



This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

DynFS-MoE: Dynamic Functional-Structural Mixture-of-Experts for Post-Traumatic Epilepsy Diagnosis

Jun-En Ding^{1,2}, Spencer Chen³, Henry Noren³, Daniel Valdivia³, Christine Yohn³, Suhina Patel³, Taylor Zink³, Hai Sun³, and Feng Liu^{1,2}(✉)

¹ Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ, USA

² Department of Systems Engineering, Stevens Institute of Technology, Hoboken, NJ, USA

³ Department of Neurosurgery, Robert Wood Johnson Medical School, Rutgers University, New Brunswick, NJ, USA
{junen.ding,feng.liu}@rutgers.edu

Abstract. Post-traumatic epilepsy (PTE) is a severe complication of traumatic brain injury (TBI). Yet, early identification remains challenging due to the complex structural and functional alterations it induces in the brain. To address this, we propose a dynamic multimodal Mixture-of-Experts (MoE) framework that integrates functional and structural connectivity through time-aware functional-structural encoding and class-conditioned expert routing. Within this framework, modality-specific and cross-modal experts learn complementary representations, while a Modality-Class MoE (MCoE) module dynamically adjusts expert weights according to each classification objective. Experimental results across three binary classification tasks demonstrate that the framework consistently outperforms static fusion baselines, and interpretability analyses further reveal meaningful interactions among regions of interest (ROIs). This dynamic multimodal expert framework effectively captures class-dependent brain interaction patterns and provides an interpretable approach for PTE diagnosis and risk stratification.

Keywords: Post-traumatic epilepsy · Traumatic brain injury · Structural connectivity · Functional connectivity · Mixture-of-Experts · Multimodal fusion

1 Introduction

Traumatic brain injury (TBI) remains a leading cause of morbidity and mortality worldwide, affecting millions annually and imposing substantial socioeconomic burdens [1, 2]. Among its long-term complications, TBI carries a significant latent risk of progression to post-traumatic epilepsy (PTE), which is associated with high recurrence rates and, in severe cases, may lead to mortality. PTE is characterized by recurrent unprovoked seizures occurring more than seven days after injury [3–5].

Currently, structural and functional connectivity offer high spatial resolution and network-level insights that effectively complement electroencephalogram (EEG) limitations. Structural MRI can precisely detect microstructural damage, brain atrophy, and hippocampal or thalamic volume loss [6–8]. Moreover, resting-state functional magnetic resonance imaging (fMRI) further reveals functional connectivity abnormalities associated with the transition from TBI to PTE, such as thalamocortical hyperconnectivity, excessive integration, and reduced segregation patterns, thereby capturing the network mechanisms underlying epileptogenesis [9, 10].

However, most existing studies employ only structural MRI [7, 8] or fMRI individually [11], focusing on structural damage (such as lesion volume and hippocampal or thalamic atrophy) or functional connectivity abnormalities (such as hyperconnectivity patterns) [9, 12]. Although these studies reveal significant differences in specific brain regions (e.g., temporal lobe and cerebellum), unimodal models generally yield lower predictive accuracy compared to multimodal integration and cannot simultaneously capture temporal dynamics and spatial network information, resulting in insufficient explanatory power for the heterogeneous progression of PTE [13–16].

Emerging evidence suggests that multimodal integration can substantially improve PTE prediction accuracy. For instance, machine learning models combining structural MRI, functional MRI, and EEG have achieved AUC values above 0.78 [11]. However, multimodal fusion faces several fundamental challenges. First, most fusion strategies project all modalities into a single shared feature space, potentially leading to information loss. Second, modality competition may arise during fusion, where dominant modalities overshadow weaker ones, and inter-individual variability in modality informativeness is often neglected. Finally, different TBI or PTE phenotypes may benefit from distinct primary modalities, necessitating adaptive rather than uniform weighting of modality contributions. To address the above limitations, we propose three modules to handle the challenges of dynamic fMRI and diffusion MRI (dMRI)-derived structural connectivity integration. Our contributions are as follows: (1) We design a temporal MoE mechanism that leverages structural connectivity to guide the selection of the most informative patches within fMRI temporal segments for cross-modal fusion and routing. (2) Modality-specific experts (MSoEs) are employed to extract individual feature representations from each modality, producing dedicated embeddings that preserve modality-unique information. (3) A class-based MoE module adaptively selects the most class-relevant modality patches to facilitate more discriminative classification.

2 Methodology

As illustrated in Fig. 1, **DynFS-MoE** consists of three components: (1) a time-aware functional-structural encoder that extracts temporal fMRI and structural connectivity representations into refined modality patches through segment-wise fusion; (2) a mixture-of-experts layer composed of modality-specific experts and

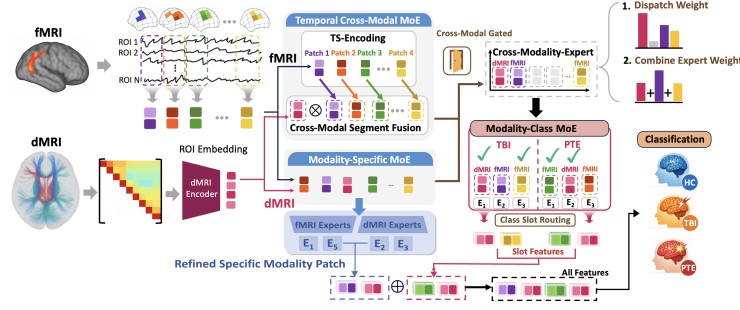


Fig. 1. Overview of the proposed DynFS-MoE framework.

cross-modal experts, which capture both modality-specific characteristics and cross-modal interactions; and (3) a modality-class MoE (MCoE) routing module that dynamically assigns expert weights using a gating mechanism conditioned on class-aware representations. The routed expert outputs are combined to select the most suitable multimodal features for classification.

2.1 Expert Space and Routing Formulation

We define the expert set as $\mathcal{E} = \{e_1, e_2, \dots, e_J\}$, where J denotes the total number of experts and each expert is parameterized as a learnable transformation $\phi_j : \mathbb{R}^D \rightarrow \mathbb{R}^D$. The expert set comprises modality-specific experts, which specialize in processing representations from either fMRI or dMRI, and cross-modal experts, which capture interactions between functional and structural modalities. To formalize the routing process, we introduce three discrete random variables: the modality patch variable $P \in \mathcal{P}$, the expert variable $E \in \mathcal{E}$, and the diagnostic label $Y \in \{1, \dots, C\}$, where C denotes the number of diagnostic categories. The routing mechanism is modeled as a class-conditioned joint distribution $P_\theta(P, E | Y)$, parameterized by θ , which defines the probability of assigning modality patches to experts conditioned on the class-specific representation. This probabilistic formulation enables class-aware expert selection and adaptive integration of modality-specific and cross-modal information, allowing the model to dynamically allocate expert capacity according to diagnostic relevance.

2.2 Time-Aware Functional-Structural MoE Encoder

Throughout this paper, we denote fMRI BOLD signals $\mathbf{F} \in \mathbb{R}^{R \times T}$ and the structural connectivity matrix, denoted as $\mathbf{S} \in \mathbb{R}^{R \times R}$, where R is the number of regions of interest (ROIs) and T is the number of time steps. To jointly model the temporal dynamics of fMRI and the structural information from dMRI, we propose a time-aware functional-structural encoder that converts raw multimodal inputs into refined modality-specific patches. Our goal is to learn a

classifier $f_\theta : (\mathbf{F}, \mathbf{S}) \rightarrow \mathcal{Y}$, parameterized by θ , which predicts the diagnostic label $y \in \mathcal{Y} = \{1, \dots, C\}$.

Modality Patch-Wise Representation. To explicitly model temporal dynamics, the fMRI sequence is partitioned into $K = \lfloor \frac{T}{L} \rfloor$ non-overlapping temporal segments of fixed length L as patches, where each modality patch represents an embedding of modality-specific temporal (fMRI) or structural connectivity information. Specifically, the fMRI temporal patches $\mathbf{F} = \{F^{(1)}, F^{(2)}, \dots, F^{(K)}\}$ with each $F^{(k)} \in \mathbb{R}^{R \times L}$. For the structural modality, we obtain a compact representation by extracting the upper-triangular entries of the connectivity matrix, denoted as $\text{vec}_\Delta(S) \in \mathbb{R}^{\frac{R(R+1)}{2}}$, where $\text{vec}_\Delta(\cdot)$ collects elements $S_{i,j}$ satisfying $1 \leq i \leq j \leq R$. We then encode temporal patches and structural patches to MLP layer $\phi(\cdot)$ as same dimensional embedding $f^{(k)} \in \mathbb{R}^D$ and $s \in \mathbb{R}^D$, respectively. The projected unified modality patch space is $\mathcal{P} = \{f^{(1)}, \dots, f^{(K)}, s\}$, with cardinality $|\mathcal{P}| = K + 1$. This unified patch formulation enables both temporal functional patches and structural connectivity features to be processed within the same MoE routing framework described in Section 2.3 and 2.4, facilitating structured expert assignment and cross-modal integration.

Cross-Modal Interaction Fusion. Unlike previous approaches that adopt a traditional unified modality paradigm, we propose a lightweight cross-modal fusion to integrate structural connectivity features into each fMRI temporal patch prior to MoE routing. For each temporal patch, a scaled dot-product gating is performed as:

$$\tilde{f}^{(k)} = f^{(k)} + \alpha^{(k)} \mathbf{W}_V s, \quad (1)$$

where the gating coefficient $\alpha^{(k)} = \sigma\left(\frac{(\mathbf{W}_Q f^{(k)})^\top (\mathbf{W}_K s)}{\sqrt{D}}\right) \in (0, 1)$ is a scalar sigmoid gate that adaptively modulates the contribution of structural connectivity to each functional patch. Next, a feed-forward projection is applied to produce the cross-modal fused embedding $z_{fused}^{(k)}$ with MLP $\phi(\tilde{f}^{(k)})$.

2.3 Modality-Specific MoEs (MSoEs)

To capture modality-specific characteristics, we employ modality-specific MoEs operating on modality patch embeddings $p_i \in \mathcal{P}$, $p_i \in \mathbb{R}^D$. Given the expert set $\mathcal{E}$, and a gating function computes routing probabilities

$$\pi_j(p_i) = P_\theta(E = e_j \mid P = p_i) = \text{softmax}_j(\mathbf{W}_g p_i). \quad (2)$$

where $\pi_j(p_i)$ represents the routing probability of assigning patch p_i to expert e_j , and $\sum_{j=1}^J \pi_j(p_i) = 1$. We can calculate the modality-specific representation with a weighted combination of expert outputs $\mathbf{z}_i^{\text{spec}} = \sum_{j=1}^J \pi_j(p_i) \phi_j(p_i)$, which allows different experts to capture heterogeneous modality-specific patterns in functional dynamics and structural connectivity, thereby improving representation expressiveness prior to class-conditioned routing in the next Section 2.4

2.4 Modality-Class MoE (MCoE) with Class-Conditioned Routing

Inspired by existing work [17], we introduce the Modality-Class Mixture of Experts (MCoE) module, which enables different classes to learn the most suitable modality representations according to their assigned modality combinations and expert specialization. Specifically, the joint distribution $P_\theta(P, E|Y)$ allows different diagnostic classes, e.g., healthy control (HC), TBI, and PTE, to dynamically route modality patches to specialized experts, capturing heterogeneous modality dependencies.

Formally, the input to this module is the set of refined modality patch embeddings $\mathbf{Z} = \{\mathbf{z}_i\}_{i=1}^{|P|}$, where each $\mathbf{z}_i \in \mathbb{R}^D$ incorporates cross-modal fusion and modality-specific processing from prior layers (Sections 2.2 and 2.3). We compute class-conditioned routing logits as:

$$\mathbf{L} = f(\mathbf{Z}) \odot \Phi \in \mathbb{R}^{B \times |P| \times J \times C}. \quad (3)$$

where B denotes the batch size, and $\Phi \in \mathbb{R}^{|P| \times J \times C}$ is the parameter matrix. Next, we design the dispatch weights $D_{ijy} = \frac{\exp(L_{ijy})}{\sum_{i'=1}^{|P|} \exp(L_{i'jy})}$, which aggregate patches for each expert per class, representing $P(P_i | E_j, Y = y)$. For each class y and expert j , the expert input is a weighted sum augmented by class embedding $\mathbf{c}_y \in \mathbb{R}^D$ with the class-conditioned expert input $\mathbf{u}_{jy}$ and expert output $\mathbf{o}_{jy}$ can be expressed as:

$$\mathbf{u}_{jy} = \left(\sum_{i=1}^{|P|} D_{ijy} \cdot \mathbf{z}_i \right) + \mathbf{c}_y, \quad \mathbf{o}_{jy} = \phi_j(\mathbf{u}_{jy}), \quad (4)$$

To combine expert outputs into class-specific features, we compute combine weights $C_{jy} = \frac{\exp(\bar{L}_{jy})}{\sum_{j'=1}^J \exp(\bar{L}_{j'y})}$ via softmax over experts, using logits from mean-pooled $\mathbf{Z}$, and we can yield the fused feature $\mathbf{h}$ for class y as $\mathbf{h}_y = \sum_{j=1}^J C_{jy} \cdot \mathbf{o}_{jy}$, where class-specific features $\{\mathbf{h}_y\}_{y=1}^C$ are concatenated with modality-specific representations and passed to a final MLP classifier for predicting y .

2.5 Brain Interpretability Regularizer

Our training aims to promote diagnostic relevance using class-conditioned routing, allowing expert capacity to be allocated adaptively. To encourage class-distinctive patterns, we add a brain interpretability regularizer maximizing conditional mutual information $I(P; E | Y) = H(P | Y) - H(P | E, Y)$. Using logits $\mathbf{L}$, joint and marginal probabilities are $P(P_i, E_j | Y) = \frac{\exp(L_{ijy})}{\sum_{i'=1}^{|P|} \sum_{j'=1}^J \exp(L_{i'j'y})}$, and class-conditioned mutual information (CMI) can be written as:

$$\begin{aligned}
& I_\theta(P; E | Y) \\
&= \sum_{y=1}^C P(Y = y) \sum_{i=1}^{|P|} \sum_{j=1}^J P_\theta(P_i, E_j | Y = y) \cdot \log \left(\frac{P_\theta(P_i, E_j | Y = y)}{P_\theta(P_i | Y = y) \cdot P_\theta(E_j | Y = y)} \right)
\end{aligned} \tag{5}$$

Formally, the overall objective is to minimize the following loss over all model parameters θ :

$$\min_{\theta} \mathbb{E}_{(\mathbf{F}, \mathbf{S}, y) \sim \mathcal{D}} \left[\mathcal{L}_{\text{cls}}(f_\theta(\mathbf{F}, \mathbf{S}), y) - \lambda I_\theta(P; E | Y) \right] \tag{6}$$

where $\mathcal{L}_{\text{cls}}$ is the cross-entropy loss, and $I_\theta(P; E | Y)$ is the CMI encouraging class-specific modality routing, and λ is the balancing hyperparameter.

3 Experimental Setting and Results

In this study, we acquired MRI data from participants, including HC (n=26), TBI (n=24), and PTE (n=19) using a 3T Siemens Prisma scanner, including diffusion-weighted imaging. Preprocessing was performed using SPM, MRtrix3, and FSL. T1 images were segmented and normalized to MNI space, and functional images were motion-corrected, co-registered, and normalized. Particularly, structural connectivity matrices were constructed from diffusion-weighted imaging via DTI tractography using the AAL atlas (116 regions) [18].

3.1 Classification Performance Results

In this section, we evaluated **DynFS-MoE** on three binary classification tasks. As shown in Table 1, **DynFS-MoE** achieves the best AUC across all three tasks and competitive ACC/F1, with the strongest improvements observed for HC vs. PTE and TBI vs. PTE. For HC vs. TBI, **DynFS-MoE** achieves an AUC of 0.77, surpassing the strongest baseline EarlyFusion (AUC=0.74) by a substantial margin. The performance gains are most pronounced in the clinically challenging HC vs. PTE and TBI vs. PTE tasks, where **DynFS-MoE** attains AUCs of 0.84 and 0.82, respectively. Notably, several graph-based methods (Fusion+GCN, FBNetGen) yield near-zero F1 scores on PTE-related tasks, suggesting that static fusion strategies fail to capture the heterogeneous multimodal patterns underlying epileptogenesis. For HC vs. PTE, **DynFS-MoE** further achieves statistically significant improvements in both AUC and F1 ($p < 0.001$), with F1=0.81 outperforming the best baseline MultiViT (F1=0.74). For TBI vs. PTE, **DynFS-MoE** also yields significant gains ($p < 0.001$) with low variance (ACC=0.77±0.01, AUC=0.82±0.07), demonstrating stable and reliable performance under dynamic multimodal routing.

Table 1. Performance comparison across different classification tasks. Statistical significance was evaluated between DynFS-MoE and the best remaining baseline using Welch’s t-test (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$).

Cohort	Model	ACC	AUC	F1
HC vs. TBI	fMRI-FC+MLP	0.76±0.19	0.72±0.28	0.62±0.35
	dMRI+MLP	0.76±0.10	0.70±0.12	0.73±0.12
	Fusion+GCN [19]	0.52±0.07	0.56±0.15	0.40±0.23
	FBNetGen [20]	0.54±0.08	0.56±0.05	0.25±0.30
	EarlyFusion	0.76±0.10	0.74±0.12	0.70±0.12
	IBGNN [21]	0.68±0.12	0.67±0.18	0.65±0.16
	IBrainGNN [22]	0.60±0.09	0.46±0.23	0.41±0.25
	AGIIBM [23]	0.70±0.11	0.55±0.14	0.64±0.19
	MultiViT [24]	0.74±0.10	0.71±0.14	0.76±0.07
	Contrasformer [25]	0.72±0.07	0.60±0.09	0.73±0.06
	DynFS-MoE	0.76±0.05	0.77±0.09	0.72±0.16
HC vs. PTE	fMRI-FC+MLP	0.78±0.10	0.73±0.13	0.73±0.11
	dMRI+MLP	0.76±0.08	0.78±0.06	0.63±0.18
	Fusion+GCN [19]	0.58±0.04	0.42±0.19	0.00±0.00
	FBNetGen [20]	0.62±0.09	0.44±0.08	0.25±0.31
	EarlyFusion	0.71±0.09	0.65±0.18	0.51±0.16
	IBGNN [21]	0.60±0.05	0.70±0.27	0.18±0.22
	IBrainGNN [22]	0.62±0.09	0.39±0.26	0.16±0.32
	AGIIBM [23]	0.71±0.13	0.61±0.31	0.40±0.35
	MultiViT [24]	0.78±0.07	0.70±0.10	0.74±0.06
	Contrasformer [25]	0.80±0.11	0.68±0.18	0.72±0.20
	DynFS-MoE	0.84±0.05*	0.84±0.10***	0.81±0.08***
TBI vs. PTE	fMRI-FC+MLP	0.72±0.06	0.67±0.13	0.69±0.10
	dMRI+MLP	0.68±0.08	0.55±0.08	0.58±0.12
	Fusion+GCN [19]	0.56±0.04	0.51±0.26	0.00±0.00
	FBNetGen [20]	0.58±0.03	0.41±0.21	0.08±0.16
	EarlyFusion	0.67±0.15	0.65±0.12	0.60±0.19
	IBGNN [21]	0.61±0.04	0.57±0.10	0.42±0.35
	IBrainGNN [22]	0.56±0.04	0.55±0.24	0.00±0.00
	AGIIBM [23]	0.63±0.11	0.52±0.14	0.25±0.32
	MultiViT [24]	0.70±0.11	0.59±0.13	0.50±0.20
	Contrasformer [25]	0.59±0.10	0.31±0.18	0.37±0.21
	DynFS-MoE	0.77±0.01***	0.82±0.07***	0.70±0.04

3.2 Ablation Studies

Class-Conditioned Routing Mechanism. We compare class-conditioned routing against a uniform prior baseline in an ablation study. As shown in Table 2, class-conditioned routing consistently improves AUC across all tasks, with pronounced gains in TBI vs. PTE (0.79 vs. 0.58) and HC vs. PTE (0.81 vs. 0.72), attributed to its ability to adaptively allocate expert capacity per diagnostic class, where PTE down-weights dMRI in favor of early functional patches while TBI up-weights mid-temporal fMRI and structural features. Although F1 variability may reflect class imbalance, the consistent AUC gains confirm that class-conditioned routing is essential for learning discriminative, disease-specific multimodal representations.

Table 2. Ablation study on conditioning strategy

Cohort	Setting	AUC	F1
HC vs TBI	Uniform Prior	0.78 ± 0.05	0.73 ± 0.09
	Class-Conditioned (proposed)	0.79 ± 0.13	0.67 ± 0.20
HC vs PTE	Uniform Prior	0.72 ± 0.16	0.72 ± 0.13
	Class-Conditioned (proposed)	0.81 ± 0.04	0.58 ± 0.15
TBI vs PTE	Uniform Prior	0.58 ± 0.13	0.54 ± 0.29
	Class-Conditioned (proposed)	0.79 ± 0.12	0.71 ± 0.06

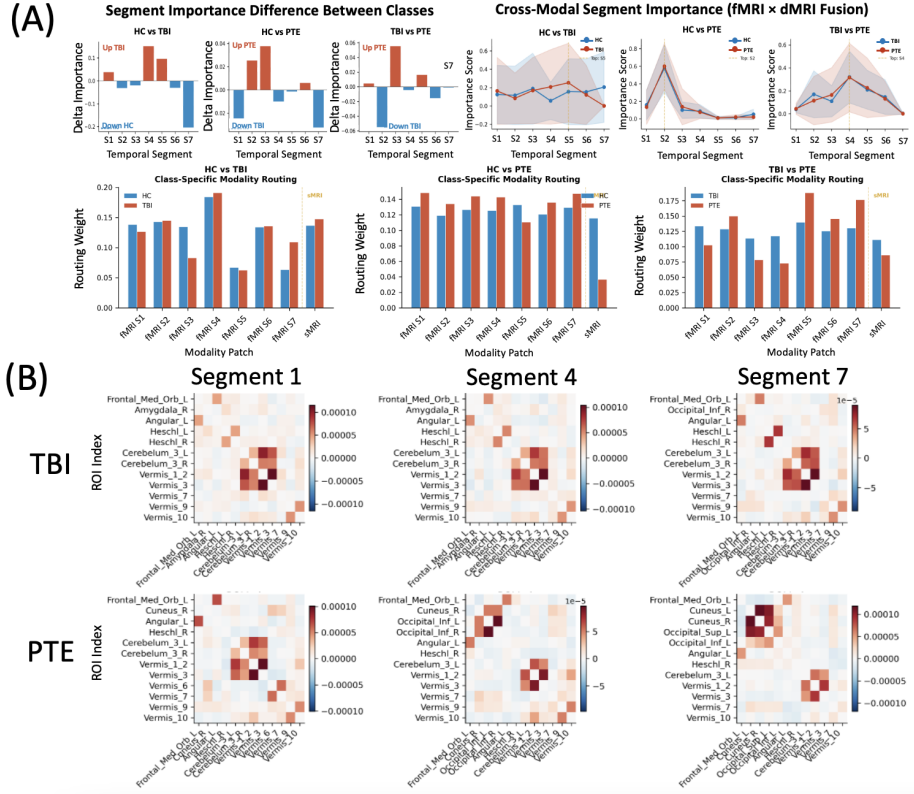


Fig. 2. Class-conditioned multimodal importance and ROI interaction visualization.

4 Functional Brain Interpretability and Structural Analysis

Fig. 2 (A) and Fig. 2 (B) illustrate the learned cross-modal temporal importance in brain regions and structural connectivity patterns. Fig. 2 (A) depicts three complementary interpretability analyses across the three binary classification tasks. PTE consistently exhibits elevated segment importance at early-to-mid temporal segments (notably S2–S4), with cross-modal importance curves peaking sharply at S2 in HC vs. PTE and S4 in TBI vs. PTE, suggesting that early functional-structural interactions are particularly informative for PTE identification. The class-specific modality routing further reveals that TBI up-weights mid-temporal fMRI patches and dMRI, HC distributes weights more uniformly, and PTE markedly down-weights dMRI in favor of functional patches, demonstrating adaptive modality reliance across diagnostic classes. Fig. 2 (B) displays pairwise ROI interaction heatmaps for TBI and PTE across Segments 1, 4, and 7. While both classes share a prominent cerebellar-vermis interaction core, PTE uniquely exhibits progressive recruitment of frontal-occipital (Frontal_Med_Orb_L, Cuneus, Occipital_Sup/Inf) and auditory cortical re-

gions (Heschl_R) in later segments, reflecting evolving functional-structural disruptions involving frontal-occipital, auditory, and cerebellar-vermis regions

5 Conclusion

In this study, we proposed DynFS-MoE, a dynamic multimodal mixture-of-experts framework for post-traumatic epilepsy (PTE) diagnosis after traumatic brain injury (TBI). By combining time-aware functional-structural encoding with class-conditioned expert routing, the model adaptively integrates fMRI and dMRI information and captures class-specific brain patterns. This dynamic routing strategy overcomes limitations of static fusion methods and improves discrimination performance across multiple classification tasks. Moreover, interpretability analyses reveal meaningful ROI interaction patterns, such as thalamic hyperconnectivity and limbic disruptions, providing clinically relevant insights. Overall, DynFS-MoE may provide an interpretable tool for PTE classification and could support future studies on early detection and risk stratification.

Funding Statement.

This work was supported by the U.S. Dept of Defense and Congressionally Directed Medical Research Programs (CDMRP) under Award No. W81XWH-18-1-0655 (PI: Hai Sun).

Disclosure of Interest.

All authors declare no potential conflict of interests.

References

1. Michael C Dewan, Abbas Rattani, Saksham Gupta, Ronnie E Baticulon, Ya-Ching Hung, Maria Punchak, Amit Agrawal, Amos O Adeleye, Mark G Shrimel, Andrés M Rubiano, et al. Estimating the global incidence of traumatic brain injury. *Journal of neurosurgery*, 130(4):1080–1097, 2018.
2. GBDTB Injury et al. Global, regional, and national burden of traumatic brain injury and spinal cord injury, 1990-2016: a systematic analysis for the global burden of disease study 2016. *Lancet Neurol*, 18(1):56–87, 2019.
3. Oliver Acosta and Jaclyn Barcikowski. An update on post-traumatic epilepsy. *Current Physical Medicine and Rehabilitation Reports*, 13(1):2, 2024.
4. Xiaoyu Yang, Siyang Chen, Haozhou Wang, Tong Sun, and Ke Wu. Post-traumatic epilepsy: bridging pathogenesis, diagnosis, and pharmacotherapeutic strategies. *Frontiers in Pharmacology*, 16:1697391, 2025.
5. Shihao Yang, Xiyang Huang, Danilo Bernardo, Jun-En Ding, Andrew Michael, Guoan Wang, Jingmei Yang, Alison Anderson, Dinesh Giritharan, Patrick Kwan, et al. Large-scale ai and foundation models for neuroscience: A comprehensive review. *Meta-Radiology*, page 100224, 2026.

6. Abdalla Z Mohamed, Paul Cumming, and Fatima A Nasrallah. Traumatic brain injury augurs ill for prolonged deficits in the brain’s structural and functional integrity following controlled cortical impact injury. *Scientific reports*, 11(1):21559, 2021.
7. H Akrami, RM Leahy, A Irimia, PE Kim, CN Heck, and AA Joshi. Neuroanatomic markers of posttraumatic epilepsy based on mr imaging and machine learning. *American Journal of Neuroradiology*, 43(3):347–353, 2022.
8. Riikka Immonen, Neil G Harris, David Wright, Leigh Johnston, Eppu Manninen, Gregory Smith, Afshin Paydar, Craig Branch, and Olli Grohn. Imaging biomarkers of epileptogenicity after traumatic brain injury—preclinical frontiers. *Neurobiology of disease*, 123:75–85, 2019.
9. Marianna La Rocca, Giuseppe Barisano, Rachael Garner, Sebastian F Ruf, Nicola Amoroso, Martin Monti, Paul Vespa, Roberto Bellotti, Deniz Erdoğan, Arthur W Toga, et al. Functional connectivity alterations in traumatic brain injury patients with late seizures. *Neurobiology of Disease*, 179:106053, 2023.
10. Rachael Garner, Marianna La Rocca, Paul Vespa, Nigel Jones, Martin M Monti, Arthur W Toga, and Dominique Duncan. Imaging biomarkers of posttraumatic epileptogenesis. *Epilepsia*, 60(11):2151–2162, 2019.
11. Haleh Akrami, Wenhui Cui, Paul E Kim, Christianne N Heck, Andrei Irimia, Karim Jerbi, Dileep Nair, Richard M Leahy, and Anand A Joshi. Prediction of post traumatic epilepsy using mr-based imaging markers. *Human brain mapping*, 45(17):e70075, 2024.
12. Rachael Garner, Marianna La Rocca, Giuseppe Barisano, Arthur W Toga, Dominique Duncan, and Paul Vespa. A machine learning model to predict seizure susceptibility from resting-state fmri connectivity. In *2019 Spring Simulation Conference (SpringSim)*, pages 1–11. IEEE, 2019.
13. Mingzhu Fang, Wanyu Liu, Jinmei Tuo, Mei Liu, Fangjing Li, Lijia Zhang, Changyin Yu, and Zucui Xu. Advances in understanding the pathogenesis of post-traumatic epilepsy: a literature review. *Frontiers in Neurology*, 14:1141434, 2023.
14. Silvia F Storti, Emanuela Formaggio, Enrica Franchini, Luigi G Bongiovanni, Roberto Cerini, Antonio Fiaschi, Christoph M Michel, and Paolo Manganotti. A multimodal imaging approach to the evaluation of post-traumatic epilepsy. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 25(5):345–360, 2012.
15. Elias Ebrahimzadeh, Saber Saharkhiz, Lila Rajabion, Homayoun Baghaei Oskouei, Masoud Seraji, Farahnaz Fayaz, Sarah Saliminia, Seyyed Mostafa Sadjadi, and Hamid Soltanian-Zadeh. Simultaneous electroencephalography-functional magnetic resonance imaging for assessment of human brain function. *Frontiers in Systems Neuroscience*, 16:934266, 2022.
16. Md Navid Akbar, Sebastian F Ruf, Ashutosh Singh, Razieh Faghihpourayesh, Rachael Garner, Alexis Bennett, Celina Alba, Marianna La Rocca, Tales Imbiriba, Deniz Erdoğan, et al. Advancing post-traumatic seizure classification and biomarker identification: Information decomposition based multimodal fusion and explainable machine learning with missing neuroimaging data. *Computerized Medical Imaging and Graphics*, 115:102386, 2024.
17. Chenwei Wu, Zitao Shuai, Zhengxu Tang, Luning Wang, and Liye Shen. Dynamic modeling of patients, modalities and tasks via multi-modal multi-task mixture of experts. In *ICLR*, 2025.
18. Nathalie Tzourio-Mazoyer, Brigitte Landeau, Dimitri Papathanassiou, Fabrice Crivello, Octave Etard, Nicolas Delcroix, Bernard Mazoyer, and Marc Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the mni mri single-subject brain. *Neuroimage*, 15(1):273–289, 2002.

19. Thomas N Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907*, 2016.
20. Xuan Kan, Hejie Cui, Joshua Lukemire, Ying Guo, and Carl Yang. Fbnetgen: Task-aware gnn-based fmri analysis via functional brain network generation. In *International conference on medical imaging with deep learning*, pages 618–637. PMLR, 2022.
21. Hejie Cui, Wei Dai, Yanqiao Zhu, Xiaoxiao Li, Lifang He, and Carl Yang. Interpretable graph neural networks for connectome-based brain disorder analysis. In *MICCAI*, pages 375–385. Springer, 2022.
22. Gang Qu, Ziyu Zhou, Vince D. Calhoun, Aiying Zhang, and Yu-Ping Wang. Integrated brain connectivity analysis with fmri, dti, and smri powered by interpretable graph neural networks. *Medical Image Analysis*, 103:103570, 2025.
23. Changxu Dong, Dengdi Sun, Zhenda Yu, and Bin Luo. Multi-view brain network classification based on adaptive graph isomorphic information bottleneck mamba. *Expert Systems with Applications*, 267:126170, 2025.
24. Yuda Bi, Anees Abrol, Zening Fu, and Vince Calhoun. Multivit: Multimodal vision transformer for schizophrenia prediction using structural mri and functional network connectivity data. In *ISBI*, pages 1–5. IEEE, 2023.
25. Jiaxing Xu, Kai He, Mengcheng Lan, Qingtian Bian, Wei Li, Tieying Li, Yiping Ke, and Miao Qiao. Contrasformer: a brain network contrastive transformer for neurodegenerative condition identification. In *CIKM*, pages 2671–2681, 2024.