Transfer Learning Based Motor Imagery Classification with Reduced Data

Presented by Istiaqur Rahman Fuad

ID: 2110576108
Azad Mahmud

ID: 2110476135
Md Yeasin Arafat
ID: 2110676145

Supervisor: MAHBOOB QAOSAR

Department of Computer Science and Engineering
University of Rajshahi

Introduction

What is BCI?

- A Brain-Computer Interface (BCI) is a technology that establishes a direct communication and control pathway between the human brain's neural system and an external computer or device.

What is Motor Imagery?

- Motor imagery (MI) means mentally imagining a movement without actually performing it
- It is like thinking about moving your hand, foot, or any body part but staying completely still.
- This is a crucial technology for Brain-Computer Interfaces (BCIs)

Introduction

Acquisition Of MI Signal

- Invasive Methods
- Semi-invasive Methods
- Non-invasive Methods (most common)

Why Non-invasive?

- Non-invasive techniques avoid any surgical penetration into the skull.
- No risk of infection or damage to brain tissue
- Suitable for healthy participants in research
- Easy to set up
- Affordable for universities and research labs

Introduction

What is EEG?

- EEG (Electroencephalography) is a technique used to record the electrical activity of the brain using electrodes placed on the scalp

How EEG Works?

- Neurons in the brain communicate using tiny electrical impulses.
- EEG electrodes placed on the scalp pick up these signals.
- The signals are amplified and sent to a computer.
- Finally, you see the brain activity as waveforms.

The Challenge with MI-EEG

Noisy & Complex Signal

EEG signals are inherently noisy, nonstationary, and high-dimensional, making robust feature extraction challenging.

The "Reduced Data" Problem

Data Scarcity: Training robust models requires vast amounts of data per user, which is difficult to acquire.

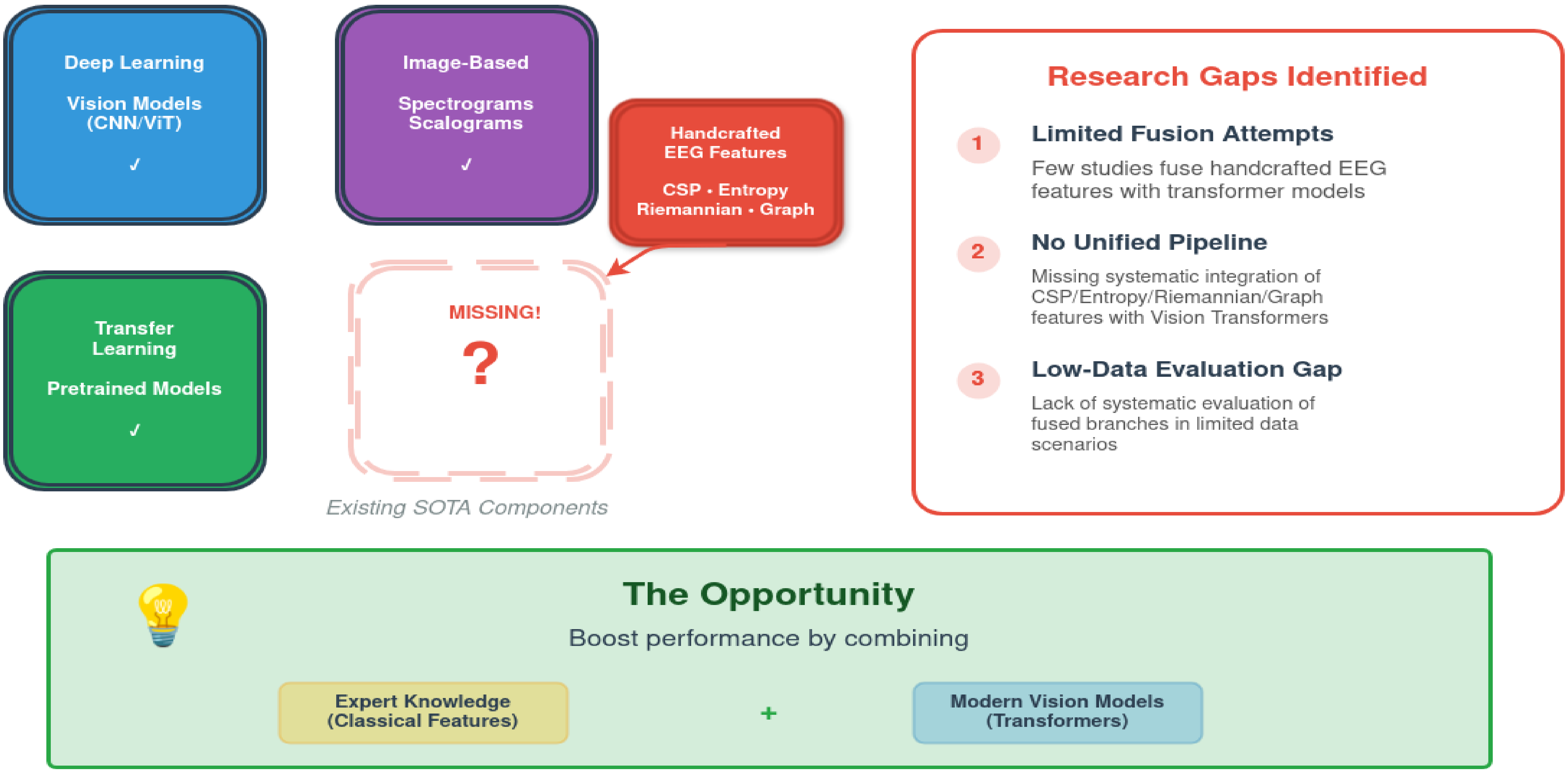
High Variability

EEG signals differ greatly between subjects and even between sessions for the same subject.

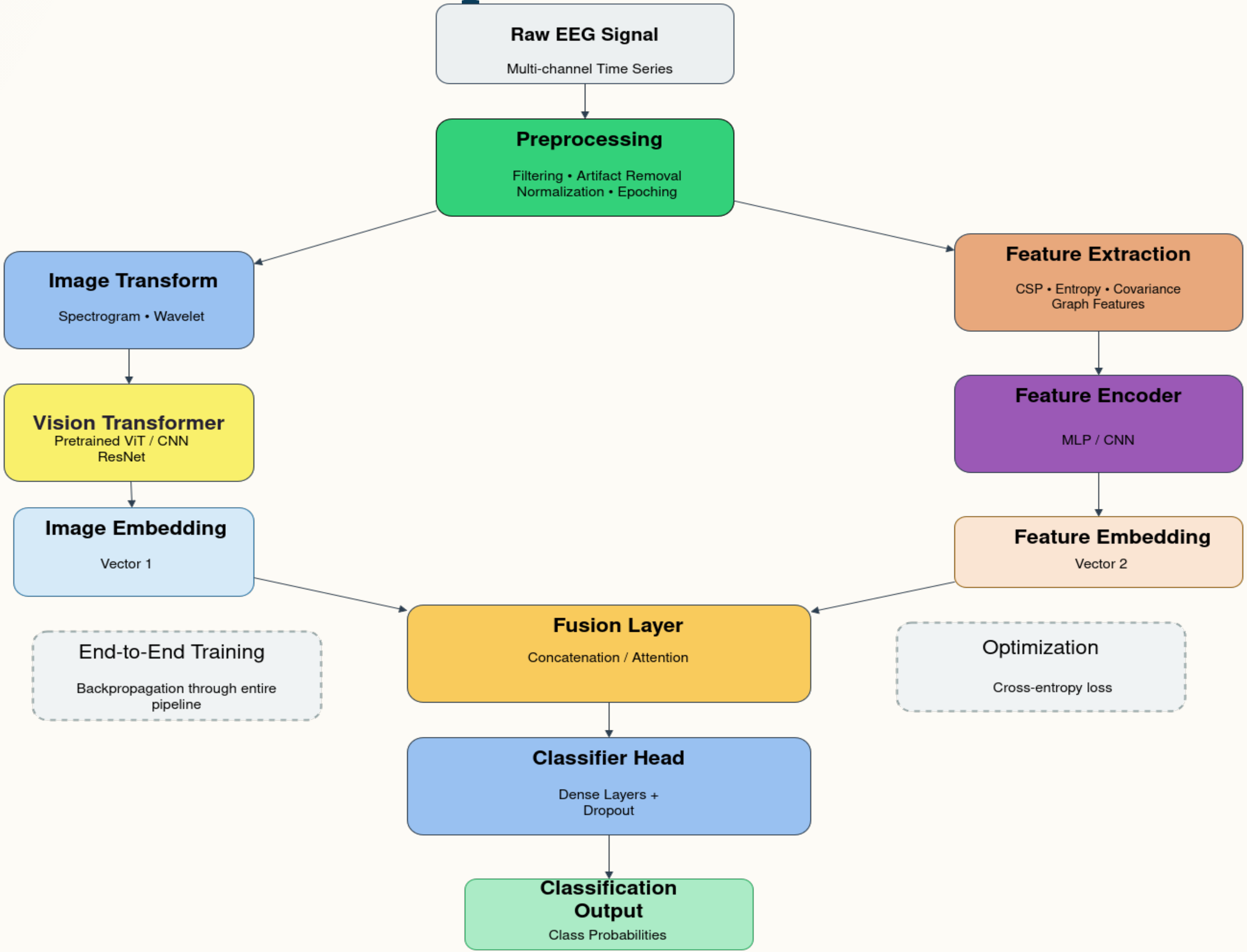
The Need: Robust Methods

We need methods that effectively handle limited data, high variability, and ensure strong cross-subject generalization.

The Missing Piece



The Proposed Framework



The Proposed Framework

- **Dual-Branch Architecture:** Dual-branch captures more EEG information.
- **Branch A- Image-Based ViT:** ViT processes EEG (spectrograms, GAF) images.
- **Branch B- Handcrafted Features:** Handcrafted features processed by MLP/CNN.
- **Strategic Fusion Layer:** Integrates outputs for final classification
- **Optimized for Small Datasets:** Integrates outputs for final classification

A multi-source domain adaptation framework that leverages a hybrid feature extraction model to overcome data scarcity and improve generalization.

Literature Evidence Supporting Our Approach

CSP + Swin Transformer (2025)

Achieved state-of-the-art accuracy on Motor Imagery (MI) tasks by effectively combining classical features with modern transformers[1].

Cross-dataset TL + DL (2023)

Proposes a cross-dataset framework using multi-task learning and pre-training to improve classification across heterogeneous datasets[2].

Multi-Scale CNN + TL (2022)

Combines multi-scale CNNs and attention mechanisms to capture diverse spatial–temporal features, improving cross-subject generalization[3].

Riemannian TL (2024)

Leverages log-Euclidean metrics for unsupervised Riemannian alignment, improving both classification accuracy and computational efficiency[4].

Public MI-EEG Datasets

PhysioNet MMIDB

Includes 64-channel EEG and 4-channel EMG data for real and imagined hand movements. It contains over 1,500 recordings from 109 volunteers.

Left/Right Hand 1D/2D movements

19-electrode data of one subject with various combinations of 1D and 2D hand movements (actual execution).

Open-BMI

Includes 52 subjects, results of physiological and psychological questionnaires, EMG Datasets, location of 3D EEG electrodes, and EEGs for non-task related states

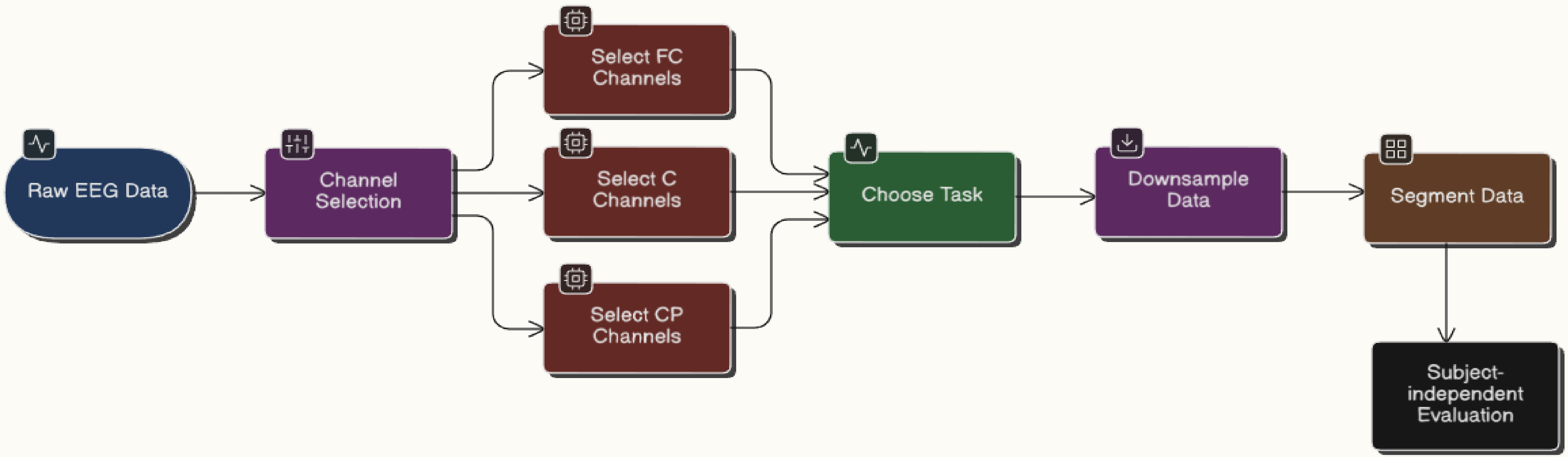
High-Gamma Dataset

Includes 128-electrode EEG recordings from 14 healthy subjects. It includes approximately 1,000 four-second trials per subject, divided into 13 runs. The trials cover four classes: left hand movement, right hand movement, both feet movement, and rest.

BCI Competition IV-2a

22-electrode EEG motor-imagery dataset, with 9 subjects and 2 sessions, each with 288 four-second trials of imagined movements per subject. Includes movements of the left hand, the right hand, the feet and the tongue.

EEG Preprocessing



EEG Preprocessing

Channel Selection

Only motor-related EEG channels were chosen

- FC-5/3/1/2/4/6
- C-5/3/1/z/2/4/6 and
- CP-5/3/1/z/2/4/6

Band-pass filter

- Band-pass filtering = keep useful frequencies + remove noise

Notch filter (50/60 Hz)

- To remove electrical line noise

EEG Preprocessing

ICA

- To identify and remove
- Eye blinks (EOG)
- Muscle noise (EMG)
- Heartbeat (ECG)
- Movement artifacts

Downsampling

- Each trial was reduced to same time steps

Epoching (Segmentation)

- Split signal into trials

Tools & Frameworks

PyTorch

Deep learning framework, providing flexibility and powerful GPU acceleration.

TorchAudio

Library for audio and signal processing, crucial for generating spectrograms from raw EEG signals for ViT inputs.

Hugging Face

Access to state-of-the-art pretrained Vision Transformers (ViTs) for robust spectrogram processing.

MNE

An open-source Python package for exploring, analysing, and visualising human neurophysiological data, essential for EEG preprocessing and feature extraction

TorchEEG

Library built on PyTorch for EEG signal analysis. It aims to provide a plug-and-play EEG analysis tool, so that researchers can quickly reproduce EEG analysis work.

Step 1: Multi-Source Data Strategy

Source Selection

Aggregate multiple existing MI-EEG datasets

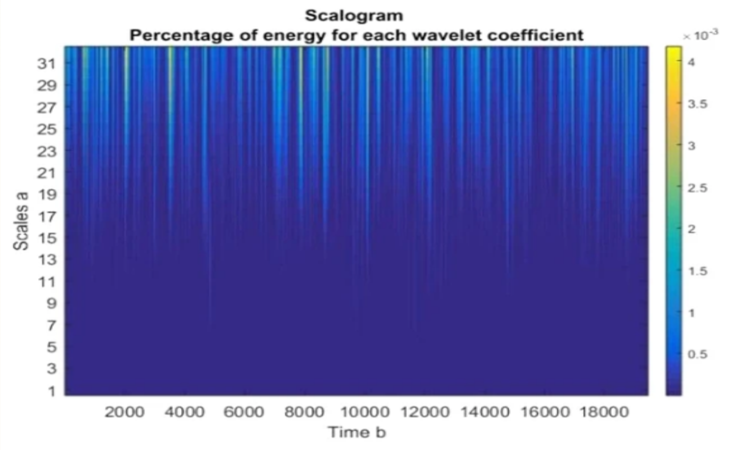
Class Consistency

Focus strictly on **Left-vs-Right Hand MI** to ensure consistent class labels.

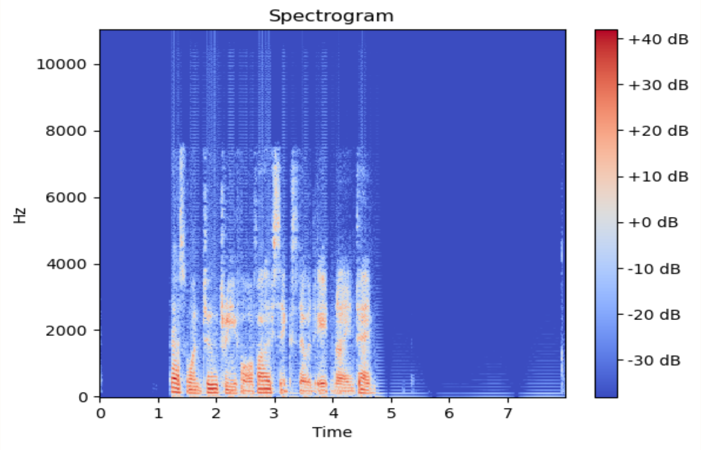
Channel Alignment

Align channels by selecting common motor electrodes (C3, Cz, C4 etc.)

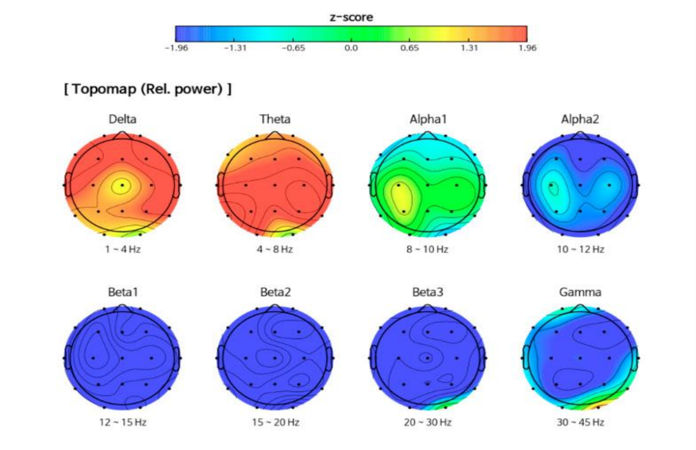
Step 2: EEG-to-Image Transformation



CWT Spectrograms: Convert 1D EEG signals into 2D time-frequency scalograms. These capture rich textures[5].

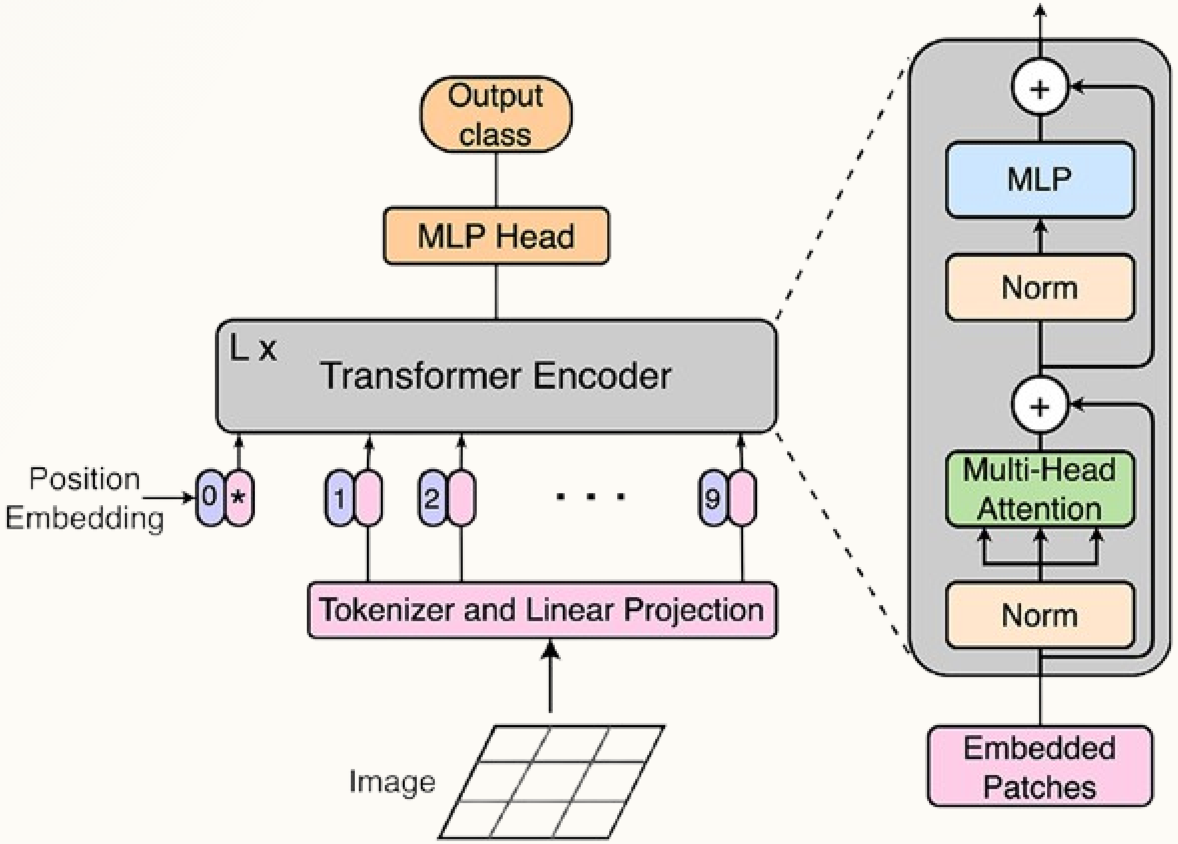


STFT Spectrograms: A common alternative, often focusing on specific bands (mu/beta rhythms)[6].



Topographic Maps: Stack 2D scalp maps of band power over time to create a "video-like" sequence[7].

Step 3: Feature Extraction via Vision Transformer



Learning from EEG Images

The Vision Transformer (ViT) processes these EEG-derived images to learn rich spectral, temporal, and spatial patterns.[1]

Self-Attention is the key mechanism: it allows the model to capture long-range dependencies in the data means how different frequency bands relate over time.

Why a Vision Transformer (ViT)?

vs. CNNs (Local Focus)

Traditional CNNs excel at learning local features (like edges in a photo). In spectrograms, this means local time-frequency "blobs."

Their fixed, local receptive fields struggle to capture the non-local dependencies inherent in brain activity, such as long-range temporal correlations or functional connections between distant electrode channels.

ViT (Global Context)

ViT's self-attention mechanism models the entire image at once.

It can link a pattern at the **start** of a trial to one at the **end**, or correlate a low-frequency band with a high-frequency one.

This captures **global context** and long-range dependencies, which CNNs struggle with.

Step 4: Parallel Branch (Mathematical Features)

Riemannian Geometry

Compute **Covariance Matrices** for each trial. These matrices are Symmetric Positive-Definite (**SPD**).

Use Riemannian geometry such as, tangent space projection, to handle these features, respecting their unique manifold structure.

Common Spatial Patterns (CSP)

An alternative/complementary method.

CSP finds spatial filters (channel weights) that **maximize** the variance for one class (left hand) while **minimizing** it for the other (right hand).

Differential Entropy

A measure of signal complexity often used to characterise brain states, such as emotional responses or cognitive load, from EEG data.

Step 5: Feature Fusion Strategy

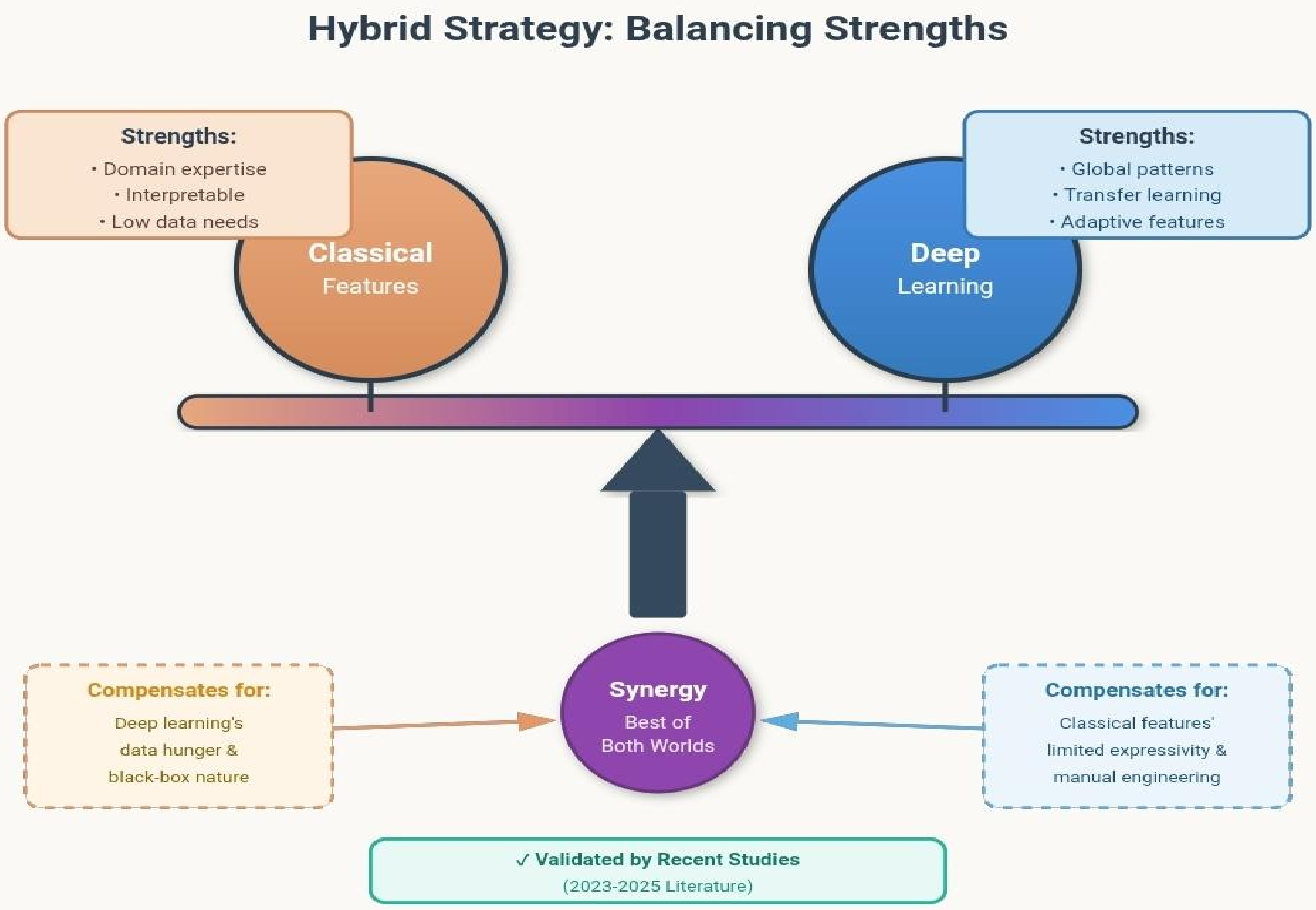
Early Fusion

Simply concatenate the ViT and Math feature vectors. Simple, but may not be optimal.

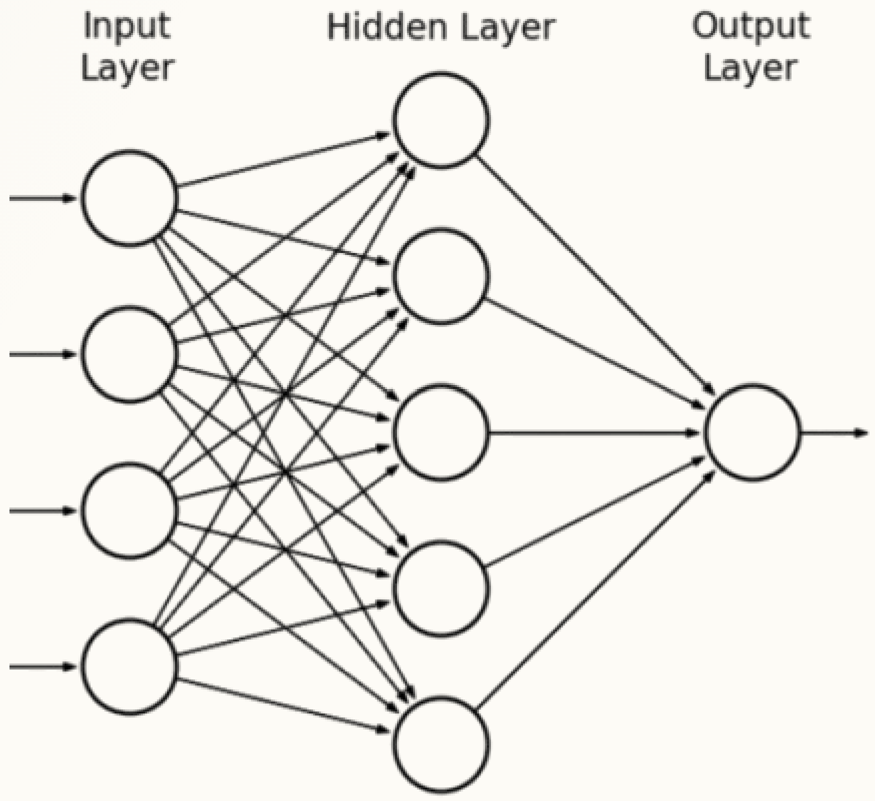
Attention-Based Fusion

Use a learnable gating or cross-attention to intelligently weight the two branches.

Why This Hybrid Strategy Makes Sense



Step 6: Classifier Design



Lightweight Classifier Design

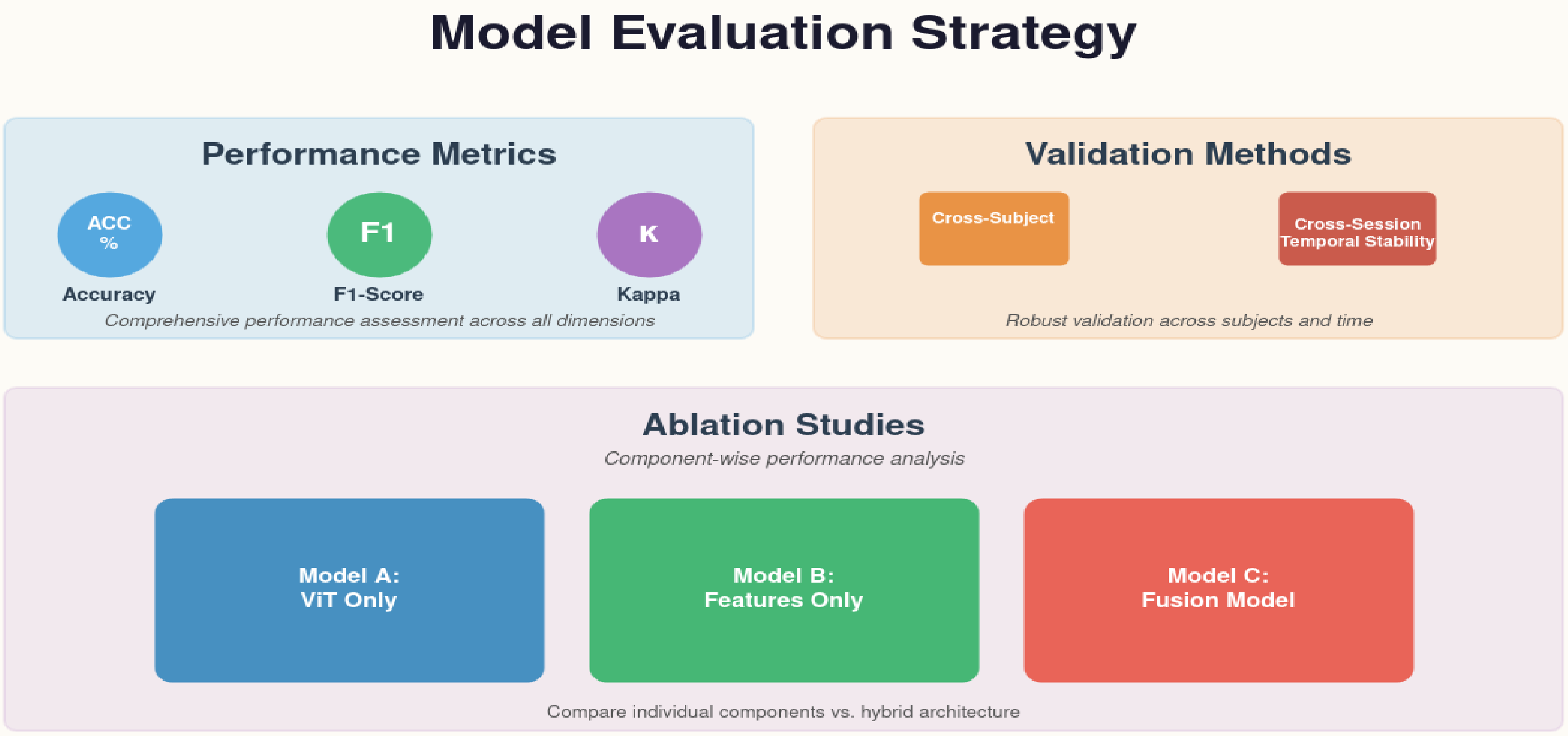
The fused feature vector is passed into a lightweight classifier. Since most of the learning occurs in the feature extraction stage, the classifier only needs minimal depth.

Options

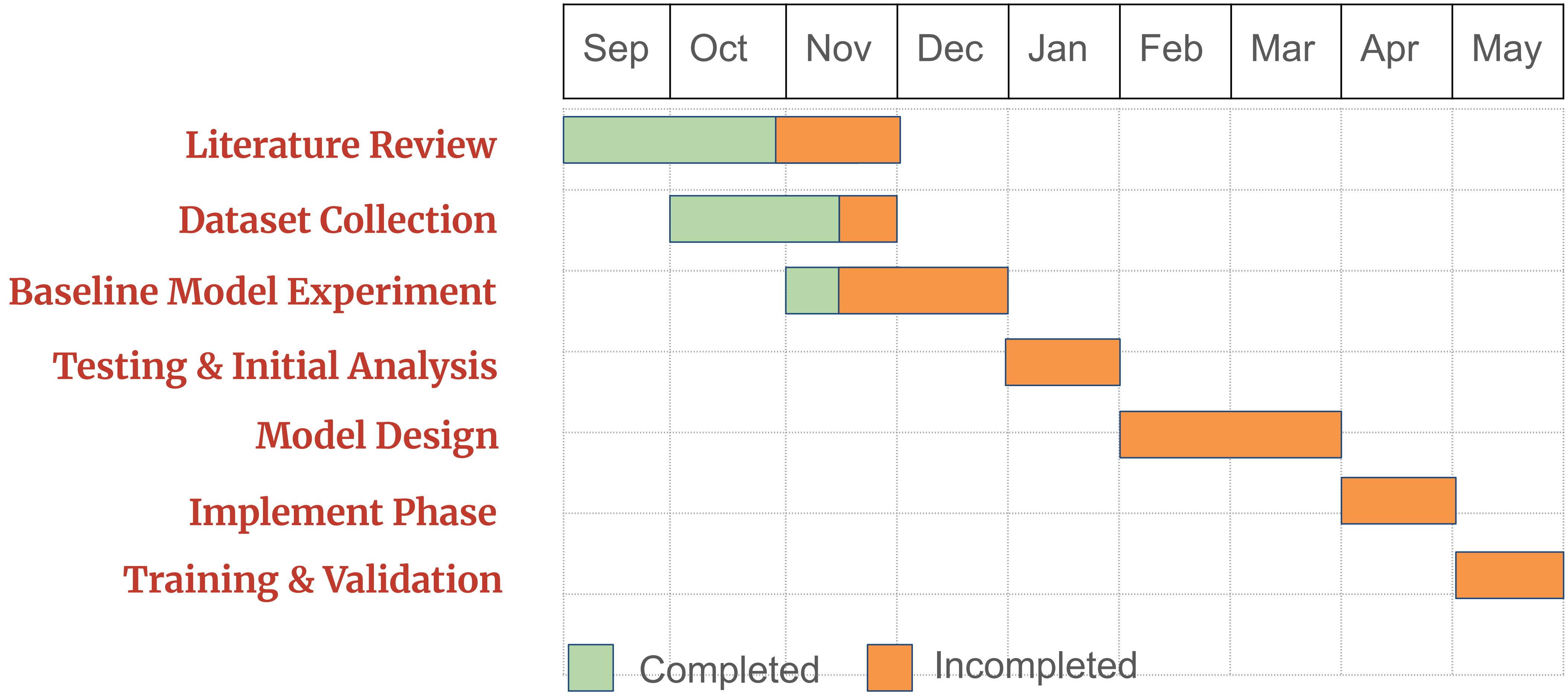
MLP: A Multi-Layer Perceptron with 1-2 hidden layers.

Linear Models: A linear SVM or LDA could also be trained on the powerful fused features.

Step 7: Evaluation



Timelines (Gantt Chart)



Key Advantages and Performance Benefits

- **Better Generalisation with Limited Data:** Leveraging established features helps models learn more effectively from smaller datasets, reducing the need for extensive data collection
- **Complementary Features Reduce Overfitting:** Combining diverse feature sets provides a richer representation, making the model more robust and less prone to overfitting on noise.
- **Elevated Cross-Subject Robustness:** The integration of different feature types aims to capture more invariant patterns across individuals, leading to improved performance in cross-subject generalisation.
- **Potential to Outperform Pure Spectrogram+ViT:** Pipelines By augmenting vision transformer models with expert-crafted features, we expect to achieve superior results compared to purely data-driven approaches.

Our Contributions

Systematic Feature Fusion

First to systematically integrate handcrafted and Vision Transformer (ViT) features for EEG data.

Complete Dual-Branch Architecture

Developed a comprehensive architecture with distinct branches for different feature types, ensuring optimal processing

Comprehensive Comparative Analysis

Extensive evaluation against pure ViT and pure handcrafted feature-based models, highlighting the

Promise for Low-Data Scenarios

Demonstrated strong potential for effective EEG classification even with limited data availability.

" By combining multi-source domain adaptation with a hybrid (ViT + Riemannian) feature model, we can create a data-efficient and robust classifier for MI-EEG. "

References

[1] X. Deng, H. Huo, L. Ai, D. Xu, and C. Li, “A Novel 3D Approach with a CNN and Swin Transformer for Decoding EEG-Based Motor Imagery Classification,” Sensors, vol. 25, no. 9, p. 2922, May 2025, doi: 10.3390/s25092922.

[2] Y. Xie et al., “Cross-dataset transfer learning for motor imagery signal classification via multi-task learning and pre-training,” Journal of Neural Engineering, vol. 20, no. 5, p. 056037, Oct. 2023, doi: 10.1088/1741-2552/acfe9c.

[3] Z. Chang, C. Zhang, and C. Li, “Motor Imagery EEG Classification Based on Transfer Learning and Multi-Scale Convolution Network,” Micromachines, vol. 13, no. 6, p. 927, Jun. 2022, doi: 10.3390/mi13060927.

[4] F. Zhuo, X. Zhang, F. Tang, Y. Yu, and L. Liu, “Riemannian transfer learning based on log-Euclidean metric for EEG classification,” Frontiers in Neuroscience, vol. 18, May 2024, doi: 10.3389/fnins.2024.1381572.

[5] S. Pattnaik, B. N. Rao, N. K. Rout, and S. K. Sabut, “Transfer learning based epileptic seizure classification using scalogram images of EEG signals,” Multimedia Tools and Applications, vol. 83, no. 36, pp. 84179–84193, Apr. 2024, doi: 10.1007/s11042-024-19129-4.

[6] Z. Ong, "Exploring the Short-Time Fourier Transform: Analyzing Time-Varying Audio Signals," Medium, 2023. [Online]. Available: exploring-the-short-time-fourier-transform-analyzing-time-varying-audio-signals-98157d1b9a12

[7] NeuroTraining Strategies, "How to interpret qEEG results?," eegstrategies.com, Jun. 17, 2024. [Online]. Available: how-to-interpret-qeeg-results

Thank You

Any Questions?