PATCHCODE: Discrete Latent Predictive Learning for EEG Foundation Model

Kieren Yu [1] Ziyang Liu [1] Chang Huang [1] Kaishun Wu [1]

[1]The Hong Kong University of Science and Technology (Guangzhou), Guangzhou, China

Problem and Core Idea

EEG foundation models face **low signal-to-noise ratio**, heterogeneous montages, and cross-subject variability. Raw waveform reconstruction can reward memorizing artifacts rather than stable neural dynamics.

PATCHCODE changes the target: the encoder receives continuous EEG patches, while masked predictions are supervised by **noise-filtered discrete codes**.

It learns **what should be predictable**, not every noisy sample.

Contributions

- **Two-stage learning:** Stage A discovers discrete EEG codes; Stage B trains predictive representations.
- **Continuous input, discrete supervision:** Discrete codes are targets only; they are never fed into the Student encoder.
- **Spectral tokenizer:** FFT amplitude/phase supervision preserves neural rhythms.
- **Multi-pathway prediction:** P2P, P2C, and C2P combine continuous, categorical, and geometric supervision.

PATCHCODE Framework Overview

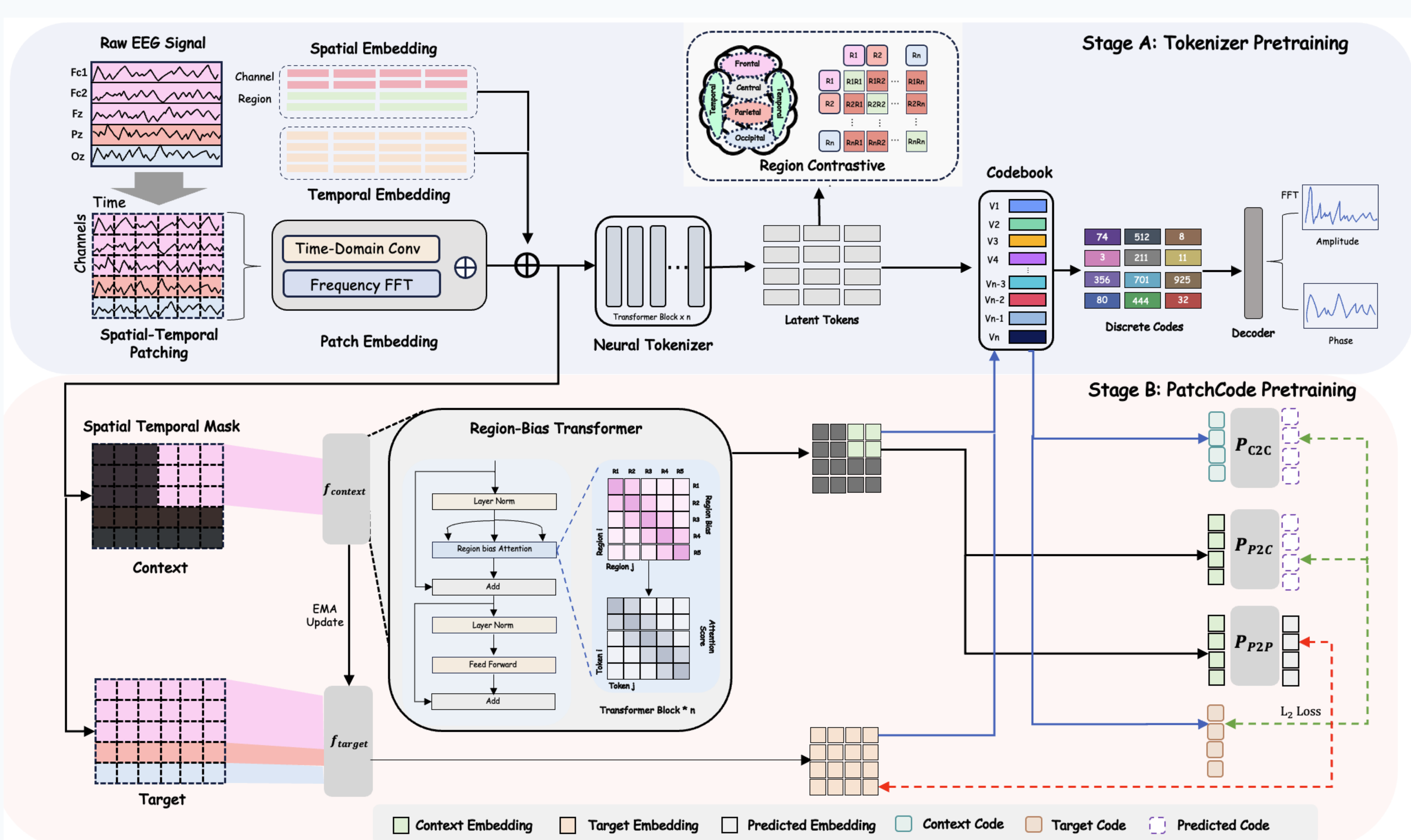

Figure 1. **PATCHCODE Framework.** Stage A trains a region-aware tokenizer with spectral reconstruction. Stage B employs Student-Teacher architecture with JEPA-style objectives (P2P, P2C, C2P).

The frozen tokenizer provides **discrete labels** for masked positions; the Student predicts both **Teacher embeddings** and **tokenizer code indices** from continuous context.

Stage A: Region-Aware Tokenizer

Frequency-Domain Discrete Code Discovery

Each 1-second EEG patch is encoded and assigned to a **nearest codebook entry**:

$$q_i = \arg\min_k \|\mathbf{z}_i - \mathbf{e}_k\|_2, \qquad \mathbf{z}_i^q = \mathbf{e}_{q_i}, \quad K = 8192.$$

The decoder reconstructs **spectral amplitude and phase**:

$$\mathcal{L}_{spec} = \frac{1}{L}\sum_{i=1}^{L}\left(\|\widehat{\mathbf{A}}_i - \mathbf{A}_i\|_2^2 + \|\widehat{\mathbf{\Phi}}_i - \mathbf{\Phi}_i\|_2^2\right).$$

Stable EEG Vocabulary

VQ commitment, **spectral reconstruction**, region contrast, and code usage are optimized jointly:

$$\mathcal{L}_A = \mathcal{L}_{vq} + \lambda_{spec}\mathcal{L}_{spec} + \lambda_{reg}\mathcal{L}_{reg} - \lambda_{usage}\mathcal{L}_{usage}.$$

Stage B: Multi-Pathway Predictive Learning

Student-Teacher Predictive Encoder

The Student receives masked **continuous EEG patches**; the EMA Teacher provides stable targets:

$$\bar{\theta} \leftarrow \tau\bar{\theta} + (1-\tau)\theta, \quad \tau : 0.996 \rightarrow 0.9999$$

Region-bias attention adds anatomical priors.

Three Complementary Losses

P2P matches **Teacher features**:

$$\mathcal{L}_{P2P} = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}} \|\mathrm{Pred}_{P2P}(\mathbf{h}_i^s) - \mathrm{sg}(\mathbf{h}_i^t)\|_2^2.$$

P2C supplies the **core discrete supervision**:

$$\mathcal{L}_{P2C} = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}} \mathrm{CE}(\mathrm{Cls}_{P2C}(\mathbf{h}_i^s), q_i).$$

C2P aligns **code embeddings** with Student representations:

$$\mathcal{L}_{C2P} = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}} \|\mathrm{Pred}_{C2P}(\mathbf{e}_{q_i}) - \mathbf{h}_i^s\|_2^2.$$

PATCHCODE Downstream Evaluation

Task	Dataset	Balanced Acc.	Metric 2	Metric 3
Emotion	SEED-V	**0.4346±0.0018**	κ 0.2594±0.0289	W-F1 0.4377±0.0193
Emotion	FACED	**0.5950±0.0072**	κ 0.5124±0.0089	W-F1 0.5888±0.0078
Imagined speech	BCIC2020-3	**0.6189±0.0105**	κ 0.5240±0.0138	W-F1 0.6193±0.0112
Motor imagery	BCI-IV-2a	**0.5842±0.0076**	κ 0.4452±0.0085	W-F1 0.5815±0.0081
Sleep staging	HMC	**0.7556±0.0041**	κ 0.7052±0.0045	W-F1 0.7715±0.0042
Sleep staging	ISRUC	**0.8120±0.0068**	κ 0.7520±0.0085	W-F1 0.8095±0.0075
Seizure detection	CHB-MIT	**0.7620±0.0128**	AUC-ROC 0.9085±0.0095	AUC-PR 0.3910±0.0152
Abnormal detection	TUAB	**0.8250±0.0035**	AUC-ROC 0.9102±0.0038	AUC-PR 0.9125±0.0052
Event typing	TUEV	**0.7080±0.0052**	κ 0.7025±0.0065	W-F1 0.8495±0.0045
Stress	MentalArithmetic	**0.7694±0.0091**	AUC-ROC 0.8632±0.0217	AUC-PR 0.6915±0.0184
Diagnosis	Mumtaz2016	**0.9721±0.0083**	AUC-ROC 0.9971±0.0017	AUC-PR 0.9952±0.0028

Ablation Study

Ablation	Setting	FACED	BCI-IV-2a	ISRUC	Avg.
Codebook size	$K = 1024$	.545	.802	.775	.707
	$K = 2048$	.563	.815	.788	.722
	$K = 4096$	.578	.825	.802	.735
	$K = 8192$	**.595**	**.835**	**.812**	**.747**
Spectral bins	$P_f = 25$	.572	.820	.786	.726
	$P_f = 50$	**.595**	**.835**	**.812**	**.747**
	$P_f = 75$	.592	.833	.809	.745
Mask ratio	20%	.568	.812	.785	.722
	40%	**.595**	**.835**	**.812**	**.747**
	60%	.572	.820	.795	.729
Region bias	Disabled	.583	.821	.805	.736
	Enabled	**.595**	**.835**	**.812**	**.747**

Best recipe: $K = 8192$, **50 bins, 40% masking, region bias enabled**. Larger codebooks help separate **sleep rhythms**, **motor patterns**, and **affective signatures**.

50 bins preserve frequency detail; **40% masking** balances difficulty and context.

Interpretable Attention

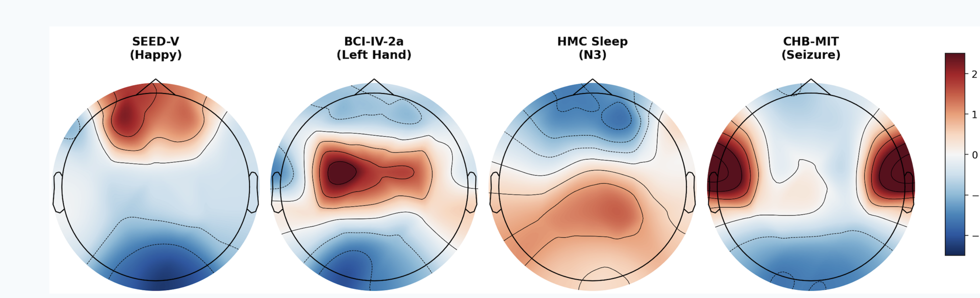

Figure 2. **Task-Specific Spatial Selectivity.** Attention maps align with known regional signatures across emotion, motor imagery, sleep, and seizure detection.

Attention is **task-specific**: emotion emphasizes frontal/temporal regions, motor imagery highlights sensorimotor areas, and seizure detection localizes high-response regions.

This supports **brain-region bias** and **region-guided masking**.

Prediction Pathway

Ablation logic: each row removes exactly one predictive route while keeping the same tokenizer and encoder. The full design is consistently best because **the three routes supervise different aspects of masked EEG**.

SEED-V					HMC					BCI-2a				
Variant	P2P	P2C	C2P	**Acc.**	**Variant**	P2P	P2C	C2P	**Acc.**	**Variant**	P2P	P2C	C2P	**Acc.**
Full	✓	✓	✓	**.414**	Full	✓	✓	✓	**.756**	Full	✓	✓	✓	**.584**
w/o C2P	✓	✓	×	.407	w/o C2P	✓	✓	×	.742	w/o C2P	✓	✓	×	.572
w/o P2C	✓	×	✓	.410	w/o P2C	✓	×	✓	.745	w/o P2C	✓	×	✓	.574
w/o P2P	×	✓	✓	.412	w/o P2P	×	✓	✓	.748	w/o P2P	×	✓	✓	.578

Removing any pathway reduces accuracy: **P2P** preserves continuous dynamics, **P2C anchors predictions to discrete neural codes**, and **C2P** keeps code embeddings geometrically aligned with Student states.

kyu219@connect.hkust-gz.edu.cn