HiMDA: Transferring Hierarchical Microstate Dynamics, Not Labels, for Cross-Task EEG Decoding

Anonymous authors
Paper under double-blind review

Abstract

Cross-task EEG decoding requires transferring knowledge between datasets whose label spaces share no meaningful correspondence, as when operational states must inform emotion recognition or clinical outcome prediction. Conventional feature alignment can conflate acquisition variation with neural temporal organization, whereas transferring a classifier imposes source semantics on an unrelated target task. We propose Hierarchical Microstate Dynamics Adaptation (HiMDA), a label-space-agnostic framework that transfers hierarchical temporal dynamics rather than labels or decision boundaries. HiMDA first constructs a domain-specific observation interface and calibrates polarity-invariant target microstate prototypes under a finite source reference. It then fits Bayesian switching dynamical systems independently in each domain, aligns their slow regimes through label-free dynamics fingerprints, and lets soft regime posteriors condition time-local fast microstate switches. Source slow-regime and regime-conditioned fast-transition operators parameterize finite Dirichlet priors whose posterior means are updated row-wise by target evidence, so well-supported target dynamics depart from the source while sparsely observed transitions remain stabilized. A random-forest readout trained from scratch on the adapted representation defines the target semantics, and no source logit or prediction head is transferred. Under a subject-dependent, trial-level five-fold protocol on SEED-IV, HiMDA achieves $77.04\% \pm 1.09\%$ accuracy. Under patient-disjoint, leave-one-hospital-out evaluation on I-CARE, it attains an AUROC of 0.85 and an AUPRC of 0.90. Together, these results show that hierarchical dynamics transfer across affective and clinical EEG tasks without source–target class correspondence.

1 Introduction

EEG provides millisecond-scale access to distributed brain activity, yet models trained on one dataset often degrade when applied to another. Montage, reference, sampling rate, filtering, equipment, acquisition context, and population can all alter the observed distribution, so even recordings collected under similar protocols rarely align directly.

Cross-task transfer introduces a deeper difficulty, because source and target datasets may describe fundamentally different phenomena. Operational states, discrete emotions, and clinical outcomes do not share a defensible class ontology. A useful transfer method must therefore jointly address acquisition shifts and label spaces with no meaningful correspondence, without assuming source–target class correspondence.

Most existing transfer methods do not separate these two problems: feature alignment can conflate measurement variation with neural temporal organization, while class prototypes, logits, or decision boundaries impose source semantics on an unrelated target task. Self-supervised and foundation models improve reuse across heterogeneous EEG collections Foumani et al. (2024); Wang et al. (2024), but their interface is still mainly a feature space or parameter initialization and does not distinguish sensor-level shifts from changes in the temporal process.

We instead ask whether temporal organization itself can provide a transferable object without imposing semantic correspondence. EEG microstates summarize brief, quasi-stable scalp topographies

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

whose temporal statistics characterize fast whole-scalp organization Tarailis et al. (2024), while recent work studies the sequential syntax formed by successive states Artoni et al. (2023). At a slower scale, latent-state models describe recurrent dynamical regimes, and the two descriptions capture complementary temporal scales Coquelet et al. (2022). Existing pipelines nevertheless analyze them separately and provide no explicit transfer mechanism in which slow contexts condition fast switches across datasets. A global microstate transition matrix cannot express that the same fast switch may behave differently under distinct slow contexts.

Hierarchical Microstate Dynamics Adaptation (HiMDA) addresses this gap through an ordered mechanism chain. First, a target-specific observation interface accounts for montage and reference variation and estimates axial microstate maps, for which a topography and its polarity reversal are equivalent; a finite prototype penalty preserves source-referenced state correspondence while allowing target topographies to adapt. Second, Bayesian switching dynamical systems (BSDSs) are fitted independently in the source and target domains, their slow regimes are aligned through label-free dynamics fingerprints, and soft regime posteriors weight time-local microstate switches, producing fast-transition operators conditioned on slow context rather than an independent concatenation of fast and slow summaries. Third, source slow-regime and regime-conditioned fast-transition operators parameterize finite Dirichlet priors, and target transition counts update these priors row by row, allowing well-supported target dynamics to depart from the source while stabilizing sparsely observed transitions. Finally, a random forest is trained from scratch on the adapted target representation. No source class label, class logit, or prediction head is transferred to define target semantics.

The finite-prior formulation addresses the bias–variance tension in cross-dataset dynamics: freezing source rows ignores target change, whereas target-only estimates discard structure when counts are sparse. HiMDA expresses source influence as finite equivalent counts, so target evidence dominates well-supported rows while the prior regularizes limited observations.

Overall, our key contributions are as follows:

- We formulate cross-task EEG transfer as the transfer of hierarchical dynamics rather than labels: source-referenced axial prototype adaptation learns target microstate maps through a domain-specific observation interface and a polarity-invariant finite reference, without requiring any source–target label mapping.
- We construct BSDS-conditioned hierarchical dynamics in which soft slow-regime posteriors weight time-local fast microstate switches, capturing context-specific transition structure beyond independent multiscale summaries.
- We derive finite source-centered Dirichlet updates for slow- and fast-scale transition operators, allowing target evidence to override well-supported rows while regularizing sparse ones; a newly fitted target classifier preserves independence from source label semantics.
- We provide unified evidence on affective and clinical targets, with protocol-matched comparisons on SEED-IV and patient-disjoint, hospital-held-out evaluation on I-CARE, together with one-mechanism-at-a-time ablations that assess individual components.

2 Related Work

2.1 EEG Representation Transfer

EEG recognition has moved from handcrafted descriptors toward learned representations. EEGNet supports efficient decoding Lawhern et al. (2018), domain-adversarial networks reduce domain-specific information through distributional alignment Ganin et al. (2016), and spatiotemporal models exploit electrode relations and temporal structure Li et al. (2023). Self-supervised learning broadens this interface: BENDR learns contrastive contextual features Kostas et al. (2021), BIOT supports cross-data biosignal learning Yang et al. (2023), EEG2Rep reconstructs informative masked inputs Foumani et al. (2024), and robust representations support EEG emotion recognition Liu et al. (2024). Foundation models including EEGPT, Neuro-GPT, CBraMod, and EEGMamba further target reuse across recording settings and downstream tasks Wang et al. (2024); Cui et al. (2024); Wang et al. (2025a;b). These methods transfer features or parameter initializations and still assume that a shared representation can map to each task's labels. They do not specify what structure persists when label

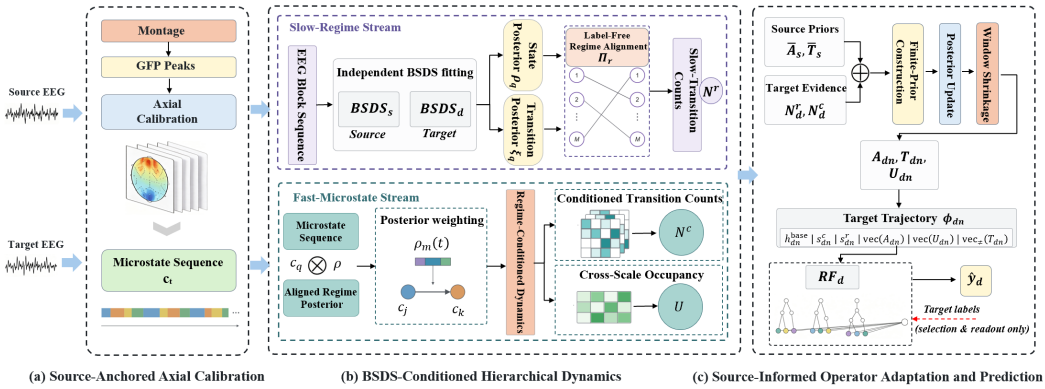

Figure 1: Overview of HiMDA. (a) Source-anchored axial calibration establishes polarity-invariant microstate correspondence under a finite source reference. (b) Independent BSDS models are aligned through label-free regime fingerprints; their posteriors yield slow-transition counts N^r, conditioned fast-transition counts N^c, and occupancy U. (c) Source operators define finite Dirichlet priors updated by target evidence to produce ϕ_{dn}; target labels enter only model selection and the task-specific readout.

spaces are unrelated, nor do they separate acquisition shifts from changes in the underlying temporal process.

2.2 Microstates and Hierarchical Brain Dynamics

EEG microstates describe approximately stable scalp fields Pascual-Marqui et al. (1995); polarity reversal therefore motivates axial map comparison. Standard statistics include occupancy, duration, and transitions Tarailis et al. (2024), with recent work studying sequential syntax Artoni et al. (2023). Switching state-space and probabilistic BSDS models capture slower recurrent regimes and uncertainty Linderman et al. (2017); Taghia et al. (2018); the scales are complementary Coquelet et al. (2022) but usually analyzed separately. HiMDA couples them by using slow posteriors to weight time-local switches and by adapting both scales as transition operators.

2.3 Transfer without Label Correspondence

Most domain-adaptation methods assume compatible class semantics. Partial, open-set, and universal variants still presume a shared subset or rejection class You et al. (2019); source-free adaptation removes source data but not source-task semantics Fang et al. (2024). Cross-task EEG transfer is stricter: class-conditional alignment, source prototypes, or a reused head can impose invalid correspondence when no target category has a source counterpart. In that setting, reusing a class prototype or decision head does not merely add noise; it changes the meaning assigned to the target representation. HiMDA instead estimates observation models separately, transfers hierarchical operators, and lets target labels enter only model selection and a new readout, not microstate estimation, BSDS fitting, regime matching, or count estimation. Source dynamics can therefore inform the target representation without defining its class geometry. This also differs from within-cohort handcrafted outcome prediction after cardiac arrest Amorim et al. (2023a): we test dynamics from an unrelated resting-state cohort under patient-disjoint evaluation.

3 Method

3.1 Overview and Problem Formulation

We consider cross-task transfer between a source EEG domain and a target EEG domain whose label spaces need not correspond. Let $q \in \{s, d\}$ index the source and target domains, respectively. The n-th analysis window, $X_{qn} \in \mathbb{R}^{C_q \times L_{qn}}$, contains C_q channels and L_{qn} frames. Target windows

have labels $y_{dn} \in \mathcal{Y}_d$. HiMDA requires no mapping, shared ordering, or semantic correspondence between $\mathcal{Y}_d$ and the source label space.

As summarized in Figure 1, HiMDA first establishes source-referenced microstate identities under domain-specific observations, then conditions fast microstate switches on independently inferred slow regimes, and finally updates finite source priors with target evidence. The model contains K fast microstates and M slow regimes. Its transferable quantities are the slow-regime operator $\bar{\mathbf{A}}_q \in [0,1]^{M\times M}$ and the regime-conditioned microstate operator $\bar{\mathbf{T}}_q \in [0,1]^{M\times K\times K}$; observation transforms, prototypes, and BSDS parameters remain domain specific.

We use *label-space-agnostic* to mean that no source–target label mapping is assumed and that target labels enter only nested model selection and the final readout. Microstate estimation, BSDS fitting, regime matching, and transition-count estimation are label-free.

3.2 Source-Anchored Axial Calibration

Montage and reference differences are handled before state identities are compared. Each observed scalp frame is transformed into a common channel geometry, normalized, and compared using a polarity-invariant distance:

$$a_{qnt} = H_q R_q(\widetilde{x}_{qnt}), \qquad x_{qnt} = \frac{a_{qnt}}{\|a_{qnt}\|_2}, \\ d_{\mathrm{ax}}(u,v) = 1 - \left(u^\top v\right)^2, \tag{1}$$

where $\widetilde{x}_{qnt} \in \mathbb{R}^{C_q}$ is the observed scalp frame, R_q applies the prescribed re-referencing and channelwise scaling estimated from domain-q training data, and H_q is a fixed selection or interpolation operator constructed from montage metadata. The resulting vector a_{qnt} is normalized to obtain the unit-norm scalp map x_{qnt}. Both operators are frozen before validation or test data are transformed, and near-zero-norm frames are discarded. For unit-norm maps, the axial distance is invariant to global polarity reversal because $d_{\mathrm{ax}}(u,v) = d_{\mathrm{ax}}(u,-v)$. Source prototypes $\widehat{\boldsymbol{\mu}}_s = \{\widehat{\mu}_{sk}\}_{k=1}^K$ are estimated from source-training global-field-power peaks. Target-only axial clustering is first run from multiple initializations, and the solution with the lowest within-target objective is retained. A Hungarian assignment Kuhn (1955) then matches the initial target clusters to the source prototypes using a cost matrix whose (k,j)-th entry is $d_{\mathrm{ax}}(\widetilde{\mu}_{dk}^{(0)}, \widehat{\mu}_{sj})$. This determines the initial source-referenced ordering without using task labels. Starting from this matched initialization, the target prototypes are calibrated by

$$\widehat{\boldsymbol{\mu}}_d = \arg\min_{\boldsymbol{\mu}_d \in \mathbb{S}^K} \left\{ \frac{1}{|\mathcal{S}_d^{\mathrm{GFP}}|} \sum_{(n,t)\in\mathcal{S}_d^{\mathrm{GFP}}} \min_k d_{\mathrm{ax}}(x_{dnt}, \mu_{dk}) + \frac{\lambda_\mu}{K} \sum_{k=1}^K d_{\mathrm{ax}}(\mu_{dk}, \widehat{\mu}_{sk}) \right\}, \tag{2}$$

where $\boldsymbol{\mu}_d = \{\mu_{dk}\}_{k=1}^K$ is the target prototype set, $\mathcal{S}_d^{\mathrm{GFP}}$ contains target-training global-field-power peak frames, and $\mathbb{S} = \{u : \|u\|_2 = 1\}$ is the unit sphere. The coefficient λ_μ controls the finite source reference. The first term fits target-specific scalp topographies, whereas the second preserves source-referenced state correspondence without forcing source and target prototypes to coincide.

Equation (2) is solved by alternating frame assignment and prototype update. Each target GFP peak is first assigned to its nearest current prototype under d_{ax}. For each state, the prototype is then updated as the leading eigenvector of the sum of two weighted outer-product matrices: the assigned target frames contribute with weight $1/|\mathcal{S}_d^{\mathrm{GFP}}|$, and the corresponding source prototype contributes with weight λ_μ/K. The eigenvector sign is immaterial under the axial distance. We use a weak anchoring strength of $\lambda_\mu = 0.20$. The two steps alternate until the assignments stabilize, the relative objective decrease falls below $\epsilon_\mu = 10^{-5}$, or the maximum iteration count $I_\mu = 100$ is reached.

Finally, every valid frame is assigned to the nearest domain-specific prototype under d_{ax}, producing the microstate sequence $c_{qnt} \in \{1,\dots,K\}$. Prototype estimation, source–target matching, and frame-level assignment use only the corresponding training pool.

3.3 BSDS-Conditioned Hierarchical Dynamics

Microstates characterize rapid frame-level changes in scalp topography but do not identify the slower dynamical context in which these changes occur. We therefore divide each EEG window into B_{qn}

contiguous blocks. Training-fold standardization followed by a training-fitted linear projection maps every block to a D-dimensional observation v_{qnb}. The resulting sequence is provided to the slow-dynamics model.

A Bayesian switching dynamical system (BSDS) models the block sequence through recurrent latent regimes Taghia et al. (2018). Let $r_{qnb} \in \{1, \ldots, M\}$ denote the slow regime of block b and z_{qnb} its latent dynamical factor. Conditioned on $r_{qnb} = m$, the block-level process is

$$\begin{aligned} z_{qnb} &= \sum_{\ell=1}^{P} F_{qm}^{(\ell)} z_{qn,b-\ell} + \epsilon_{qnb}, \\ v_{qnb} &= L_{qm} z_{qnb} + \nu_{qnb}, \end{aligned} \tag{3}$$

where P is the autoregressive order, $F_{qm}^{(\ell)}$ specifies the regime-dependent latent dynamics, L_{qm} maps the latent factor to the block observation, and ϵ_{qnb} and ν_{qnb} are Gaussian process and observation noise. The regime sequence follows a first-order Markov process. Each regime therefore represents a recurrent mode of block-level EEG evolution rather than a task class or deterministic segment. The latent dimension, autoregressive order, and number of regimes are selected within the training pool.

Variational smoothing preserves uncertainty over the slow regime of each block through the single-slice and two-slice posterior probabilities

$$\begin{aligned} \rho_{qnb}(m) &= p(r_{qnb} = m \mid V_{qn}), \\ \xi_{qnb}(m, m') &= p(r_{qn,b-1} = m, r_{qnb} = m' \mid V_{qn}), \end{aligned} \tag{4}$$

where V_{qn} denotes the complete block-observation sequence, $\rho_{qnb}(m)$ is the posterior probability that block b occupies regime m, and $\xi_{qnb}(m, m')$ is the joint posterior probability of a transition from m to m' between adjacent blocks. The single-slice posterior supplies a soft slow-context weight for frame-level events, whereas the two-slice posterior supplies expected evidence for slow-regime transitions.

Source and target BSDS models are fitted independently using their respective training data. This preserves domain-specific latent dynamics but also means that their regime indices and factor coordinates are not directly comparable. We therefore align slow regimes through their observable relationships with the already matched microstates rather than through BSDS emission coordinates.

For each local regime m, let $\mathbf{o}_{qm}$ denote its posterior-weighted microstate-occupancy distribution, δ_{qm} its standardized dwell time, and $\mathbf{g}_{qm}$ its normalized distribution of regime-conditioned off-diagonal microstate switches. Together, these quantities define a label-free regime fingerprint. The source–target matching cost is

$$C_{mm'}^{r} = \frac{1}{3}\Big[D_{\mathrm{JS}}(\mathbf{o}_{sm}, \mathbf{o}_{dm'}) + |\delta_{sm} - \delta_{dm'}| + D_{\mathrm{JS}}(\mathbf{g}_{sm}, \mathbf{g}_{dm'})\Big], \tag{5}$$

where D_{JS} denotes the Jensen–Shannon divergence Lin (1991), m indexes a source regime, and m' indexes a target regime. The three terms compare complementary observable properties: microstate occupancy, temporal persistence, and regime-conditioned fast-switching structure. The occupancy and switching terms operate on normalized distributions, while dwell time is standardized using training-fold statistics. Equal weighting encodes no prior preference among the three properties and avoids tuning regime correspondence with target class labels.

The Hungarian algorithm Kuhn (1955) applied to $\mathbf{C}^r$ yields the source–target correspondence Π_r. Target posterior quantities are subsequently reordered by Π_r, after which ρ and ξ denote the aligned posteriors. This alignment uses matched microstate identities and observable dynamics fingerprints, but neither task labels nor domain-specific BSDS emission coordinates.

The aligned two-slice posterior provides expected slow-regime transition counts.

$$N_{qn,mm'}^{r} = \sum_{b=2}^{B_{qn}} \xi_{qnb}(m, m'), \tag{6}$$

where $N_{qn,mm'}^{r}$ is the posterior-expected number of transitions from slow regime m to m' in window n.

To connect slow BSDS regimes with fast microstate dynamics, let $b_{qn}(t)$ identify the block containing destination frame t. The aligned single-slice posterior weights every observed microstate switch by its local slow context:

$$N^c_{qn,mjk} = \sum_{t \in \mathcal{E}_{qn}} \rho_{qn,b_{qn}(t)}(m) \mathbf{1}[c_{qn,t-1} = j,\ c_{qnt} = k], \qquad j \neq k, \tag{7}$$

where $\mathcal{E}_{qn}$ contains valid consecutive frame pairs, c_{qnt} is the aligned microstate assignment of frame t, and $N^c_{qn,mjk}$ is the posterior-weighted count of transitions from microstate j to microstate k under slow regime m. The destination block supplies the context when a microstate transition crosses a slow-block boundary. Fast self-transitions are excluded so that microstate dwell time remains a separate descriptor.

The same posterior weights define $U_{qn,mk}$ as the normalized co-occurrence of slow regime m and microstate k. Thus, N^r, N^c, and U describe slow transitions, slow-conditioned fast switches, and cross-scale occupancy, respectively. For target windows, U_{dn} is computed directly from the target posterior and receives no source-prior update.

3.4 Source-Informed Operator Adaptation and Prediction

Training-window counts are pooled within each domain. Because overlapping windows may contain the same underlying transition event, directly summing their counts would overrepresent repeatedly covered time points. Let $\mathcal{W}_{qn}$ denote the temporal support of window n, and let e denote an underlying fast frame-pair or slow-block transition event. Its coverage count and inverse-coverage weight are

$$\kappa_q(e) = \sum_n \mathbf{1}[e \in \mathcal{W}_{qn}], \qquad \omega_q(e) = \frac{1}{\kappa_q(e)}, \tag{8}$$

where $\kappa_q(e)$ is the number of windows containing event e. Multiplying each occurrence by $\omega_q(e)$ gives every underlying event a total pooled weight of one. After coverage correction, a small symmetric pseudocount is added to each admissible transition. The resulting target counts are denoted by N^r_d and N^c_d; row-normalized source counts define $\bar{\mathbf{A}}_s$ and $\bar{\mathbf{T}}_s$.

HiMDA uses each source row as the mean of a finite Dirichlet prior. The prior concentrations are $\tau_A \bar{\mathbf{A}}_{s,m\cdot}$ for a slow-regime transition row and $\tau_T \bar{\mathbf{T}}_{s,mj\cdot}$ for a conditioned microstate-transition row. The symmetric pseudocount ensures that all admissible concentrations remain positive. Updating these priors with pooled target evidence gives

$$\begin{aligned} \bar{A}_{d,mm'} &= \frac{N^r_{d,mm'} + \tau_A \bar{A}_{s,mm'}}{N^r_{d,m\cdot} + \tau_A}, \\ \bar{T}_{d,mjk} &= \frac{N^c_{d,mjk} + \tau_T \bar{T}_{s,mjk}}{N^c_{d,mj\cdot} + \tau_T}, \qquad j \neq k, \end{aligned} \tag{9}$$

where $N^r_{d,m\cdot}$ and $N^c_{d,mj\cdot}$ are the corresponding target row totals, and τ_A and τ_T are source equivalent sample sizes for the slow and conditioned fast operators. Rows with abundant target evidence can depart from the source, whereas sparse rows retain stronger source influence. Setting $\tau_A = \tau_T = 0$ yields matched target-only operator estimates without changing the representation dimension. We use fixed equivalent sample sizes of $\tau_A = 20$ and $\tau_T = 10$, providing stronger stabilization for the more sparsely observed slow-regime transitions without scaling the prior by the original source sample size.

Short windows may contain too few transitions for stable window-specific operators. HiMDA therefore performs a second posterior-mean update using the adapted target-domain operator as the prior mean:

$$Q_{dn,i\ell} = \frac{N^Q_{dn,i\ell} + \eta_Q \bar{Q}_{d,i\ell}}{N^Q_{dn,i\cdot} + \eta_Q}, \qquad Q \in \{A, T\}, \tag{10}$$

where N^Q_{dn} is the coverage-corrected window-level count table, $\bar{Q}_d$ is the adapted target-domain operator, and η_Q controls window-level shrinkage. For $Q = A$, $N^Q_{dn} = N^r_{dn}$, and i indexes a slow regime. For $Q = T$, $N^Q_{dn} = N^c_{dn}$, and i indexes a slow-regime–microstate context. In both cases,

ℓ indexes an admissible destination state. We set the fixed window-level equivalent sample sizes to $\eta_A = 8$ and $\eta_T = 4$. Empty rows revert to the corresponding target-domain operator, whereas rows containing approximately 10–20 observed transitions are already governed primarily by window-specific evidence. The resulting operators are denoted by A_{dn} and T_{dn}. The final target-window representation is

$$\phi_{dn} = \left[h_{dn}^{\text{base}};\ s_{dn}^{c};\ s_{dn}^{r};\text{vec}(A_{dn});\ \text{vec}(U_{dn});\ \text{vec}_{\neq}(T_{dn})\right], \tag{11}$$

where h_{dn}^{base} contains conventional temporal, spectral, and complexity descriptors; s_{dn}^{c} and s_{dn}^{r} contain fast- and slow-state summaries; and $\text{vec}_{\neq}$ retains admissible off-diagonal fast transitions.

To preserve label-space independence, HiMDA separates dynamics transfer from task-specific prediction. The representation ϕ_{dn} is passed to a newly initialized random-forest readout Breiman (2001), fitted exclusively on target-domain training pairs (ϕ_{dn}, y_{dn}). Thus, the source domain contributes only a finite prior over hierarchical dynamics; source class labels, logits, and prediction heads never define the target decision function.

4 Experiments

4.1 Experimental Settings

We evaluate HiMDA with one source dataset and two transfer datasets. The source dataset is the *Resting-State EEG Data before and after Cognitive Activity across the Adult Lifespan and a 5-Year Follow-Up* dataset Getzmann et al. (2024); Wascher et al. (2024). It contains 608 participants, 64 EEG channels, and signals sampled at 1000 Hz. Preprocessing yields 52,645 analysis windows, which are divided by a fixed stratified 8:2 split. The source data have no native cognitive-state annotations, so we construct four operational classes—high motivation, high alertness, fatigue, and attention lapse—from alpha/theta ratios, spectral entropy, spatial complexity, and temporal and GFP statistics. Their counts are 8,913, 15,608, 22,482, and 5,642, respectively. HiMDA's source classifier also receives spectral and signal-complexity features that overlap with several label descriptors. Its score therefore measures recovery of the constructed rule partition, not independent cognitive-state recognition. The source random forest obtains 76.94% accuracy, 75.37% balanced accuracy, 76.74% macro-F1, and Cohen's $\kappa = 0.6621$ on held-out source windows, compared with 42.70% accuracy for a majority-class predictor and 25% expected balanced accuracy for uniform guessing. Source labels are used only for this diagnostic; the source-training partition estimates reference prototypes, BSDS fingerprints, and transition priors without labels. The four operational classes are not aligned to either target label space.

SEED-IV. SEED-IV is the four-class affective transfer dataset Zheng et al. (2019). We use a subject-dependent, trial-level five-fold protocol: all windows from a trial remain in one fold, and window probabilities are averaged into one held-out trial prediction. Fold assignment precedes filtering, prototype estimation, BSDS fitting, feature construction, and head selection. Rows in the protocol-matched block use the same folds, target supervision, seed set, and trial-aggregation rule. Consequently, each reported prediction summarizes an unseen complete trial rather than neighboring windows that may share short-range temporal structure with the training data.

I-CARE. We evaluate on the publicly released I-CARE development cohort, comprising 607 post-cardiac-arrest patients from five represented hospitals Amorim et al. (2023a;b). We predict favorable (CPC1–2) versus poor (CPC3–5) outcome from the first 72 hours of EEG, using 2-second windows with 0.5-s overlap. In leave-one-hospital-out evaluation, the held-out hospital is evaluated once and all remaining hospitals form the development partition. Model selection and fitted components use development hospitals only. Window probabilities are averaged per patient before computing AUROC, AUPRC, and the hospital-constrained poor-outcome detection rate. The latter measures the proportion of poor-outcome patients detected while the false-positive rate is constrained to at most 5% separately at every hospital. The hospitals contribute A: 139/261 poor outcomes (53.3%, 95% Wilson CI 47.2–59.2), B: 86/120 (71.7%, 63.0–79.0), D: 56/83 (67.5%, 56.8–76.6), E: 59/74 (79.7%, 69.2–87.3), and F: 42/69 (60.9%, 49.1–71.5). These intervals describe outcome prevalence, not model performance.

Method	Accuracy	Std.
Protocol-matched baselines		
HiMDA (ours)	**77.04%**	**1.09%**
BiHDMLi et al. (2021)	74.35%	14.09%
TSceptionDing et al. (2023)	71.40%	1.13%
A-LSTM	69.50%	15.65%
GRU-MCCCui et al. (2023)	61.55%	3.38%
DANNGanin et al. (2016)	60.60%	3.06%
EEGNetLawhern et al. (2018)	58.52%	2.99%
SVM (DE)Duan et al. (2013)	56.61%	20.05%
XGBoostChen & Guestrin (2016)	50.47%	1.52%
Reference-only result		
CIT-EmotionNetLu et al. (2024)	92.09%	5.33%

Table 1: SEED-IV trial-level accuracy (%) for the shared subject-dependent five-fold comparison. All windows from each trial remain in the same fold, and their predicted probabilities are averaged into one trial-level prediction. The first block contains protocol-matched results. CIT-EmotionNet is a published reference evaluated with a randomly shuffled 60:40 split, not the shared protocol.

Method	AUROC	AUPRC	Constrained Detection Rate
Protocol-matched baselines			
HiMDA (ours)	**0.85**	**0.90**	**0.61**
CBraMod (scratch)Wang et al. (2025a)	0.79	0.87	0.63
EEGNetLawhern et al. (2018)	0.75	0.82	0.40
EEG2RepFoumani et al. (2024)	0.74	0.83	0.48
RBF-SVM	0.72	0.79	0.36
Reference-only result			
BiaxialFormerNesaragi et al. (2025)	0.77	0.86	0.52

Table 2: I-CARE patient-level transfer results under 72-hour leave-one-hospital-out evaluation; probabilities are aggregated per patient. The first block contains results run with the present protocol; CBraMod uses scratch initialization. The reference-only block reports BiaxialFormer with 5-minute segments and 30-second stride under its own leave-one-hospital-out setup. It is not a direct comparison.

4.2 System Comparison

Models using the present split and preprocessing are protocol-matched; published results with different splits or input resolutions are reference-only and not ranked against them.

Affective transfer on SEED-IV. HiMDA attains $77.04\% \pm 1.09\%$ trial-level accuracy under the shared five-fold protocol. As shown in Table 1, it achieves the highest accuracy among the retained protocol-matched comparisons. CIT-EmotionNet is reported separately using its published random 60:40 split.

Clinical transfer on I-CARE. Under the 72-hour hospital-held-out protocol, HiMDA reaches an AUROC of 0.85, an AUPRC of 0.90, and a constrained detection rate of 0.61, pooled across all 607 patients. Scratch-initialized CBraMod scores 0.79/0.87/0.63 on these measures: HiMDA improves AUROC and AUPRC but not constrained detection.

4.3 Ablation Study and Analysis

Ablation design. We conduct one-mechanism-at-a-time ablations on the source diagnostic and both targets, keeping windows, feature dimensionality, readout, splits, labels, and aggregation fixed. Removing prototype anchoring sets $\lambda_\mu = 0$ while retaining target clustering and post-hoc matching; removing time-local conditioning replaces $\rho_{qn,b_{qn}(t)}(m)$ with its within-window mean; and removing operator priors sets $\tau_A = \tau_T = 0$ with symmetric smoothing. Each variant is reconstructed from its first affected stage. The prototype ablation removes the finite anchoring penalty but retains source-prototype matching, so it is not source-free. The local-BSDS ablation retains window-averaged slow-regime information.

Target-only hierarchical control. The target-only control retains the slow–fast architecture and target readout, but estimates prototypes, regime fingerprints, and both transition operators from target-training data alone. It scores 64.56%/67.92% on the separate source diagnostic, 73.68% on SEED-IV, and 0.81/0.85 AUROC/AUPRC on I-CARE. Against full HiMDA, these results are lower by 10.81/8.82 source points, 3.36 SEED-IV points, and 0.04/0.05 on I-CARE. This complete source-removal control differs from w/o prototype anchor, which retains source-prototype initialization and post-hoc state matching.

Variant	Source		SEED-IV	I-CARE	
	Bal. Acc.	Macro-F1	Acc.	AUROC	AUPRC
HiMDA	**75.37%**	**76.74%**	**77.04%**	**0.85**	**0.90**
Target-only hierarchical HiMDA[§]	64.56%	67.92%	73.68%	0.81	0.85
w/o local BSDS	69.36%	72.00%	74.38%	0.83	0.87
w/o operator prior	68.86%	61.88%	75.69%	0.81	0.88
w/o prototype anchor	63.09%	50.44%	72.21%	0.82	0.86

Table 3: One-mechanism-at-a-time ablation with fixed splits, labels, features, and readout. SEED-IV is trial-level accuracy (%); I-CARE is patient-level. [§]Target-only hierarchical HiMDA fits target prototypes, regime fingerprints, and transition operators without source information. Its source columns come from the separate source diagnostic and do not indicate source access by the target-only transfer model.

Effect of source-anchored prototypes. The complete model obtains 75.37%/ 76.74% on the source diagnostic, 77.04% on SEED-IV, and 0.85/0.90 on I-CARE. Removing prototype anchoring gives the largest source decrease, 12.28/26.30 percentage points, the largest SEED-IV decrease of 4.83 points, and lowers I-CARE AUROC/AUPRC by 0.03/0.04. Without a finite reference, axial maps remain clusterable, but less consistent state identities propagate to microstate sequences and transition features.

Effect of source-centered operator priors. Removing the source operator prior reduces source balanced accuracy/macro-F1 by 6.51/14.86 points, SEED-IV accuracy by 1.35 points, and gives the largest I-CARE AUROC decrease, from 0.85 to 0.81; AUPRC falls from 0.90 to 0.88. The finite Dirichlet update stabilizes sparse rows while allowing well-supported target evidence to override the prior.

Effect of time-local BSDS conditioning. Replacing the time-local BSDS posterior by a within-window regime mean removes the regime-specific weighting of individual fast switches. Source balanced accuracy/macro-F1 fall by 6.01/4.74 points, SEED-IV accuracy falls by 2.66 points, and I-CARE AUROC/AUPRC decrease by 0.02/0.03. Marginal slow-state and microstate summaries remain informative, but local slow context distinguishes otherwise similar switches that occur in different regimes.

Overall interpretation. Every measured ablation lowers all five metrics. Prototype anchoring has the largest effects on the source diagnostic and SEED-IV, while operator priors produce the largest I-CARE AUROC decrease. Local BSDS conditioning improves both target results. The target-only control also retains strong performance on both tasks, and source-informed components raise each reported score.

5 Conclusion

HiMDA transfers hierarchical dynamics rather than labels or decision boundaries. Target-specific calibration establishes polarity-invariant microstates; soft BSDS posteriors connect slow regimes to time-local fast switches; and finite source-centered priors adapt both transition operators to target evidence. A new random forest defines target semantics without source–target class mapping. HiMDA reaches 77.04% trial accuracy on SEED-IV and 0.85/0.90 patient-level AUROC/AUPRC on I-CARE. The target-only control scores 73.68% and 0.81/0.85, respectively, while each component ablation reduces all five reported metrics. HiMDA also improves I-CARE discrimination over scratch CBraMod (0.85/0.90 versus 0.79/0.87); CBraMod has a slightly higher constrained detection rate (0.63 versus 0.61). The results across the two targets reflect the same division of labor: source dynamics provide the transferable temporal organization, while a target-specific readout assigns task semantics. This separation supports affective trial classification and clinical patient-level ranking without reusing source decision boundaries. On I-CARE, window probabilities are aggregated only after the held-out recordings are processed, so the reported AUROC and AUPRC measure patient-level discrimination under hospital-held-out transfer. Across source recovery, affective transfer, and clinical prediction, the same operator structure remains effective despite different target label geometries. Each task uses it through a calibrated interface and its own classifier.

AI USE STATEMENT

Large language models (LLMs) were used to assist with writing and polishing this manuscript. Their role was limited to editorial and presentation support, including sentence rephrasing, grammar checking, language refinement, readability, clarity, LaTeX template conversion, and formatting checks. LLMs were not involved in conceiving the research, selecting the methodology, designing or running experiments, analyzing data, interpreting results, or formulating scientific claims. All research ideas, methods, analyses, and conclusions were developed and verified by the authors. The authors take full responsibility for the manuscript and reviewed the AI-assisted text against the source materials, equations, citations, numerical results, and compilation output to avoid unsupported claims, plagiarism, or scientific misconduct.

REFERENCES

Edilberto Amorim, Wei-Long Zheng, Mohammad M. Ghassemi, Mahsa Aghaeeaval, Pravinkumar Kandhare, Vishnu Karukonda, Jong Woo Lee, Susan T. Herman, Adithya Sivaraju, Nicolas Gaspard, Jeannette Hofmeijer, Michel J. A. M. van Putten, Reza Sameni, Matthew A. Reyna, Gari D. Clifford, and M. Brandon Westover. The international cardiac arrest research consortium electroencephalography database. *Critical Care Medicine*, 51(12):1802–1811, 2023a. doi: 10.1097/CCM.0000000000006074. URL https://doi.org/10.1097/CCM.0000000000006074.

Edilberto Amorim, Wei-Long Zheng, Jong Woo Lee, Susan Herman, Mohammad Ghassemi, Adithya Sivaraju, Nicolas Gaspard, Jeannette Hofmeijer, Michel J. A. M. van Putten, Matthew Reyna, Gari Clifford, and Brandon Westover. I-CARE: International cardiac arrest REsearch consortium database. *PhysioNet*, February 2023b. doi: 10.13026/rjbz-cq89. URL https://doi.org/10.13026/rjbz-cq89. Version 1.0.

Fiorenzo Artoni, Julien Maillard, Juliane Britz, Denis Brunet, Christopher Lysakowski, Martin R. Tramèr, and Christoph M. Michel. Microsynt: Exploring the syntax of EEG microstates. *NeuroImage*, 277:120196, 2023. doi: 10.1016/j.neuroimage.2023.120196. URL https://doi.org/10.1016/j.neuroimage.2023.120196.

Leo Breiman. Random forests. *Machine Learning*, 45(1):5–32, 2001. doi: 10.1023/A:1010933404324. URL https://doi.org/10.1023/A:1010933404324.

Tianqi Chen and Carlos Guestrin. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 785–794, 2016. doi: 10.1145/2939672.2939785. URL https://doi.org/10.1145/2939672.2939785.

Nicolas Coquelet, Xavier De Tiège, Larysa Roshchupkina, Philippe Peigneux, Serge Goldman, Mark Woolrich, and Vincent Wens. Microstates and power envelope hidden markov modeling probe bursting brain activity at different timescales. *NeuroImage*, 247:118850, 2022. doi: 10.1016/j.neuroimage.2021.118850. URL https://doi.org/10.1016/j.neuroimage.2021.118850.

Heng Cui, Aiping Liu, Xu Zhang, Xiang Chen, Jun Liu, and Xun Chen. EEG-based subject-independent emotion recognition using gated recurrent unit and minimum class confusion. *IEEE Transactions on Affective Computing*, 14(4):2740–2750, 2023. doi: 10.1109/TAFFC.2022.3179717. URL https://doi.org/10.1109/TAFFC.2022.3179717.

Wenhui Cui, Woojae Jeong, Philipp Thölke, Takfarinas Medani, Karim Jerbi, Anand A. Joshi, and Richard M. Leahy. Neuro-GPT: Towards a foundation model for EEG. In *2024 IEEE International Symposium on Biomedical Imaging*, pp. 1–5, 2024. doi: 10.1109/ISBI56570.2024.10635453. URL https://doi.org/10.1109/ISBI56570.2024.10635453.

Yi Ding, Neethu Robinson, Su Zhang, Qiuhao Zeng, and Cuntai Guan. TSception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition. *IEEE Transactions on Affective Computing*, 14(3):2238–2250, 2023. doi: 10.1109/TAFFC.2022.3169001. URL https://doi.org/10.1109/TAFFC.2022.3169001.

Ruo-Nan Duan, Jia-Yi Zhu, and Bao-Liang Lu. Differential entropy feature for EEG-based emotion classification. In *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, pp. 81–84, 2013. doi: 10.1109/NER.2013.6695876. URL https://doi.org/10.1109/NER.2013.6695876.

Yuqi Fang, Pew-Thian Yap, Weili Lin, Hongtu Zhu, and Mingxia Liu. Source-free unsupervised domain adaptation: A survey. *Neural Networks*, 174:106230, 2024. doi: 10.1016/j.neunet.2024.106230.

Navid Mohammadi Foumani, Geoffrey Mackellar, Soheila Ghane, Saad Irtza, Nam Nguyen, and Mahsa Salehi. EEG2Rep: Enhancing self-supervised EEG representation through informative masked inputs. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 5544–5555, 2024. doi: 10.1145/3637528.3671600. URL https://doi.org/10.1145/3637528.3671600.

Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario March, and Victor Lempitsky. Domain-adversarial training of neural networks. *Journal of Machine Learning Research*, 17(59):1–35, 2016. URL https://jmlr.org/papers/v17/15-239.html.

Stephan Getzmann, Patrick D. Gajewski, Daniel Schneider, and Edmund Wascher. Resting-state EEG data before and after cognitive activity across the adult lifespan and a 5-year follow-up. *Scientific Data*, 11(1):988, 2024. doi: 10.1038/s41597-024-03797-w. URL https://doi.org/10.1038/s41597-024-03797-w.

Demetres Kostas, Stéphane Aroca-Ouellette, and Frank Rudzicz. BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data. *Frontiers in Human Neuroscience*, 15:653659, 2021. doi: 10.3389/fnhum.2021.653659. URL https://doi.org/10.3389/fnhum.2021.653659.

Harold W. Kuhn. The hungarian method for the assignment problem. *Naval Research Logistics Quarterly*, 2(1–2):83–97, 1955. doi: 10.1002/nav.3800020109. URL https://doi.org/10.1002/nav.3800020109.

Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering*, 15(5):056013, 2018. doi: 10.1088/1741-2552/aace8c. URL https://doi.org/10.1088/1741-2552/aace8c.

Jingcong Li, Weijian Pan, Haiyun Huang, Jiahui Pan, and Fei Wang. STGATE: Spatial-temporal graph attention network with a transformer encoder for EEG-based emotion recognition. *Frontiers in Human Neuroscience*, 17:1169949, 2023. doi: 10.3389/fnhum.2023.1169949. URL https://doi.org/10.3389/fnhum.2023.1169949.

Yang Li, Lei Wang, Wenming Zheng, Yuan Zong, Lei Qi, Zhen Cui, Tong Zhang, and Tengfei Song. A novel bi-hemispheric discrepancy model for EEG emotion recognition. *IEEE Transactions on Cognitive and Developmental Systems*, 13(2):354–367, 2021. doi: 10.1109/TCDS.2020.2999337. URL https://doi.org/10.1109/TCDS.2020.2999337.

Jianhua Lin. Divergence measures based on the shannon entropy. *IEEE Transactions on Information Theory*, 37(1):145–151, 1991. doi: 10.1109/18.61115. URL https://doi.org/10.1109/18.61115.

Scott Linderman, Matthew Johnson, Andrew Miller, Ryan Adams, David Blei, and Liam Paninski. Bayesian learning and inference in recurrent switching linear dynamical systems. In Aarti Singh and Jerry Zhu (eds.), *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics*, volume 54 of *Proceedings of Machine Learning Research*, pp. 914–922. PMLR, 2017. URL https://proceedings.mlr.press/v54/linderman17a.html.

Huan Liu, Yuzhe Zhang, Xuxu Chen, Dalin Zhang, Rui Li, and Tao Qin. Self-supervised EEG representation learning for robust emotion recognition. *ACM Transactions on Sensor Networks*, 20(5):1–22, 2024. doi: 10.1145/3674975. URL https://doi.org/10.1145/3674975.

Wei Lu, Lingnan Xia, Tien Ping Tan, and Hua Ma. Cit-emotionnet: convolution interactive transformer network for eeg emotion recognition. *PeerJ Computer Science*, 10:e2610, 2024.

Naimahmed Nesaragi, Hemin Ali Qadir, Per Steiner Halvorsen, and Ilangko Balasingham. Biaxialformer: Leveraging channel independence and inter-channel correlations in EEG signal decoding for predicting neurological outcomes, 2025. URL https://arxiv.org/abs/2507.02879. Preprint.

Roberto D. Pascual-Marqui, Christoph M. Michel, and Dietrich Lehmann. Segmentation of brain electrical activity into microstates: Model estimation and validation. *IEEE Transactions on Biomedical Engineering*, 42(7):658–665, 1995. doi: 10.1109/10.391164. URL https://doi.org/10.1109/10.391164.

Jalil Taghia, Weidong Cai, Srikanth Ryali, John Kochalka, Jonathan Nicholas, Tianwen Chen, and Vinod Menon. Uncovering hidden brain state dynamics that regulate performance and decision-making during cognition. *Nature Communications*, 9(1):2505, 2018. doi: 10.1038/s41467-018-04723-6. URL https://doi.org/10.1038/s41467-018-04723-6.

Povilas Tarailis, Thomas Koenig, Christoph M. Michel, and Inga Griškova-Bulanova. The functional aspects of resting EEG microstates: A systematic review. *Brain Topography*, 37(2):181–217, 2024. doi: 10.1007/s10548-023-00958-9. URL https://doi.org/10.1007/s10548-023-00958-9.

Guagnyu Wang, Wenchao Liu, Yuhong He, Cong Xu, Lin Ma, and Haifeng Li. EEGPT: Pretrained transformer for universal and reliable representation of EEG signals. In *Advances in Neural Information Processing Systems*, volume 37, pp. 39249–39280. Curran Associates, Inc., 2024. doi: 10.52202/079017-1239. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/4540d267eeec4e5dbd9dae9448f0b739-Abstract-Conference.html.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. CBraMod: A criss-cross brain foundation model for EEG decoding. In *International Conference on Learning Representations*, 2025a. URL https://openreview.net/forum?id=NPNUHgHF2w.

Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Shijian Li, and Gang Pan. EEGMamba: An EEG foundation model with Mamba. *Neural Networks*, pp. 107816, 2025b. doi: 10.1016/j.neunet.2025.107816. URL https://doi.org/10.1016/j.neunet.2025.107816.

Edmund Wascher, Daniel Schneider, Patrick D. Gajewski, and Stephan Getzmann. Resting-state EEG data before and after cognitive activity across the adult lifespan and a 5-year follow-up. OpenNeuro, 2024. URL https://doi.org/10.18112/openneuro.ds005385.v1.0.0. Dataset ds005385, version 1.0.0.

Chaoqi Yang, M. Brandon Westover, and Jimeng Sun. BIOT: Biosignal transformer for cross-data learning in the wild. In *Advances in Neural Information Processing Systems*, volume 36, pp. 78240–78260. Curran Associates, Inc., 2023. URL https://proceedings.neurips.cc/paper_files/paper/2023/hash/f6b30f3e2dd9cb53bbf2024402d02295-Abstract-Conference.html.

Kaichao You, Mingsheng Long, Zhangjie Cao, Jianmin Wang, and Michael I. Jordan. Universal domain adaptation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2720–2729, 2019. doi: 10.1109/CVPR.2019.00283.

Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki. EmotionMeter: A multimodal framework for recognizing human emotions. *IEEE Transactions on Cybernetics*, 49(3):1110–1122, 2019. doi: 10.1109/TCYB.2018.2797176. URL https://doi.org/10.1109/TCYB.2018.2797176.