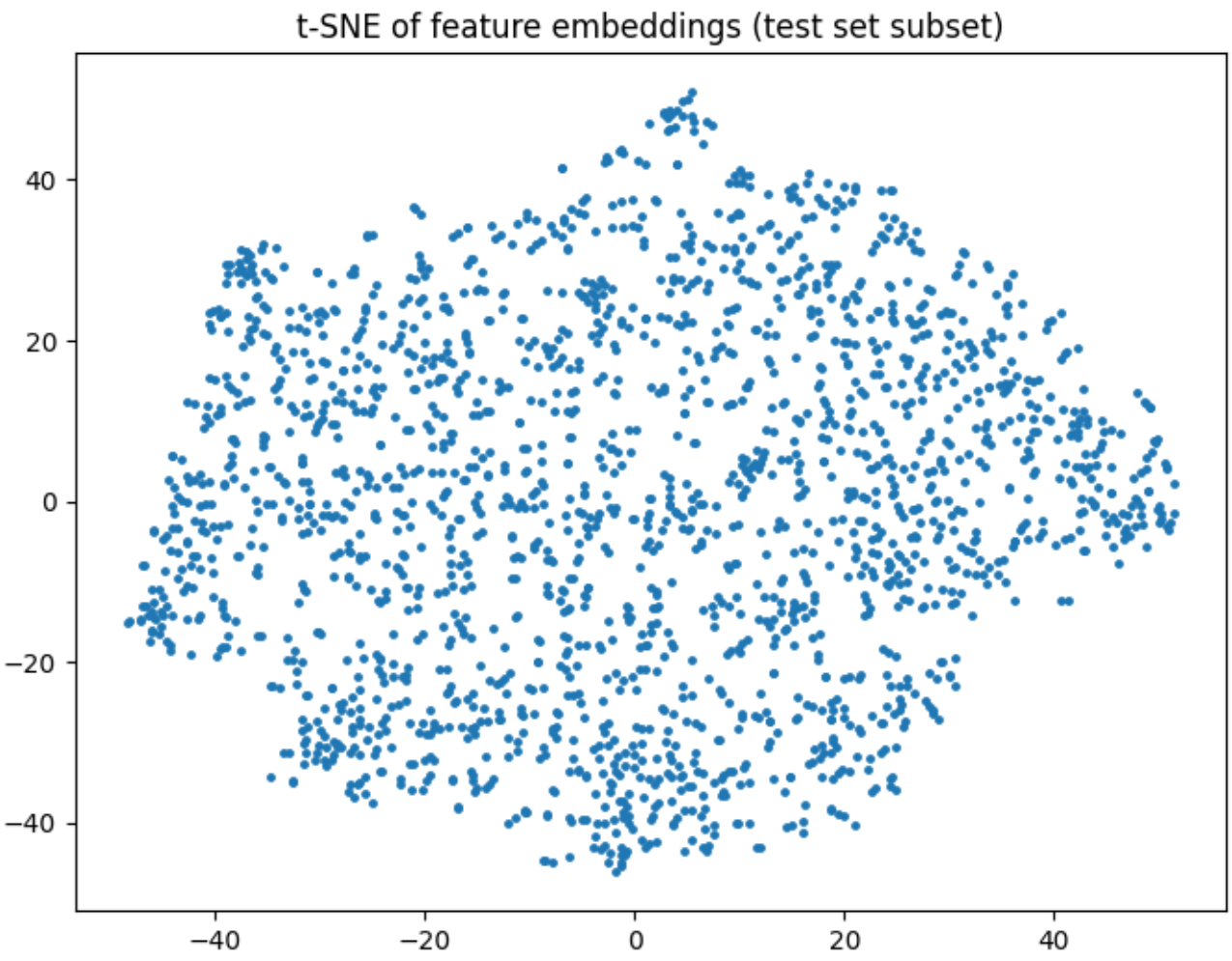


Model: "functional"

Layer (type)	Output Shape	Param #
input_layer (InputLayer)	(None, 64, 5, 1)	0
conv2d (Conv2D)	(None, 64, 5, 16)	416
batch_normalization (BatchNormalization)	(None, 64, 5, 16)	64
max_pooling2d (MaxPooling2D)	(None, 32, 2, 16)	0
conv2d_1 (Conv2D)	(None, 32, 2, 32)	4,640
batch_normalization_1 (BatchNormalization)	(None, 32, 2, 32)	128
max_pooling2d_1 (MaxPooling2D)	(None, 16, 1, 32)	0
permute (Permute)	(None, 1, 16, 32)	0
time_distributed (TimeDistributed)	(None, 1, 512)	0
gru (GRU)	(None, 1, 128)	246,528
gru_1 (GRU)	(None, 64)	37,248
dense (Dense)	(None, 109)	7,085

Total params: 296,109 (1.13 MB)
Trainable params: 296,013 (1.13 MB)
Non-trainable params: 96 (384.00 B)


Confusion matrix (top 10 classes by test frequency)
True
Predicted
98
101
102
103
100
104
105
107
108
109
30
25
20
15
10
5
0


t-SNE of feature embeddings (test set subset)
40
20
0
−20
−40
−40
−20
0
20
40

```
Classification report (top 10 classes):
              precision    recall  f1-score   support

          97       0.89      0.29      0.43        84
         100       0.62      0.15      0.25        84
         101       0.50      0.10      0.16        84
         102       0.79      0.13      0.22        84
          99       1.00      0.26      0.42        84
         103       0.51      0.25      0.34        84
         104       0.63      0.14      0.23        84
         106       0.91      0.37      0.53        84
         107       0.88      0.08      0.15        84
         108       0.38      0.07      0.12        84

   micro avg       0.71      0.18      0.29       840
   macro avg       0.71      0.18      0.28       840
weighted avg       0.71      0.18      0.28       840
```