



Technical Assessment: EEG and Foundation Models

Task 2: Transformer-based Foundation Model (BIOT)

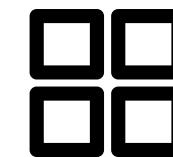
Name: Tasmin Karim

Dataset Descriptions



All Code Link:

https://github.com/TasminKarim-19/task_2



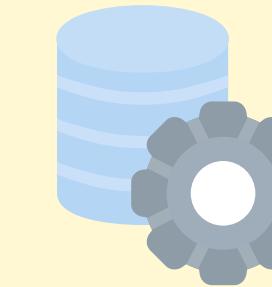
Training from Scratch:



Dataset Details

- Dataset: BNCI2014_001
- Subjects: [1, 2, 3, 4, 5, 6, 7, 8, 9]
Loaded X: (2592, 22, 801) Label dist: [1296 1296]
Prepared subjects: 9
Split: {'train': [4, 1, 8, 3, 5, 7], 'val': [2], 'test': [6, 9]}

Finetuning using pretrained BIOT:

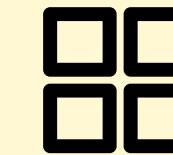


Dataset Details

- Dataset: CHB-MIT



As our project's goals are:
pretraining vs. training from scratch,
cross-dataset generalization,
interpretability via attention visualization.

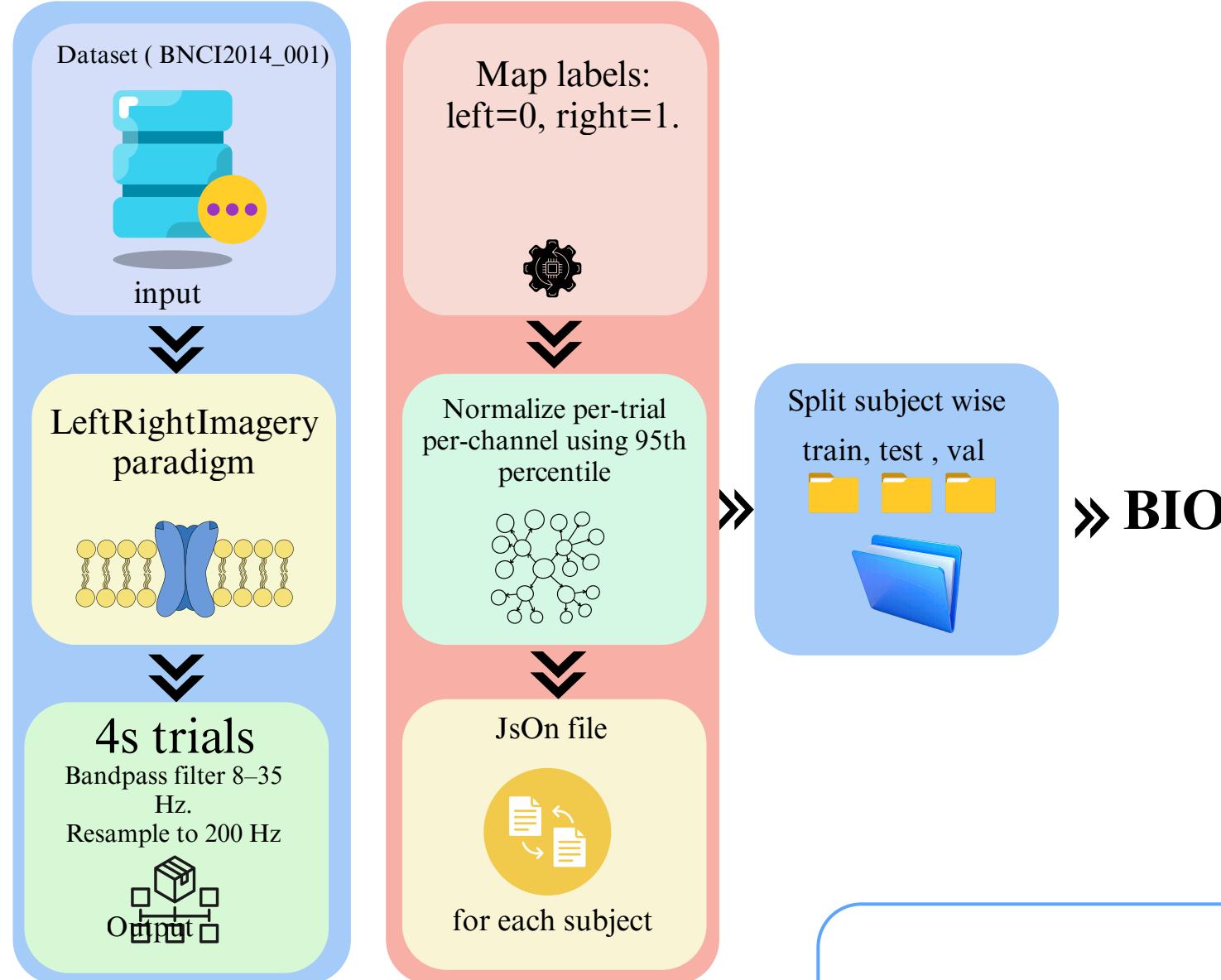


Split A
Train: chb01–chb19
Val: chb20–chb21
Test: chb22–chb23

Split B (swap)
Train same
Val/Test swapped

Training from scratch (Data Processing and Architecture)

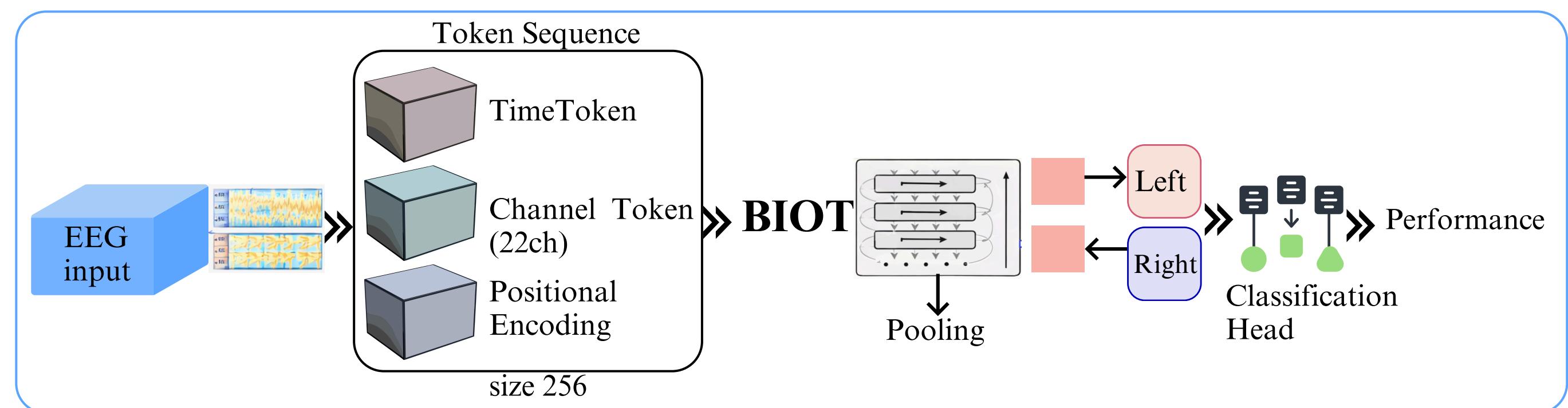
Data Processing overview



Steps

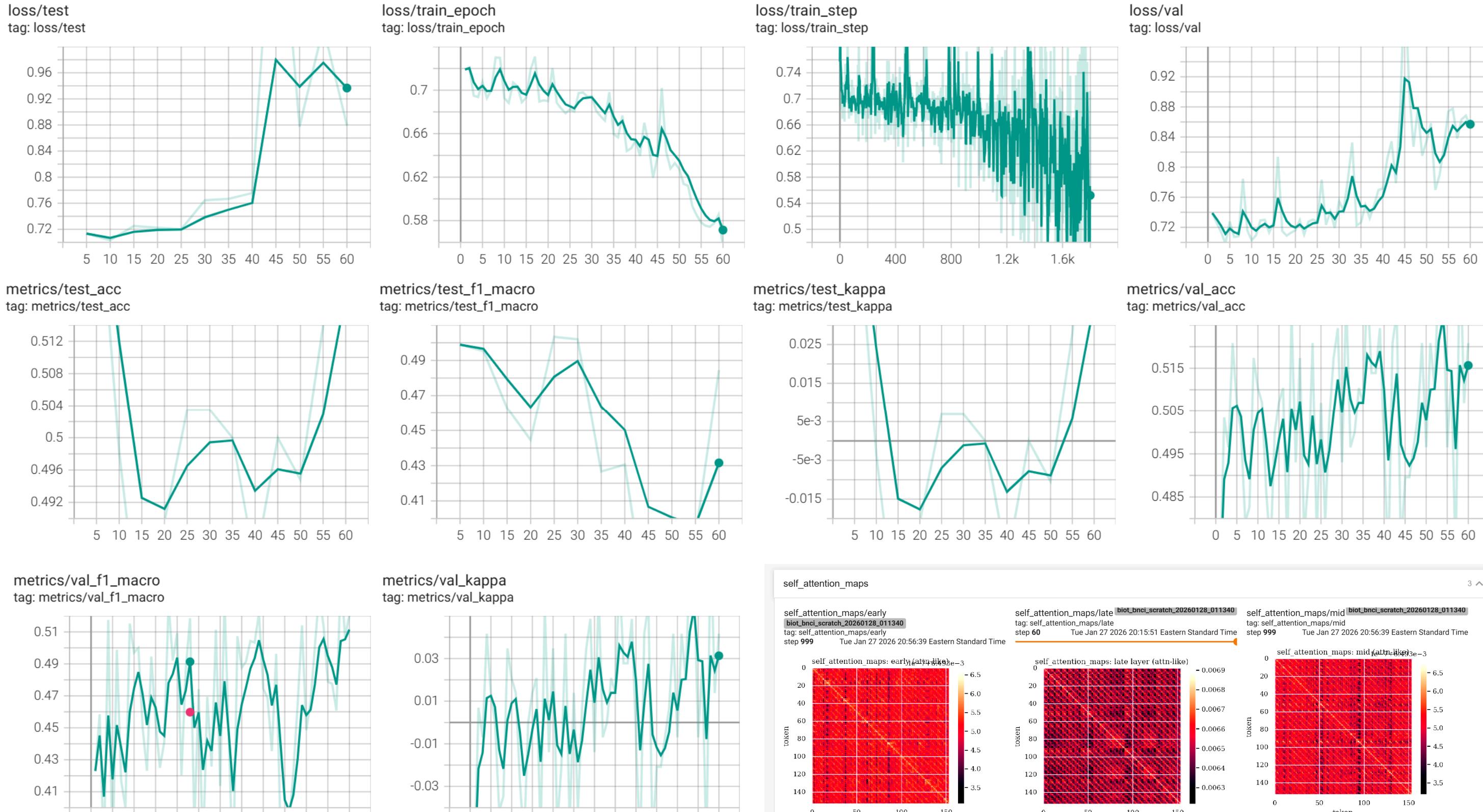
- Load BNCI2014_001 with MOABB.
- Use LeftRightImagery paradigm.
- (4s trials).
- Bandpass filter 8–35 Hz.
- Resample to 200 Hz.
- Map labels: left=0, right=1.
- Normalize per-trial per-channel using 95th percentile.
- Replace NaN/Inf with 0.
- Save subject-wise NPZ (x,y).
- Write meta.json + split_subjects.json (70/15/15, balanced).

Training from scratch architecture



Scratch Results

TensorBoard outcome:



Independent Test Results

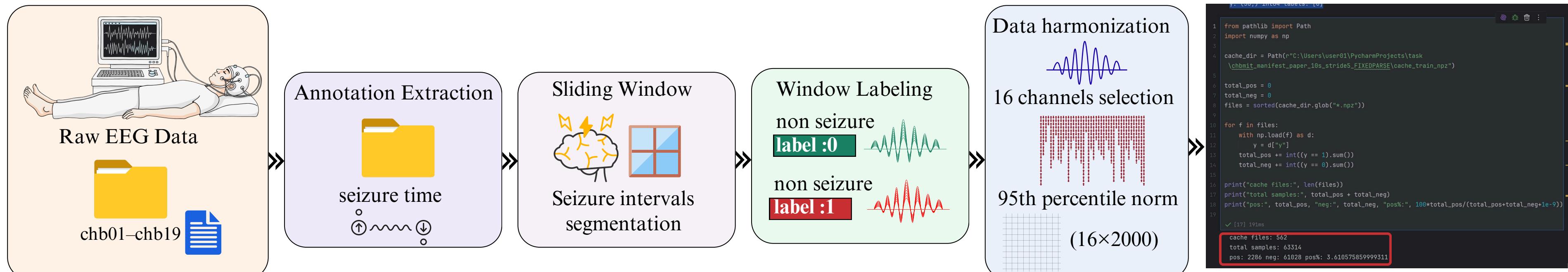
Metrics	Performance
Accuracy	0.5277
F1 macro	0.5131
Kappa	0.0555
Test Loss	0.8212

Attention based mapping:

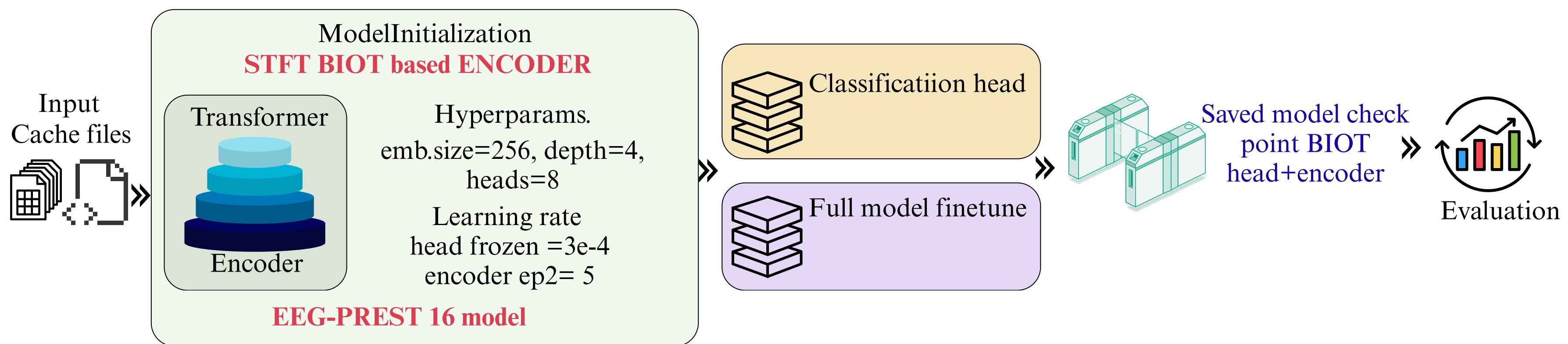
Fine tuning

Dataset processing overview

We preprocess CHB-MIT by parsing seizure intervals from subject summary files, segmenting each EDF recording into 10-second windows with 5-second stride, and labeling each window as seizure/non-seizure based on temporal overlap with annotated seizure intervals. We then harmonize samples following BIOT by selecting a fixed 16-montage channel set, applying per-channel 95th-percentile normalization, and resampling to 200Hz to obtain uniform (16×2000) inputs.

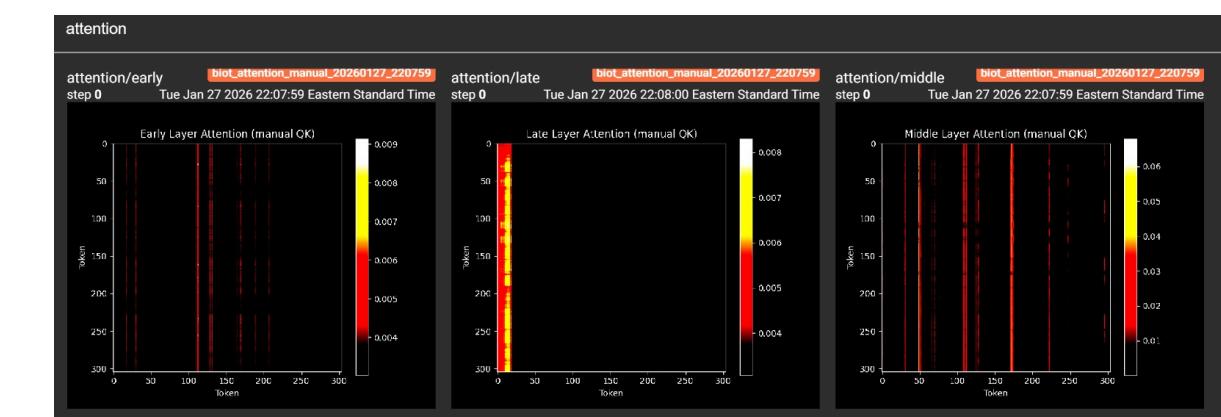
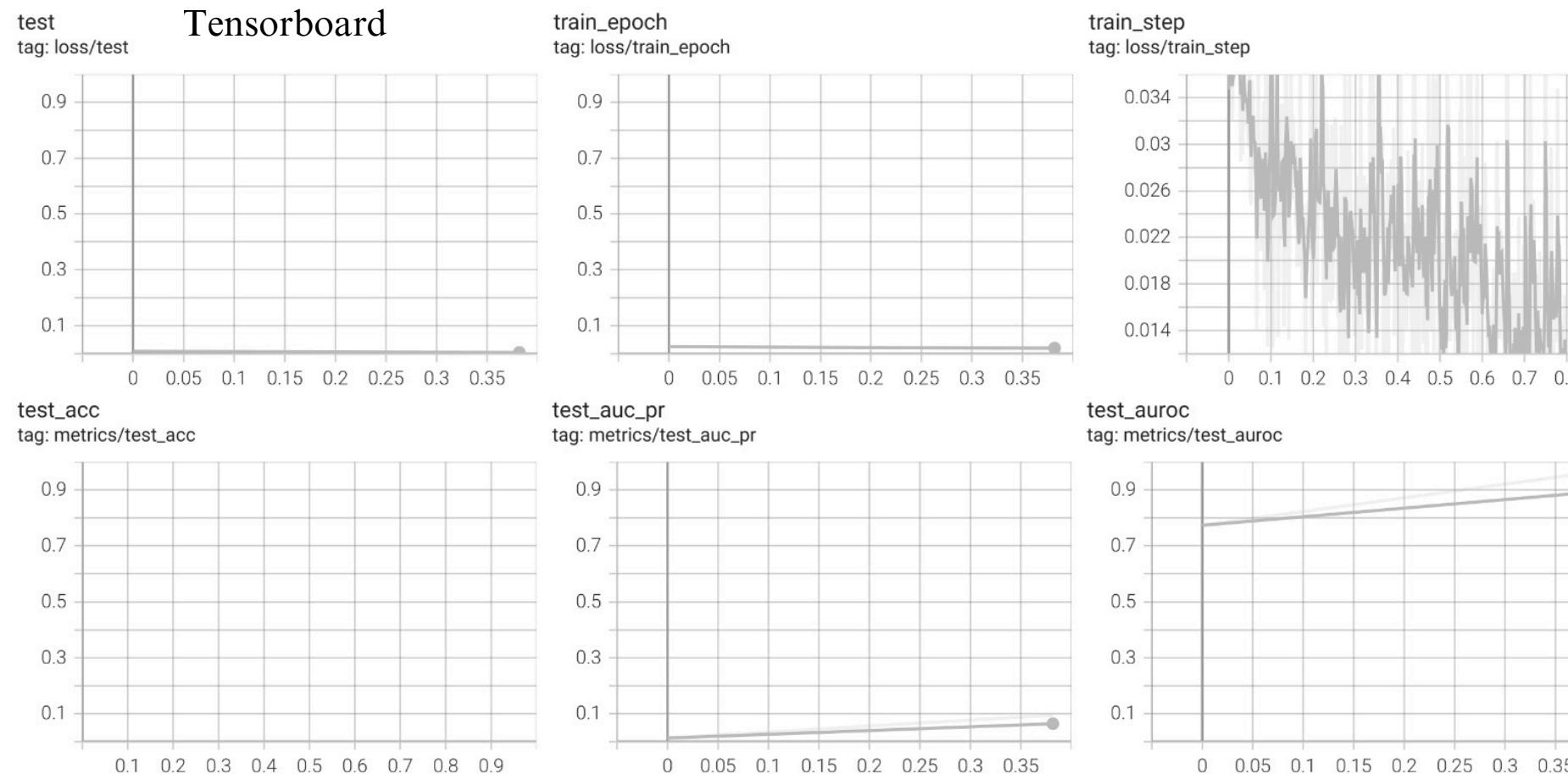


Fine tune model architecture



Fine tuning results

Cross data generalization BIOT model pretrained EEG PREST



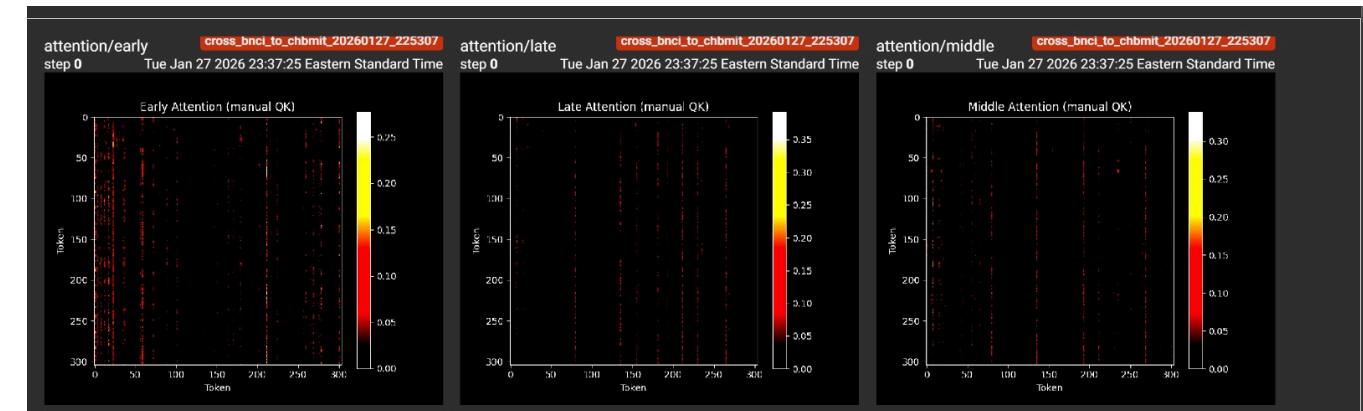
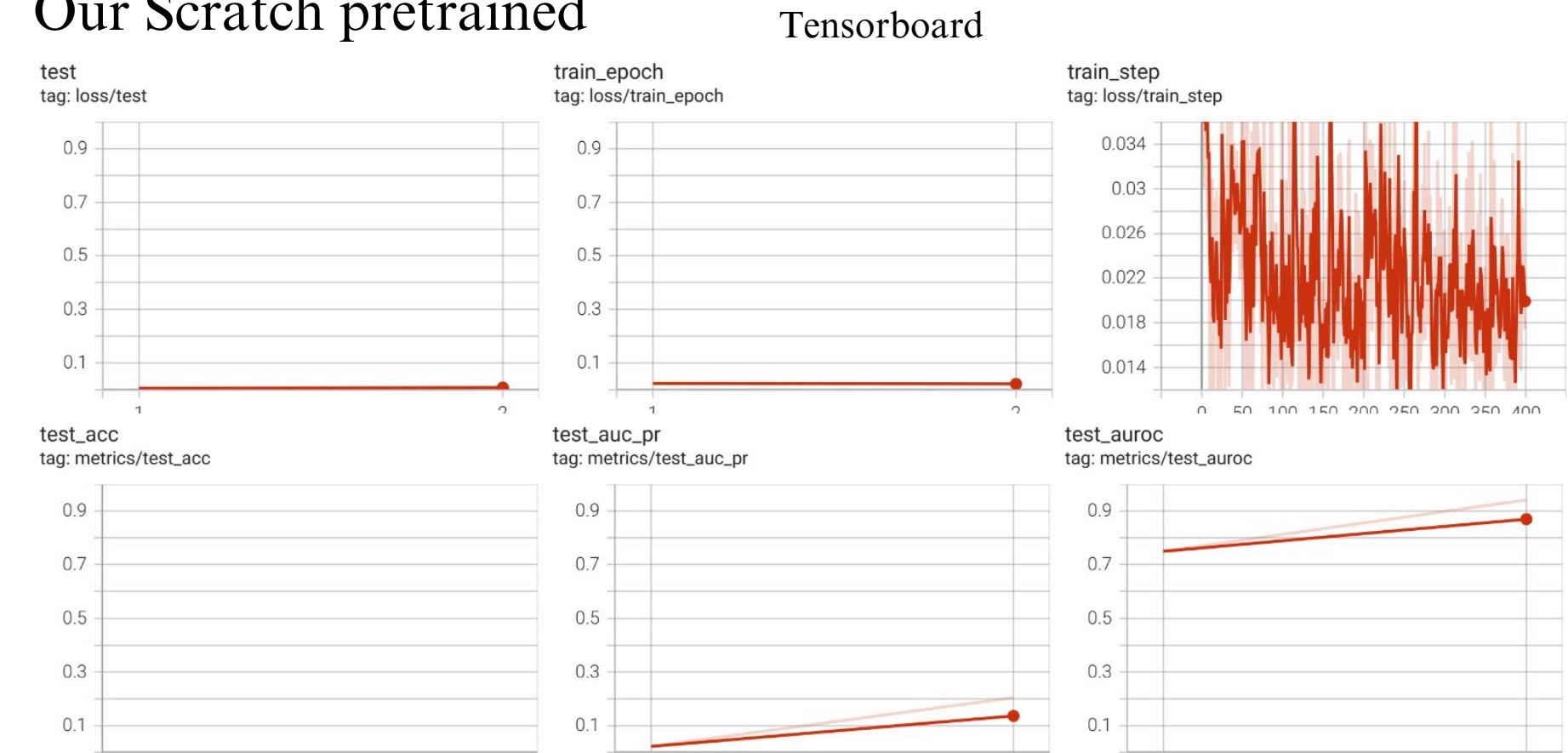
Attention

FINETUNING MODEL RESULT: BIOT PREST BASED

Scores:

Metrics	Performance
Accuracy	0.5173
AUC Roc	0.9610
AUC PR	0.0947
Test Loss	0.0027

Cross data generalization Our Scratch pretrained



Attention

FINETUNING MODEL RESULT: SCRATCH BASED

Scores:

Metrics	Performance
Accuracy	0.7296
AUC Roc	0.9414
AUC PR	0.2046
Test Loss	0.0101