FA-GPNet: When Gaussian Process Meets Auto-Encoder and FBCSP - A Hybrid Model for Motor Imagery Classification

Trung M. Pham[1,3], Hieu M. Pham[1,3], Vi K. Nguyen[1,3], Truong D. Tran[1,3], Long S. T. Nguyen[1,3], Duc Q. Nguyen[1,3], Huong T. T. Ha[2,3], and Tho T. Quan[1,3(✉)]

[1] URA Research Group, Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology (HCMUT), Ho Chi Minh City, Vietnam
qttho@hcmut.edu.vn

[2] School of Biomedical Engineering, Ho Chi Minh City International University (HCMIU), Ho Chi Minh City, Vietnam

[3] Vietnam National University Ho Chi Minh City, Ho Chi Minh City, Vietnam

Abstract. Motor imagery electroencephalography (MI-EEG) decoding is challenged by noisy, non-stationary signals, high variability, and limited labeled data. We propose FA-GPNet, a unified framework that integrates Filter Bank Common Spatial Pattern (FBCSP) for multi-band spectral–spatial feature extraction, a deep autoencoder for non-linear compression and noise reduction, and a Gaussian Process Classifier (GPC) for probabilistic, uncertainty-aware predictions. Unlike conventional FBCSP pipelines that rely on manual feature selection and deterministic classifiers, FA-GPNet replaces heuristic ranking with data-driven latent representation learning and leverages GP's Bayesian framework for calibrated outputs. Under within-subject evaluation, FA-GPNet achieves 78.19% mean accuracy on BCI Competition IV-2b, surpassing strong traditional baselines and multiple deep networks, while remaining competitive with CapsNet. On the HCM-IU hand-binary MI dataset, FA-GPNet outperforms the well-optimized classical baseline. These results demonstrate that FA-GPNet offers a robust, reproducible, and efficient solution for MI-EEG decoding.

Keywords: Motor Imagery · Electroencephalography (EEG) · Auto-Encoder · Gaussian Process · Hybrid Model

1 Introduction

Electroencephalography (EEG) is a non-invasive neuroimaging technique that records brain activity via scalp electrodes [5]. Due to its affordability, portability,

T. M. Pham and H. M. Pham—The two authors contributed equally to this work.

T. T. Quan et al. (Eds.): MIWAI 2025, LNAI 16355, pp. 127–139, 2026.
https://doi.org/10.1007/978-981-95-4963-4_11

and high temporal resolution, EEG has become a cornerstone in brain–computer interface (BCI) research [15]. Among various BCI paradigms, motor imagery (MI)—where individuals imagine movements without physical execution—has attracted significant attention for applications in neurorehabilitation and assistive technologies [6].

Compared to evoked paradigms, MI can elicit neural activity spontaneously across multiple brain regions without external stimuli [14]. However, EEG signals are notoriously noisy, non-stationary, and vary significantly both within and across sessions [9]. These challenges make MI decoding difficult, particularly under limited labeled data. Traditional pipelines typically employ Common Spatial Patterns (CSP) [4,10,17] or its extension Filter Bank CSP (FBCSP) [3], followed by conventional classifiers such as Support Vector Machines (SVM) or Linear Discriminant Analysis (LDA). Dimensionality reduction methods (e.g., PCA, LDA projection) are often added to further compress features [11,18]. While these linear methods are simple and effective to some extent, they are limited in capturing the nonlinear and complex structure of EEG data [18]. Deep learning has emerged as a powerful alternative. Autoencoders (AEs), in particular, provide unsupervised nonlinear feature extraction and denoising, yielding compact latent representations that are often more discriminative for classification [8,21]. Yet, the combination of handcrafted spectral–spatial features (e.g., FBCSP), deep representation learning, and Bayesian probabilistic modeling remains underexplored in MI decoding. Gaussian Process Classifiers (GPCs) [22], with their ability to provide calibrated uncertainty estimates, are especially attractive in noisy EEG contexts, but have rarely been integrated into such hybrid frameworks.

To address these limitations, we introduce a unified framework that first extracts spectral–spatial features using FBCSP, compresses them with a deep Autoencoder to obtain nonlinear and noise-reduced representations, and finally classifies with a GPC to achieve probabilistic decision-making. We evaluate the framework on the public BCI Competition IV 2b dataset [20] and the private HCMIU – Motor Imagery Hand–Binary Dataset. Our approach improves balanced accuracy by +5.2% points over a strong baseline (FBCSP+SVM, 77.4%→82.6%) and achieves superior results compared to several recently published deep learning models on BCI Competition IV 2b. In summary, our contributions are as follows.

- We propose a hybrid framework that integrates handcrafted spectral–spatial features, nonlinear representation learning with Autoencoders, and probabilistic classification with Gaussian Processes for MI-EEG decoding.
- We replace manual feature selection with data-driven nonlinear compression while simultaneously providing calibrated uncertainty estimates, improving both accuracy and reliability in noisy EEG settings.
- We conduct extensive experiments on two EEG datasets, demonstrating consistent improvements over strong baselines and competitive or superior performance compared to recent deep learning approaches.

2 Related Works

2.1 Traditional Methods for EEG Classification

Motor-imagery (MI) pipelines have long relied on spatial–spectral features obtained with the Common Spatial Pattern (CSP) algorithm and, more recently, its Filter-Bank extension (FBCSP) [3]. In a typical FBCSP workflow, EEG is decomposed into multiple frequency bands, CSP is applied to each band, and the resulting log-variance features are manually ranked, with a subset heuristically selected—commonly via mutual information or Fisher score—before classification using SVM, Random Forest, or k-Nearest Neighbour [3,12]. However, this manual feature selection is highly sensitive to both selection criteria and the number of retained features, often requiring extensive trial-and-error and expert tuning. Consequently, it risks introducing selection bias—where features optimized on the training set may not generalize robustly to the test set—and can result in models that are highly sensitive to the specific feature selection procedure, limiting reproducibility and reliability [12].

2.2 Dimensionality Reduction and Representation Learning

To overcome these limitations, recent work has shifted towards data-driven dimensionality reduction. While linear methods such as Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) offer simplicity, they fail to capture the complex, nonlinear structures inherent in EEG signals [15]. In contrast, deep Autoencoders (AEs) have emerged as a powerful unsupervised alternative, learning compact, task-relevant latent representations without manual intervention [19]. Notably, AEs automatically discover meaningful nonlinear feature combinations, thereby enhancing robustness and generalization across sessions of the same subject and experimental sessions [13].

2.3 Probabilistic Modeling and Gaussian Process Classifiers

Conventional classifiers such as SVM and LDA provide only deterministic predictions, lacking principled uncertainty quantification—a critical requirement for EEG-based brain–computer interfaces (BCIs), where signals are inherently noisy and data scarcity is common [12]. Gaussian Process Classifier (GPC) offers a rigorous Bayesian alternative, providing calibrated probabilistic predictions crucial for reliable decision-making in BCI applications [22]. Although traditional GPC suffer from scalability issues with high-dimensional inputs [16], coupling GPC with AE-derived compact representations can address these challenges, enabling scalable and uncertainty-aware EEG classification.

2.4 Hybrid Approaches and Research Gap

Although several studies have combined FBCSP-derived features with deep learning methods, they typically retain a separate feature-selection stage or rely

on deterministic classifiers, thus failing to fully leverage probabilistic modeling [1,12]. To the best of our knowledge, no existing approach unifies handcrafted spatial–spectral feature extraction (FBCSP), nonlinear representation learning (Autoencoder), and Bayesian probabilistic classification (Gaussian Process) into a single, fully integrated pipeline.

3 FA-GPNet Architecture

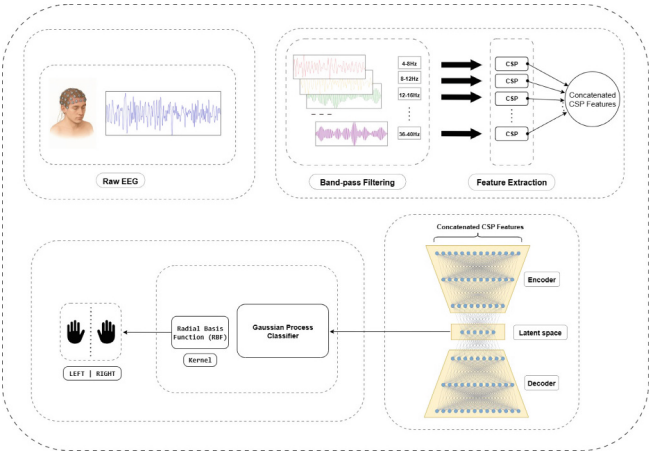

Fig. 1. Overview of FA-GPNet. Raw EEG is band-pass filtered, multi-band CSP features are extracted and concatenated, compressed by an autoencoder, and classified with a Gaussian Process using RBF kernel.

Our proposed FBCSP–AE–GP framework is designed to robustly capture motor imagery spectro-spatial signatures, while effectively addressing the data scarcity and high variability typical of within-subject EEG studies. The overall pipeline is illustrated in Fig. 1. At its core, the architecture proceeds through three sequential stages:

(i) **Spectro-spatial feature extraction:** Standardize and band-limit the raw EEG, then apply FBCSP to extract discriminative features across multiple frequency bands.
(ii) **Nonlinear dimensionality reduction:** Use a deep Autoencoder to obtain compact, task-relevant latent codes.
(iii) **Probabilistic classification:** Employ a Gaussian Process for Bayesian inference and calibrated confidence estimation.

3.1 Spectro-Spatial Feature Extraction

To robustly extract discriminative representations from EEG signals, we adopt the Filter Bank Common Spatial Pattern (FBCSP) technique as the first stage of our framework.

Let $\mathbf{X} \in \mathbb{R}^{C\times T}$ denote the standardized multi-channel EEG segment, where C is the number of channels and T is the number of time points. The raw EEG signal is first decomposed into $B = 9$ non-overlapping frequency sub-bands by a bank of band-pass filters. Specifically, we divide the frequency axis into $B = 9$ non-overlapping sub-bands: 4–8 Hz, 8–12 Hz, 12–16 Hz, 16–20 Hz, 20–24 Hz, 24–28 Hz, 28–32 Hz, 32–36 Hz, and 36–40 Hz, following standard motor imagery BCI protocols:

$$\mathbf{X}^{(b)} = \text{BandpassFilter}_b(\mathbf{X}), \quad b = 1, 2 \ldots, B \tag{1}$$

For each frequency band b, the Common Spatial Pattern (CSP) algorithm is applied to extract spatial filters that maximize variance differences between motor imagery classes. Given the band-limited signal $\mathbf{X}^{(b)}$, CSP solves the following optimization problem:

$$\mathbf{W}^{(b)} = \underset{\mathbf{W}}{\text{argmax}} \frac{\mathbf{W}^\top \mathbf{\Sigma}_1^{(b)} \mathbf{W}}{\mathbf{W}^\top \left(\mathbf{\Sigma}_1^{(b)} + \mathbf{\Sigma}_2^{(b)}\right) \mathbf{W}} \tag{2}$$

where $\mathbf{\Sigma}_1^{(b)}$ and $\mathbf{\Sigma}_2^{(b)}$ are the covariance matrices of the two classes in band b. The spatially filtered signals are then obtained as:

$$\mathbf{Z}^{(b)} = (\mathbf{W}^{(b)})^\top \mathbf{X}^{(b)} \tag{3}$$

From each CSP-projected sub-band signal, we extract *log-variance features* with (K = 2) as follows:

$$f_k^{(b)} = \log\left(\text{var}\left(Z_k^{(b)}\right)\right), \quad k = 1, \ldots, K \tag{4}$$

where $Z_k^{(b)}$ is the output of the k-th CSP component in the b-th frequency band.

Finally, features from all sub-bands and CSP components are concatenated to form a high-dimensional spectro-spatial feature vector:

$$\mathbf{f} = \left[f_1^{(1)}, \ldots, f_K^{(1)}, f_1^{(2)}, \ldots, f_K^{(2)}, \ldots, f_1^{(B)}, \ldots, f_K^{(B)}\right]^\top \in \mathbb{R}^{B\times K} \tag{5}$$

This representation serves as the input for subsequent nonlinear dimensionality reduction.

3.2 Nonlinear Compression with an Autoencoder

To replace manual feature selection and capture nonlinear relationships among FBCSP features, we employ a deep autoencoder (AE) as a dimensionality reduction module. Let $\mathbf{f} \in \mathbb{R}^{d_{\text{in}}}$ denote the FBCSP feature vector of a trial, standardized using z-score normalization (statistics computed on the training split only).

The AE consists of an encoder $E_\phi : \mathbb{R}^{d_{\text{in}}} \to \mathbb{R}^{d_z}$ and a decoder $D_\psi : \mathbb{R}^{d_z} \to \mathbb{R}^{d_{\text{in}}}$, jointly trained to minimize the mean squared reconstruction error:

$$\mathcal{L}_{\text{AE}} = \frac{1}{N} \sum_{i=1}^{N} \|\mathbf{f}_i - D_\psi(E_\phi(\mathbf{f}_i))\|_2^2 .$$

Training. We optimize the AE with Adam (learning rate 10^{-3}, weight decay 10^{-5}) for at most 250 epochs, using early stopping with patience = 10 on validation reconstruction loss; the best checkpoint is restored. Unless otherwise stated, we set the latent dimension to $d_z = 8$. After training, the encoder is frozen and its latent codes are used for probabilistic classification. *The architecture and hyperparameters of the AE are provided in Table 1.*

Table 1. Autoencoder architecture (fully connected).

Block	Layer	Hyper-parameters	Output shape	Activation
I	Input	d_{in} (feature dim)	(N, d_{in})	–
II	Dense1	64 units	$(N, 64)$	Linear
	BatchNorm1d	–	$(N, 64)$	–
	LeakyReLU	$\alpha = 0.2$	$(N, 64)$	LeakyReLU
III	Dense2	32 units	$(N, 32)$	Linear
	BatchNorm1d	–	$(N, 32)$	–
	LeakyReLU	$\alpha = 0.2$	$(N, 32)$	LeakyReLU
IV	Latent	d_z units	(N, d_z)	Linear
V	Dense3	32 units	$(N, 32)$	Linear
	BatchNorm1d	–	$(N, 32)$	–
	LeakyReLU	$\alpha = 0.2$	$(N, 32)$	LeakyReLU
VI	Dense4	64 units	$(N, 64)$	Linear
	BatchNorm1d	–	$(N, 64)$	–
	LeakyReLU	$\alpha = 0.2$	$(N, 64)$	LeakyReLU
VII	Output	d_{in} units	(N, d_{in})	Linear

3.3 Probabilistic Classification with Gaussian Processes

Given AE latents $\mathbf{z}_i \in \mathbb{R}^{d_z}$, we use a Gaussian Process Classifier (GPC) with logistic link:

$$f(\mathbf{z}) \sim \mathcal{GP}\big(0, k_\theta(\mathbf{z}, \mathbf{z}')\big), \qquad y_i \mid f_i \sim \text{Bernoulli}\big(\sigma(f_i)\big), \;\; \sigma(t) = \tfrac{1}{1+e^{-t}}. \tag{6}$$

We adopt an RBF kernel $k_\theta(\mathbf{z}, \mathbf{z}') = \sigma_f^2 \exp\big(-\|\mathbf{z} - \mathbf{z}'\|^2/(2\ell^2)\big)$, whose *hyperparameters* $\theta = \{\sigma_f, \ell\}$ are learned by maximizing the Laplace-approximated marginal likelihood via L-BFGS-B. With the Laplace approximation, letting $\mathbf{W} = \text{diag}\big(\sigma(\hat{f}_i)\,[1 - \sigma(\hat{f}_i)]\big)$ at the posterior mode $\hat{\mathbf{f}}$, the posterior is approximated by

$$q(\mathbf{f}) \approx \mathcal{N}\big(\hat{\mathbf{f}},\, \boldsymbol{\Sigma}\big), \qquad \boldsymbol{\Sigma} = (\mathbf{K}^{-1} + \mathbf{W})^{-1}. \tag{7}$$

Calibration Metrics. We report Expected Calibration Error (ECE) and Maximum Calibration Error (MCE) with $M = 10$ equal-width bins $\{B_m\}_{m=1}^{M}$. For bin B_m of size n_m,

$$\mathrm{acc}(B_m) = \frac{1}{n_m} \sum_{i \in B_m} \mathbf{1}\{\hat{y}_i = y_i\}, \quad \mathrm{conf}(B_m) = \frac{1}{n_m} \sum_{i \in B_m} \hat{p}_i,$$

$$\mathrm{ECE} = \sum_{m=1}^{M} \frac{n_m}{N} \left|\mathrm{acc}(B_m) - \mathrm{conf}(B_m)\right|, \qquad \mathrm{MCE} = \max_m \left|\mathrm{acc}(B_m) - \mathrm{conf}(B_m)\right|.$$

ECE captures *average* miscalibration, MCE highlights the *worst-bin* deviation.

4 Experiments

We extract multi-band FBCSP features and evaluate two selectors, SelectKBest (SKB) and Mutual Information-Based Individual Feature (MIBIF) [2]. Selected features are z-scored and fed to standard classifiers (KNN, SVM-RBF, Random Forest, Naive Bayes, LDA, and GPC-RBF) under the same subject-specific splits. For context only, we also reference a literature-reported deep result (CapsNet; Sensors 2019) [7] rather than re-training deep baselines. Comparative results are presented in Sect. 5.

4.1 BCI Competition IV – Dataset 2b

We evaluated our method on BCI Competition IV – Dataset 2b, a benchmark MI-EEG corpus from Graz University of Technology [20]. Nine healthy volunteers each completed five sessions: two screening sessions without feedback and three feedback sessions with a real-time "smiley" reinforcement (see Fig. 2). EEG was recorded from electrodes C3, Cz, C4 at 250 Hz. Each session comprised left-vs-right hand imagery trials. In screening sessions, the timeline was: 0 s fixation (+) → 3 s auditory cue → 1.25 s arrow cue → 4 s imagery → 1.5 s rest (total 9 s). In smiley-feedback sessions, the fixation was replaced by a neutral grey smiley, and during the imagery period, the smiley provided real-time visual feedback according to the subject's performance. Sessions 1–2 contained 120 trials; 3–5 contained 160 trials. The first three sessions of each subject were used exclusively for training, while the remaining two were reserved for evaluation. In our experiments, we used the 4-second imagery segment from 3 s to 7 s.

4.2 HCMIU – Motor Imagery Hand–Binary Dataset

To validate our method's robustness on small-scale EEG datasets, we employed a hand–binary dataset collected by the Brain Health Lab at Vietnam National University – International University. The dataset comprises EEG recordings from 14 healthy undergraduates performing motor imagery (MI) tasks, captured with a 32-channel Emotiv Flex headset at 128 Hz. Each trial commenced with a

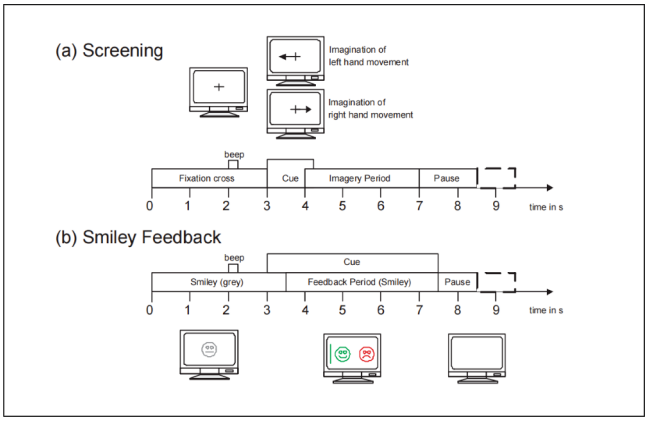

Fig. 2. Protocol in BCI Competition IV 2b: (a) Screening session with directional arrow cue and no feedback, (b) Smiley feedback session with real-time visual reinforcement.

2-second visual cue, followed by an auditory prompt instructing participants to imagine specific motor actions for 4 s without physical execution (see Fig. 3). Participants were randomly divided into three groups: G1 viewed hand movement images; G2 observed cursor images; and G3 observed cursor images with real-time Event-Related Desynchronization (ERD) feedback, enabling performance adjustments. In total, the dataset comprised 1,140 samples (see Table 2), which were analyzed through subject-specific 5-fold cross-validation. In our experiments, we used the 4-second imagery segment from 4 s to 8 s.

5 Results and Discussion

Comparison with FBCSP Pipelines (Table 3). On BCI Competition IV-2b, FA-GPNet achieves the highest average accuracy across subjects (78.19%),

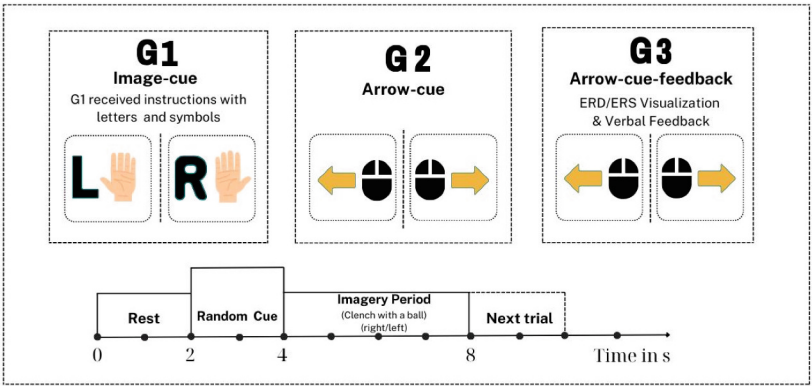

Fig. 3. Protocol in HCMIU – Motor Imagery Hand–Binary Dataset: Image-cue (G1), Arrow-cue (G2), and Arrow-cue-feedback (G3). All cues were displayed in Vietnamese.

Table 2. The HCMIU MI-EEG dataset: Number of samples for each subject.

Subject	F10	F11	F12	F14	F15	F16	F18	F20	F22	F24	F25	F26	F27	F29	Total
Dataset	114	90	90	30	90	60	90	90	96	90	84	84	72	60	1140

outperforming both FBCSP+SKB and FBCSP+MIBIF configurations. Subject-wise breakdown in Table 3 shows that the gains are not driven by a single subject: FA-GPNet improves over traditional pipelines on several challenging subjects (e.g., S2, S3, S5, S7), indicating that the method is robust to inter-subject variability. These results support the design choice of replacing manual feature selection with a deep, nonlinear compression stage (autoencoder) and using a probabilistic classifier (Gaussian Process) that provides calibrated predictions.

Table 3. Per-subject classification accuracy (%) on the BCI Competition IV-2b dataset. Bold denotes the best result per subject; the last row reports the average across subjects.

Subject	SelectKBest (SKB)						MIBIF						FA-GPNet
	KNN	SVM	RF	NB	LDA	GP	KNN	SVM	RF	NB	LDA	GP	
1	64.10	66.20	67.20	66.20	67.20	65.90	61.30	64.40	65.90	64.10	65.60	65.60	**71.88**
2	48.90	46.10	45.70	50.00	51.10	47.10	48.90	46.10	45.70	50.00	51.10	47.10	**59.64**
3	50.60	57.20	55.00	55.90	55.00	55.60	51.90	56.20	55.30	57.20	54.10	50.90	**57.50**
4	85.00	87.20	90.60	86.30	95.00	95.00	86.90	87.80	91.60	86.90	**95.30**	94.40	93.44
5	85.60	86.30	85.90	86.90	88.70	87.50	88.10	86.90	86.90	**89.40**	**89.40**	**89.40**	89.38
6	74.70	78.70	77.20	80.00	78.70	79.10	72.20	77.80	75.90	78.70	77.80	78.40	**83.43**
7	70.30	74.40	77.50	73.10	75.30	75.60	67.50	72.50	73.10	73.80	75.00	75.90	**77.81**
8	74.40	84.70	81.90	79.70	**86.60**	**86.60**	67.80	78.70	74.10	72.80	79.40	72.80	85.31
9	70.30	80.90	78.10	77.20	79.40	80.00	69.10	80.90	77.80	80.00	81.20	80.60	**85.31**
Average	69.32	73.52	73.23	72.81	75.22	74.71	68.19	72.37	71.81	72.54	74.22	72.89	**78.19**

Table 4. Per-subject accuracy (%) on BCI Competition IV-2b. Figures for ShallowNet, DeepNet, EEGNet, CNN, and CapsNet are quoted verbatim from [7] (their original preprocessing/evaluation; no re-training). They are provided for context only and are *not directly comparable* to our FA-GPNet results, which are obtained under a single fixed pipeline.

Subject	ShallowNet	DeepNet	EEGNet	CNN	CapsNet	FA-GPNet(Proposed)
1	71.56	67.25	67.18	69.78	**78.75**	71.88
2	53.57	56.10	58.21	54.75	55.71	**59.64**
3	53.12	54.87	55.62	52.88	55.00	**57.50**
4	**95.93**	94.52	95.31	95.31	**95.93**	93.44
5	85.00	84.59	86.87	85.91	83.12	**89.38**
6	76.87	74.46	77.50	78.03	**83.43**	**83.43**
7	76.56	77.03	76.87	69.75	75.62	**77.81**
8	85.93	87.75	89.68	87.56	**91.25**	85.31
9	82.18	79.25	80.00	80.91	**87.18**	85.31
Average	75.63	75.10	76.36	74.99	**78.44**	78.19

Comparison with Compact Deep Networks (Table 4). Against prior compact CNN baselines reported for context—ShallowNet, DeepNet, EEGNet, standard CNN, and CapsNet—FA-GPNet is competitive with the strongest model: its average accuracy (78.19%) is within 0.25 points of CapsNet (78.44%) while surpassing the other deep baselines in Table 4. Notably, FA-GPNet attains this performance without end-to-end deep classification and while preserving the interpretability of handcrafted spectral–spatial features and the calibrated uncertainty of a Bayesian classifier. This suggests that combining classical priors (FBCSP) with data-driven compression and Bayesian inference can match the accuracy of compact deep models under within-subject protocols, often with more transparent decision factors.

Generalization to HCMIU MI-EEG Dataset (Table 5). On the small, high-variability HCMIU hand–binary MI dataset, FA-GPNet achieves the best average accuracy (57.15%), outperforming all FBCSP+feature-selection baselines in Table 5. The consistent subject-wise gains under scarce labeled data support the claim that the hybrid handcrafted–deep–Bayesian design is well-suited to low-data regimes, where purely supervised deep classifiers tend to overfit or require heavy regularization.

Table 5. The HCMIU hand–binary MI dataset: per-subject accuracy (%) under within-subject 5-fold CV. We compare FBCSP baselines with feature selection (SKB, MIBIF) across classical classifiers against our FA-GPNet. The *Average* row macro-averages across subjects; best per subject is **bold**.

Subject	SelectKBest (SKB)						MIBIF						FA-GPNet
	KNN	SVM	RF	NB	LDA	GP	KNN	SVM	RF	NB	LDA	GP	
F10	57.87	57.04	57.04	57.83	59.60	59.60	54.31	55.30	55.18	50.00	56.09	56.96	**64.00**
F11	55.56	54.44	60.00	47.78	58.89	53.33	50.00	52.22	55.56	50.00	52.22	51.11	**63.33**
F12	50.00	45.56	45.56	47.78	41.11	40.00	46.67	40.00	45.56	42.22	42.22	45.56	**53.33**
F14	36.67	36.67	36.67	33.33	36.67	36.67	43.33	36.67	36.67	36.67	36.67	33.33	**53.33**
F15	52.22	55.56	56.67	54.44	51.11	48.89	54.44	52.22	53.33	51.11	53.33	53.33	**62.22**
F16	38.33	41.67	45.00	41.67	33.33	45.00	43.33	43.33	45.00	41.67	43.33	43.33	**51.67**
F18	54.44	53.33	53.33	53.33	54.44	54.44	52.22	51.11	52.22	51.11	52.22	53.33	**61.11**
F20	53.33	46.67	47.78	46.67	54.44	54.44	54.44	52.22	51.11	51.11	51.11	54.44	**60.00**
F22	41.79	44.63	42.22	42.22	42.22	41.76	42.00	42.74	41.67	42.22	41.19	44.63	**50.00**
F24	51.11	50.00	52.22	51.11	44.44	43.33	47.78	50.00	48.89	48.89	46.67	51.11	**54.44**
F25	54.44	53.33	53.33	51.11	51.11	54.44	52.22	53.33	51.11	52.22	53.33	52.22	**55.96**
F26	53.33	50.00	50.00	53.33	50.00	50.00	53.33	52.22	52.22	51.11	52.22	51.11	**57.06**
F27	51.25	51.11	50.00	53.33	51.11	53.33	50.00	52.22	52.22	52.22	53.33	52.22	**56.95**
F29	51.11	46.67	51.11	52.22	48.89	48.89	54.44	54.44	54.44	51.11	51.11	54.44	**56.67**
Average	50.10	49.04	50.06	49.01	48.38	48.86	49.89	49.14	49.65	47.97	48.93	49.79	**57.15**

Calibration and Uncertainty (Table 6). We report Expected Calibration Error (ECE) and Maximum Calibration Error (MCE) per subject using equal-width binning with $M = 10$ bins. Across the nine IV-2b subjects, the mean ECE is ≈ 0.080 and the mean MCE is ≈ 0.232; five subjects have ECE ≤ 0.10, and two subjects fall below 0.05. Subject 3 exhibits perfect calibration under this metric set (ECE/MCE = 0.000), whereas Subject 4 shows the largest MCE (≈ 0.373), indicating room for improvement on difficult cases.

Table 6. Per-subject calibration on BCI Competition IV-2b: Expected Calibration Error (ECE) and Maximum Calibration Error (MCE) using equal-width binning with $M = 10$ on each subject's test split. Lower is better.

Subject	1	2	3	4	5	6	7	8	9
ECE	0.1641	0.0352	0.0000	0.0589	0.1023	0.1067	0.0681	0.0803	0.1033
MCE	0.3151	0.1797	0.0000	0.3734	0.3047	0.3681	0.1539	0.1549	0.2401

6 Conclusion

We presented FA-GPNet, a novel hybrid framework that integrates FBCSP, Autoencoder, and Gaussian Process for motor imagery EEG decoding. Experiments on the BCI Competition IV-2b and HCMIU - Hand-Binary Motor Imagery dataset show consistent improvements over classical pipelines and competitive performance with compact deep networks. These results highlight the value of combining handcrafted priors, deep representations, and Bayesian inference, opening avenues for more reliable and interpretable BCI systems.

Future Work. We will extend the model to multi-class MI evaluation–e.g., the 4-class BCI Competition IV-2a–and aim to improve performance through data augmentation.

Acknowledgement. We acknowledge Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, for supporting this study, and Nguyen Huu Tuong and Nguyen Hoang Kim Long (School of Biomedical Engineering, International University – VNU-HCM, Ho Chi Minh City, Vietnam) for providing access to the HCMIU–Motor Imagery Hand–Binary Dataset.

References

1. An, Y., Lam, H.K., Ling, S.: Multi-classification for EEG motor imagery signals using data evaluation-based auto-selected regularized FBCSP and convolutional neural network. Neural Comput. Appl. **35**, 1–27 (2023). https://doi.org/10.1007/s00521-023-08336-z
2. Ang, K.K., Chin, Z.Y., Wang, C., Guan, C., Zhang, H.: Filter bank common spatial pattern algorithm on BCI competition iv datasets 2A and 2B. Front. Neurosci. **6** (2012). https://doi.org/10.3389/fnins.2012.00039
3. Ang, K.K., Chin, Z.Y., Zhang, H., Guan, C.: Filter bank common spatial pattern (FBCSP) in brain–computer interface. In: Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN), Hong Kong, China, pp. 2390–2397. IEEE (2008). https://doi.org/10.1109/IJCNN.2008.4634130
4. Antony, M.J., et al.: Classification of EEG using adaptive SVM classifier with CSP and online recursive independent component analysis. Sensors **22**(19), 7596 (2022). https://doi.org/10.3390/s22197596
5. Biasiucci, A., Franceschiello, B., Murray, M.M.: Electroencephalography. Curr. Biol. **29**(3), R80–R85 (2019). https://doi.org/10.1016/j.cub.2018.11.052
6. Chaudhary, U., Birbaumer, N., Ramos-Murguialday, A.: Brain-computer interfaces for communication and rehabilitation. Nat. Rev. Neurol. **12**(9), 513–525 (2016). https://doi.org/10.1038/nrneurol.2016.113
7. woo Ha, K., won Jeong, J.: Motor imagery EEG classification using capsule networks. Sensors **19**(13), 2854 (2019). https://doi.org/10.3390/s19132854
8. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science **313**(5786), 504–507 (2006). https://doi.org/10.1126/science.1127647

9. Huang, X., et al.: A review on signal processing approaches to reduce calibration time in EEG-based brain-computer interface. Front. Neurosci. **15**, 733546 (2021). https://doi.org/10.3389/fnins.2021.733546
10. Jiang, J., Wang, C., Wu, J., Qin, W., Xu, M., Yin, E.: Temporal combination pattern optimization based on feature selection method for motor imagery BCIs. Front. Hum. Neurosci. **14**, 231 (2020). https://doi.org/10.3389/fnhum.2020.00231
11. Kabir, M.H., Mahmood, S., Al Shiam, A., Musa Miah, A.S., Shin, J., Molla, M.K.I.: Investigating feature selection techniques to enhance the performance of EEG-based motor imagery tasks classification. Mathematics **11**(8), 1921 (2023). https://doi.org/10.3390/math11081921
12. Lotte, F., et al.: A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. J. Neural Eng. **15**(3), 031005 (2018). https://doi.org/10.1088/1741-2552/aab2f2
13. Mammone, N., Ieracitano, C., Adeli, H., Morabito, F.C.: Autoencoder filter bank common spatial patterns to decode motor imagery from EEG. IEEE J. Biomed. Health Inform. **27**(5), 2365–2376 (2023). https://doi.org/10.1109/JBHI.2023.3243698
14. Pfurtscheller, G., Neuper, C.: Motor imagery and direct brain-computer communication. Proc. IEEE **89**(7), 1123–1134 (2001). https://doi.org/10.1109/5.939829
15. Rashid, M., et al.: Current status, challenges, and possible solutions of EEG-based brain-computer interface: a comprehensive review. Front. Neurorobot. **14**, 25 (2020). https://doi.org/10.3389/fnbot.2020.00025
16. Rasmussen, C.E., Williams, C.K.I.: Gaussian Processes for Machine Learning. The MIT Press (2005). https://doi.org/10.7551/mitpress/3206.001.0001
17. Shiam, A.A., Hassan, K.M., Islam, M.R., Almassri, A.M.M., Wagatsuma, H., Molla, M.K.I.: Motor imagery classification using effective channel selection of multichannel EEG. Brain Sci. **14**(5), 462 (2024). https://doi.org/10.3390/brainsci14050462
18. Singh, A., Hussain, A.A., Lal, S., Guesgen, H.W.: A comprehensive review on critical issues and possible solutions of motor imagery based electroencephalography brain-computer interface. Sensors **21**(6), 2173 (2021). https://doi.org/10.3390/s21062173
19. Tang, R., Li, Z., Xie, X.: Motor imagery EEG signal classification using upper triangle filter bank auto-encode method. Biomed. Signal Process. Control **68**, 102608 (2021). https://doi.org/10.1016/j.bspc.2021.102608
20. Tangermann, M., et al.: Review of the BCI competition IV. Front. Neurosci. **6** (2012). https://doi.org/10.3389/fnins.2012.00055
21. Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.A.: Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th International Conference on Machine Learning (ICML), Helsinki, Finland, pp. 1096–1103. ACM (2008). https://doi.org/10.1145/1390156.1390294
22. Zhong, M., Lotte, F., Girolami, M., Lécuyer, A.: Classifying EEG for brain computer interfaces using gaussian processes. Pattern Recogn. Lett. **29**(3), 354–359 (2008). https://doi.org/10.1016/j.patrec.2007.10.009