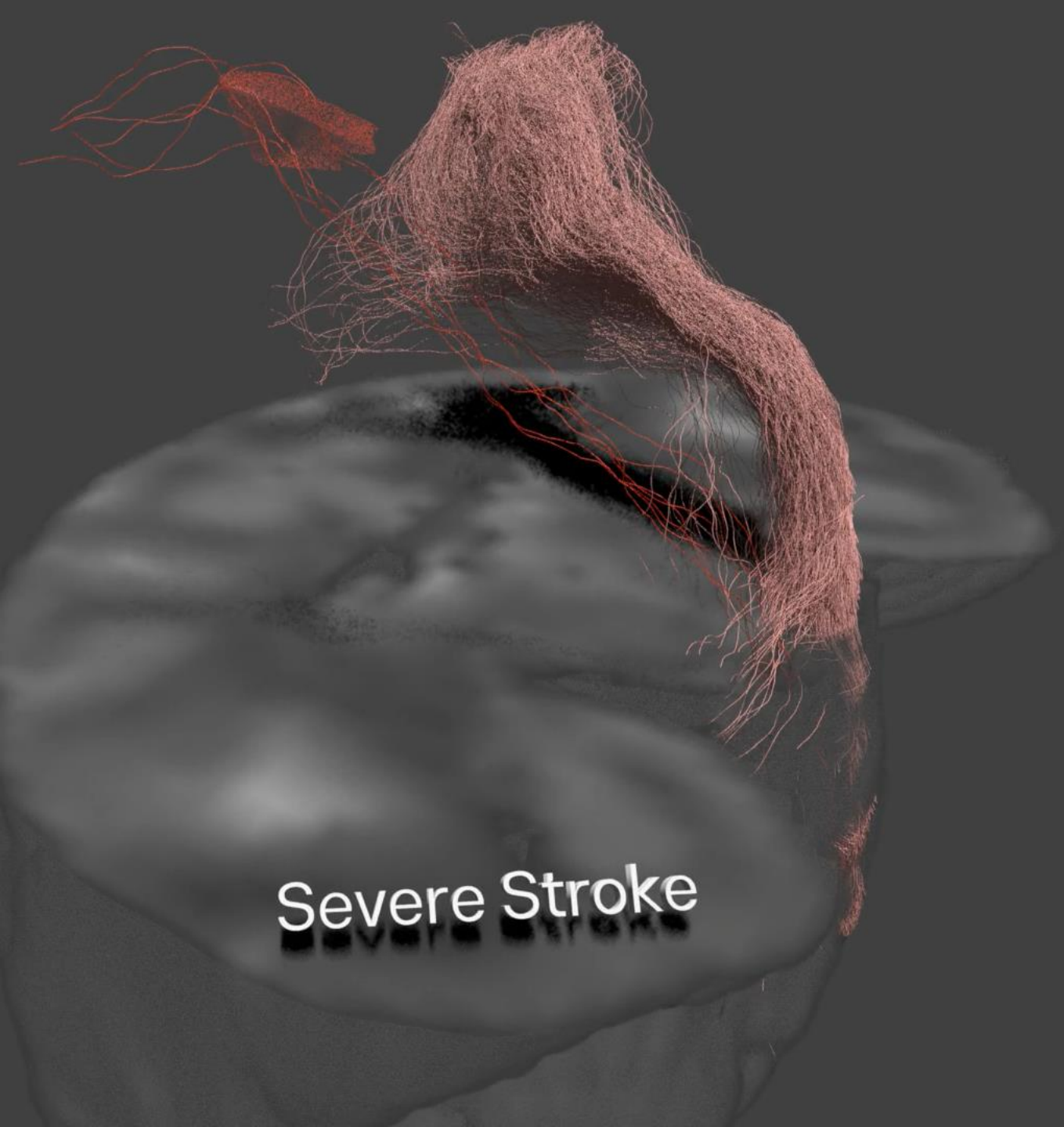
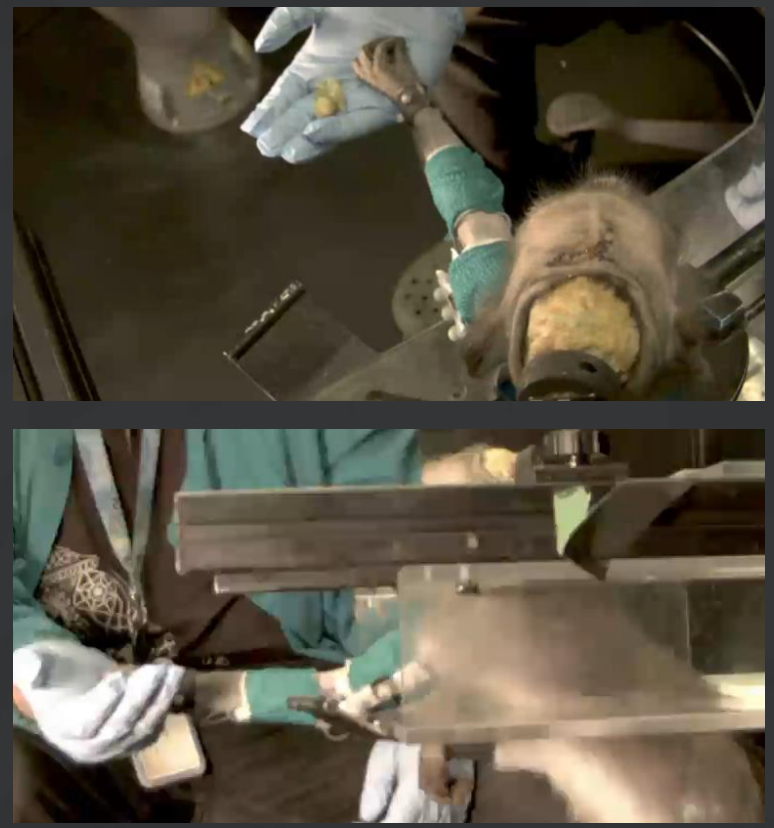
Graph representations for biology and medicine

Week 8: Neuroscience

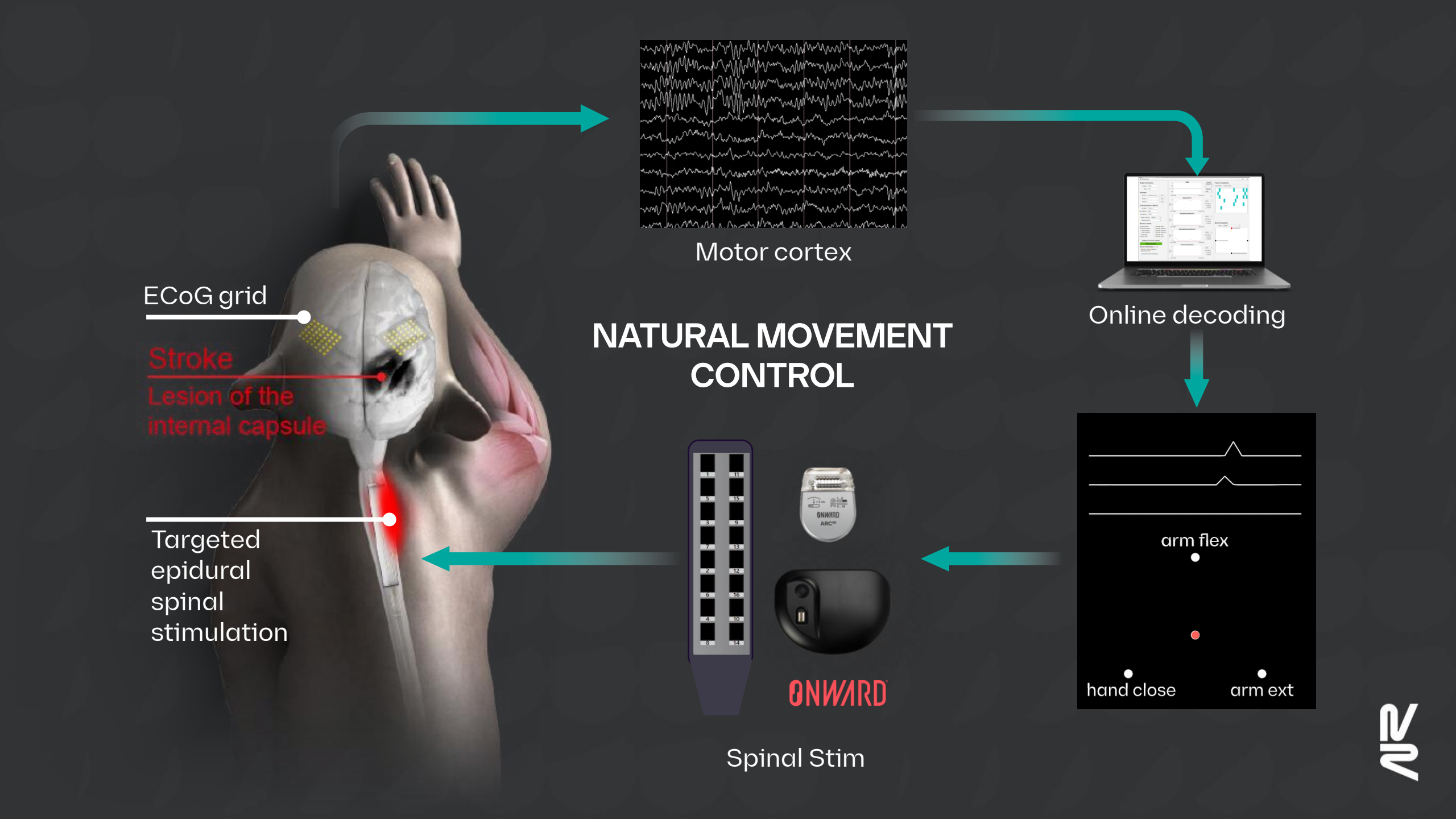
Alejandro Rodriguez Guajardo

Upper limb **severe stroke**

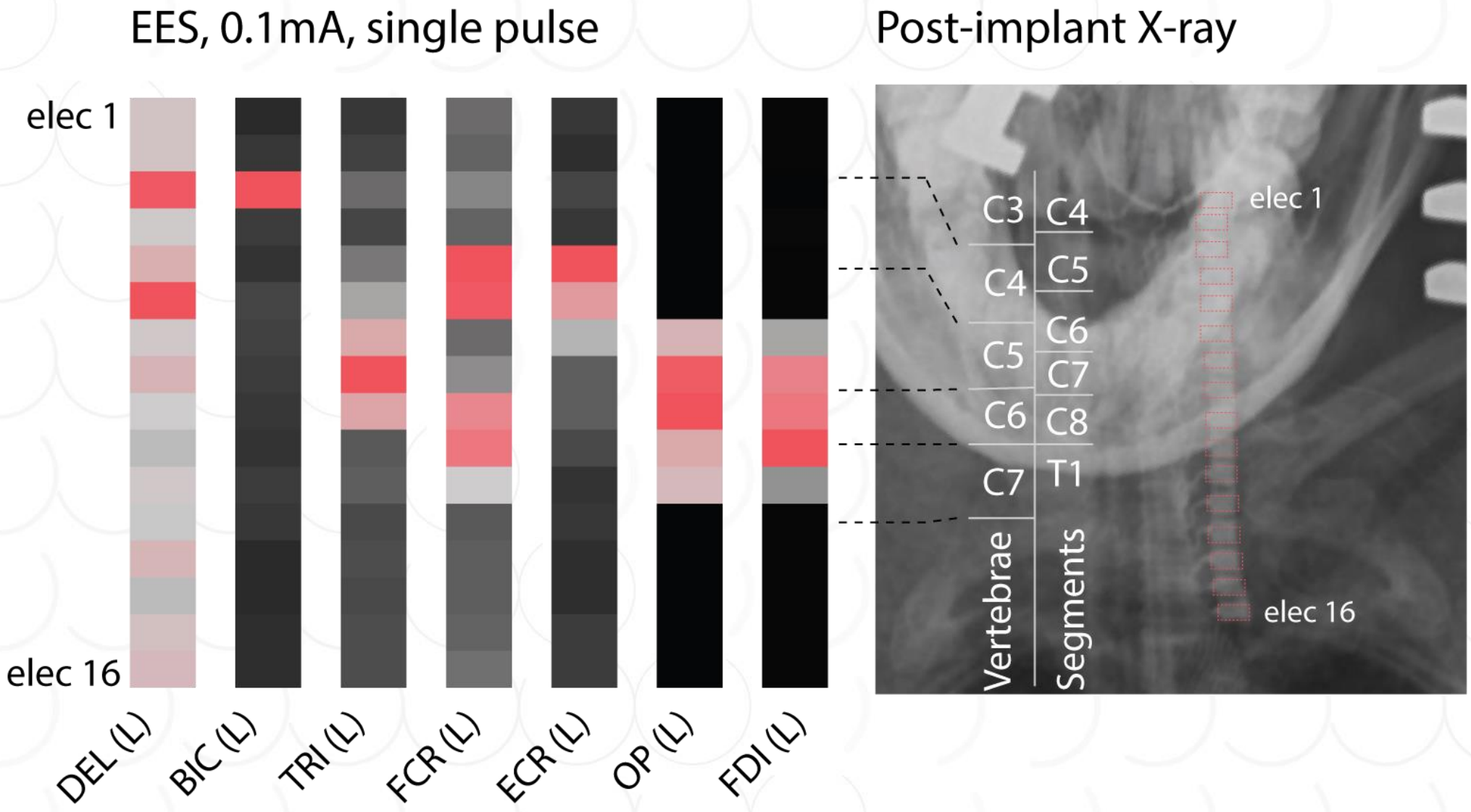


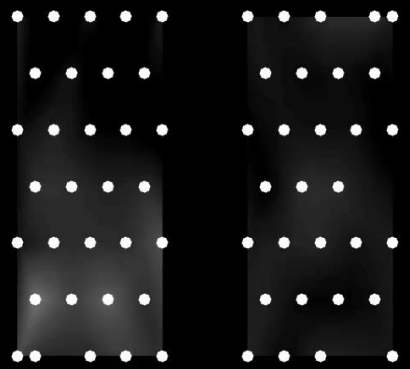


1 day post-stroke

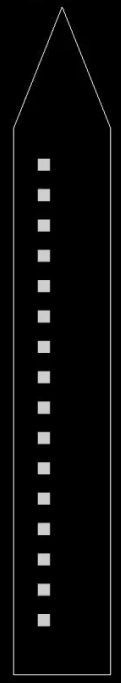

Motor cortex
Online decoding
ECoG grid
Stroke
Lesion of the internal capsule
NATURAL MOVEMENT CONTROL
Targeted epidural spinal stimulation
arm flex
hand close
arm ext
ONWARD
Spinal Stim

Intraoperative Single-Pulse Spinal Stimulation: EMG Responses





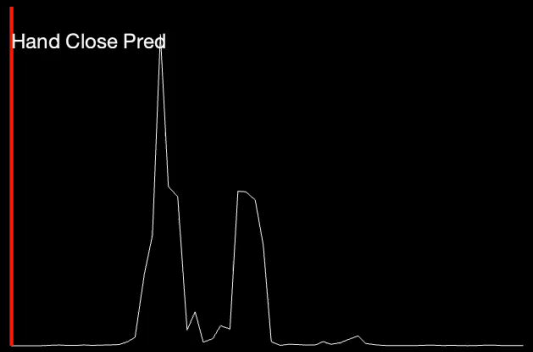
5 s

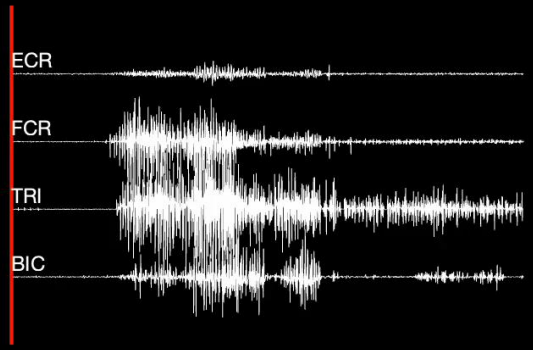


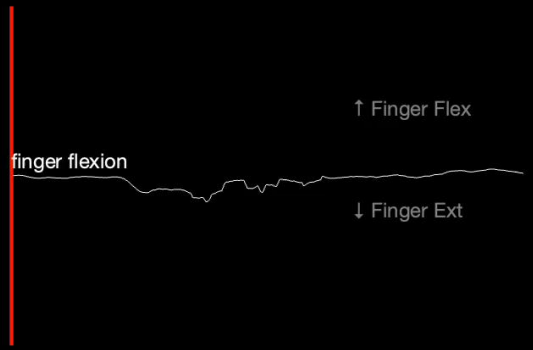
80Hz

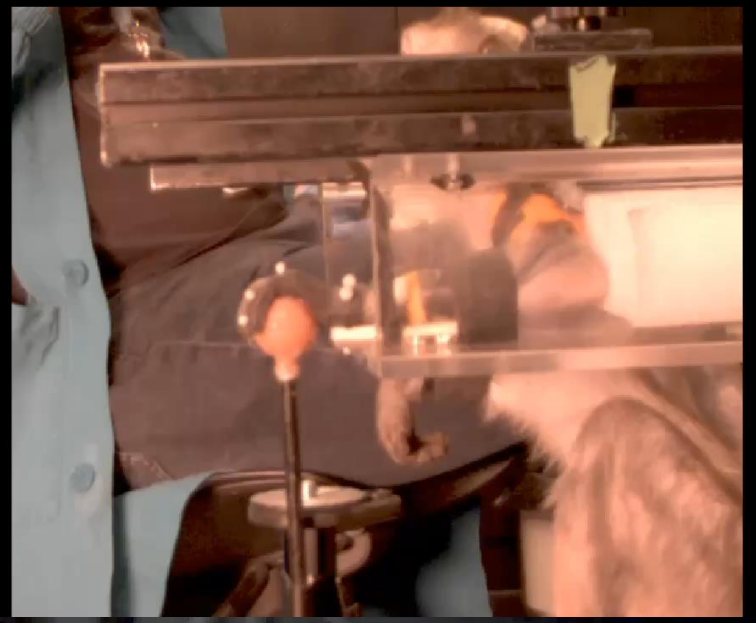
Cat:10

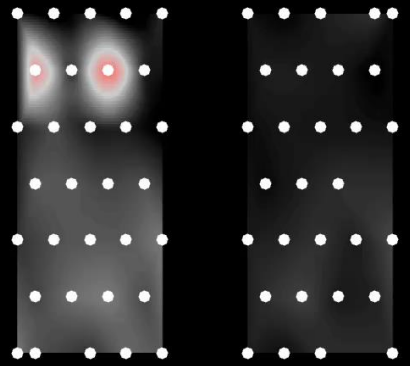
An:17



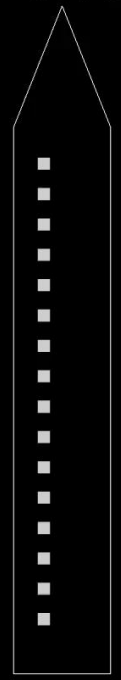








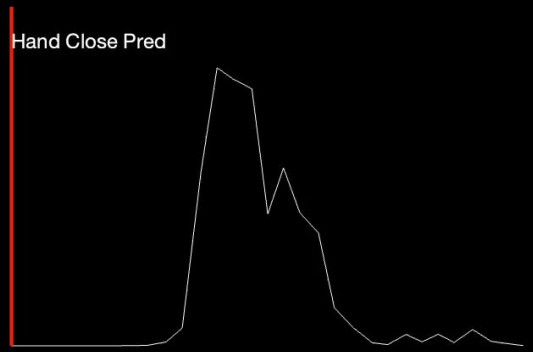
311 s

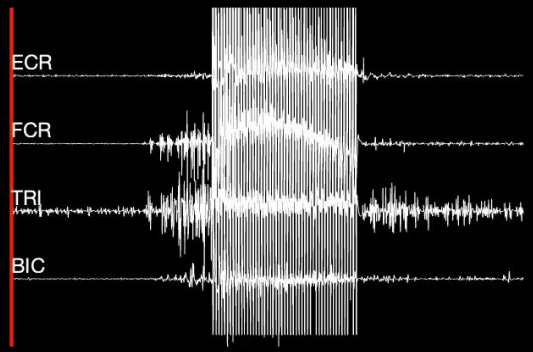


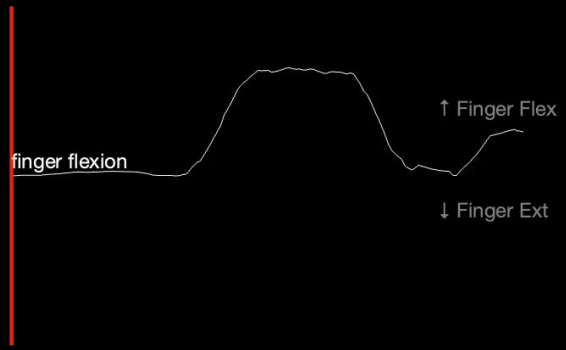
80Hz

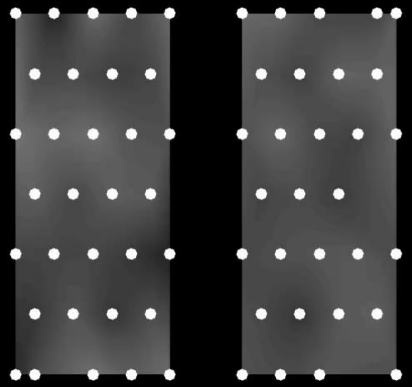
Cat:10

An:17



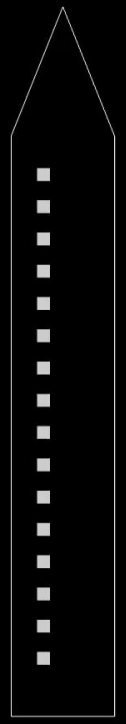






Arm Flexion Pred

572 s



ECR

FCR

TRI

BIC

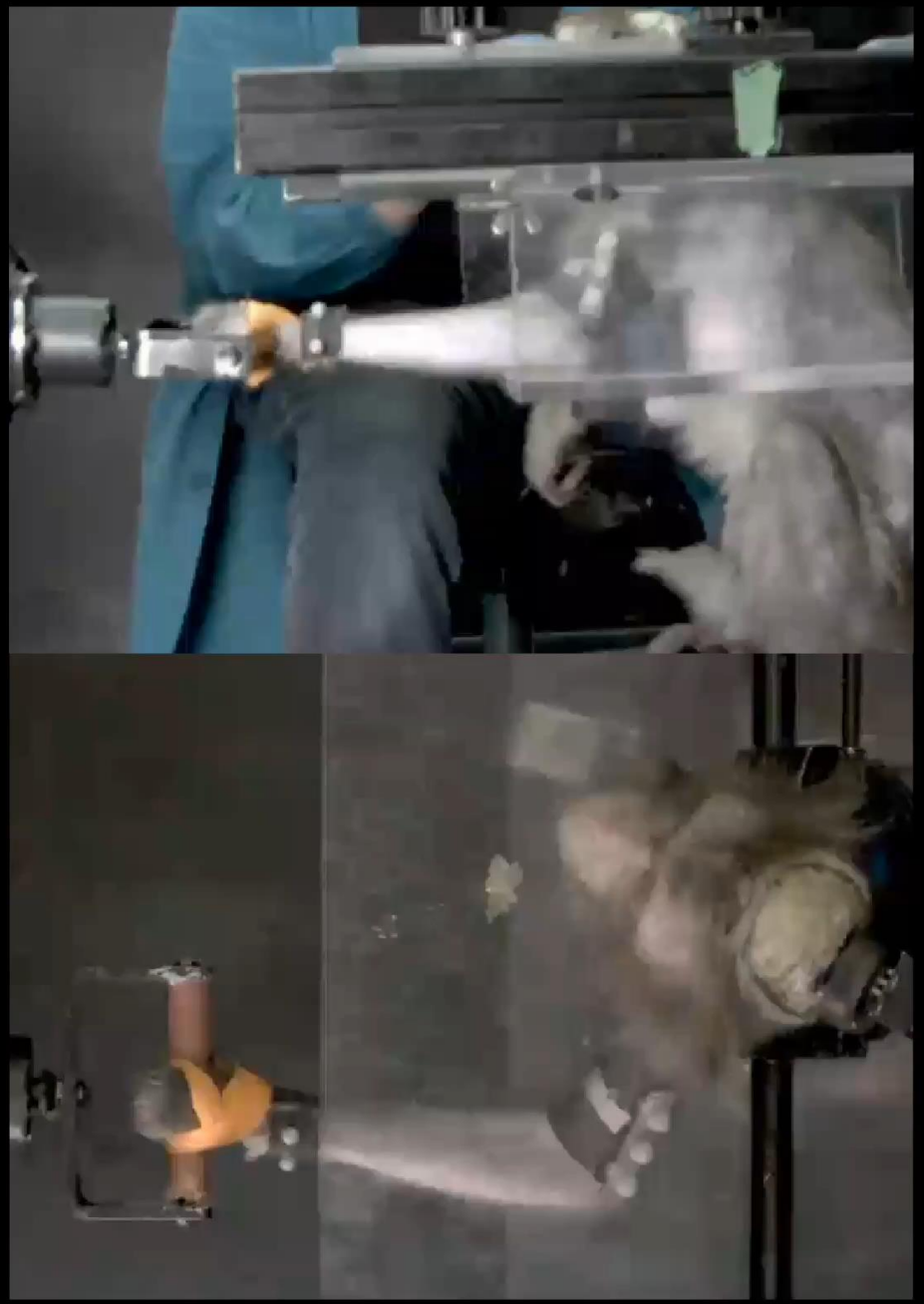
No EMG Recorded

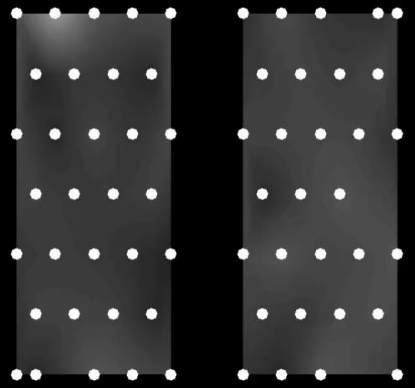
Pulling distance

80Hz

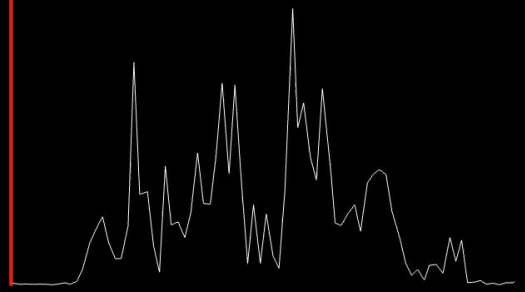
Cat:4 10

An:17

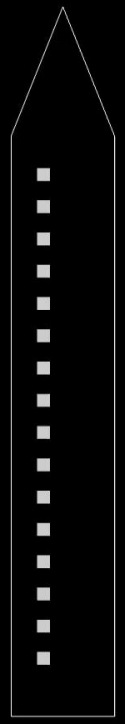




Arm Flexion Pred



537 s



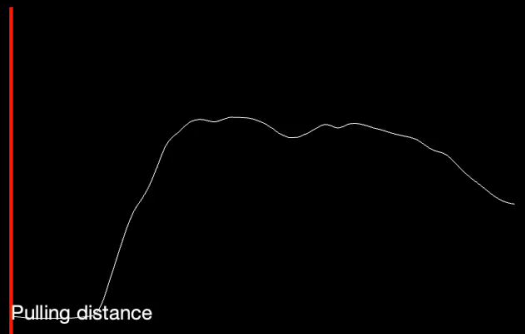
ECR

FCR

TRI

BIC

No EMG Recorded

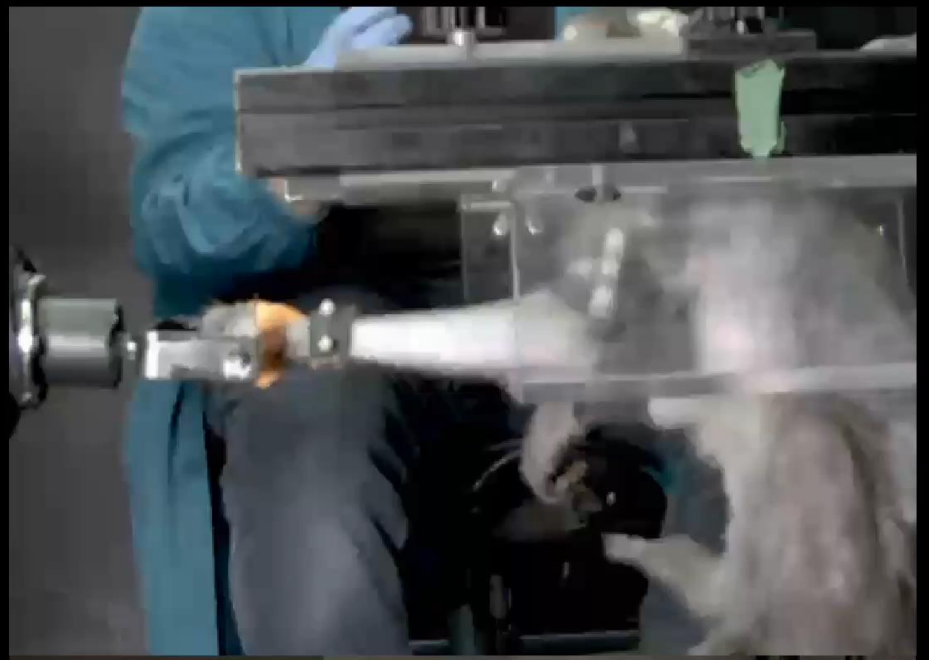


Pulling distance

80Hz

Cat:4 10

An:17

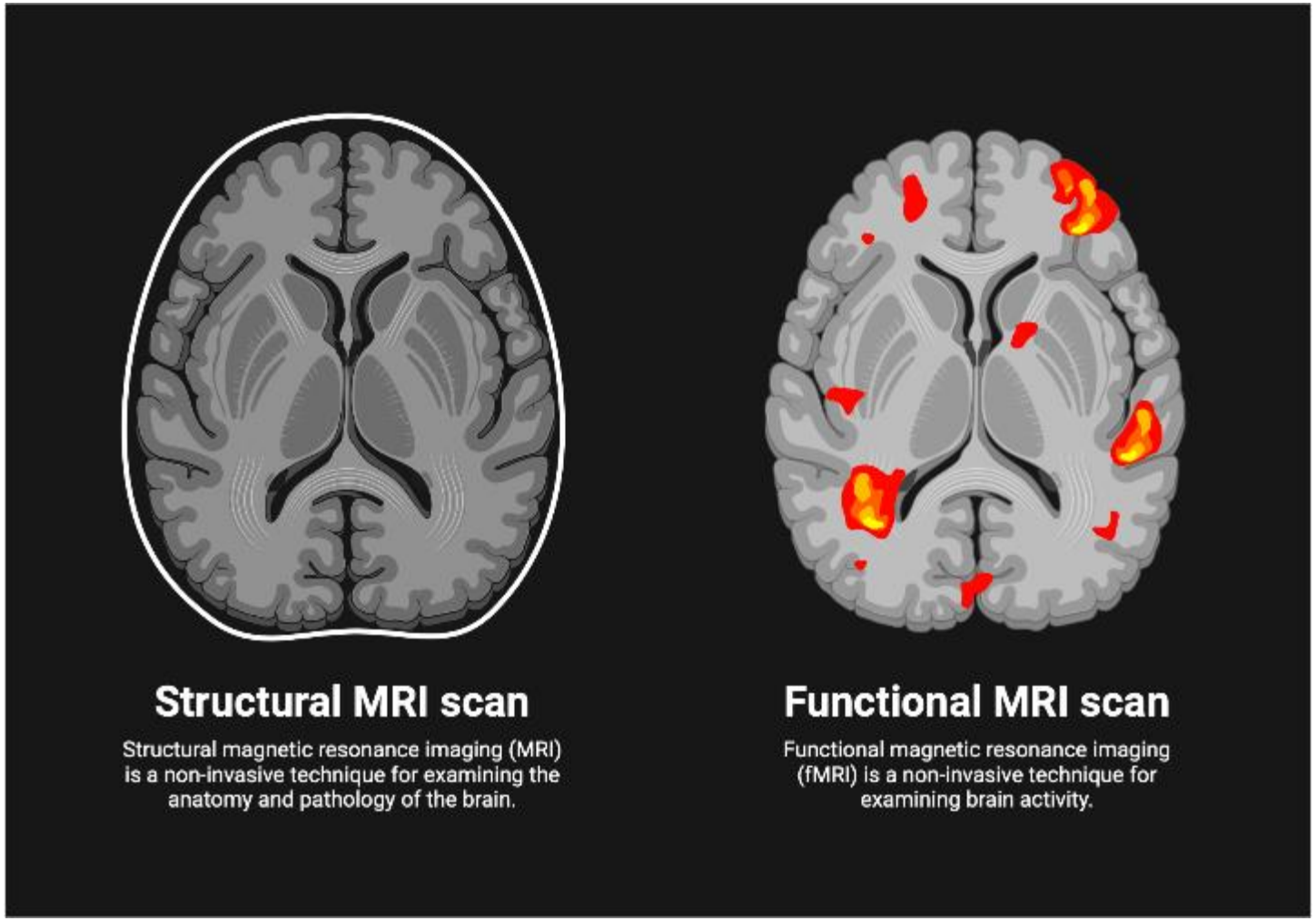


Towards Graph Neural Networks with Domain-Generalizable Explainability for fMRI-Based Brain Disorder Diagnosis

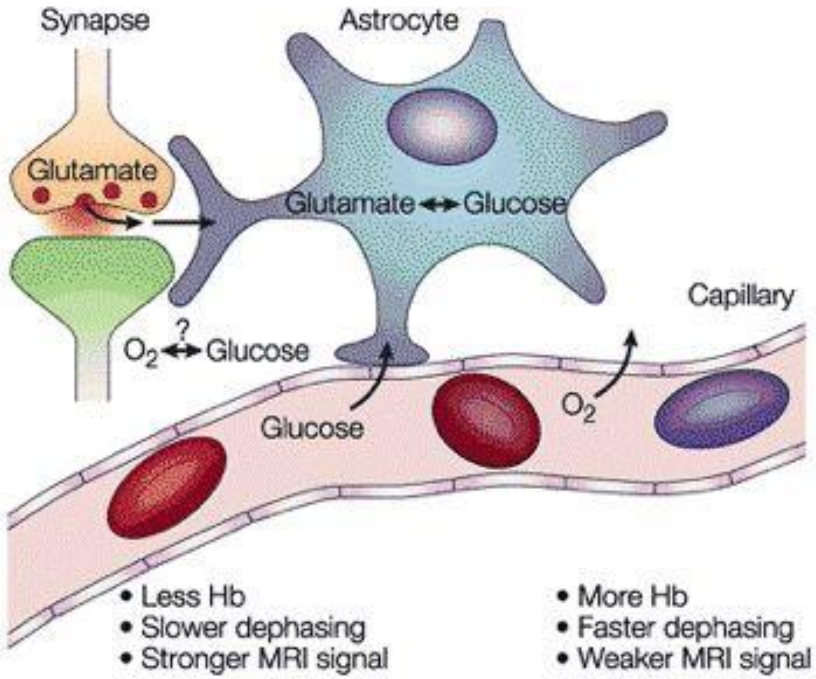
Xinmei Qiu[1], Fan Wang[2,4,✉], Yongheng Sun[1], Chunfeng Lian[1,3,✉], and Jianhua Ma[2,3,✉]

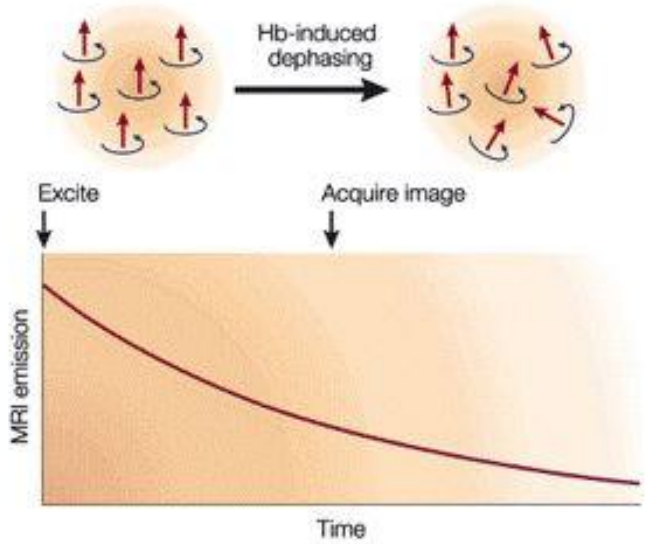
[1] School of Mathematics and Statistics, Xian Jiaotong University, Xi'an, China
chunfeng.lian@xjtu.edu.cn
[2] The Key Laboratory of Biomedical Information Engineering of Ministry of Education, School of Life Science and Technology, Xi'an Jiaotong University, Xi'an, China
{fan.wang,jhma}@xjtu.edu.cn
[3] Pazhou Lab (Huangpu), Guangzhou, China
[4] The First Affliated Hospital of Xi'an Jiao Tong University, Xi'an, China

Functional magnetic resonance imaging (fMRI)

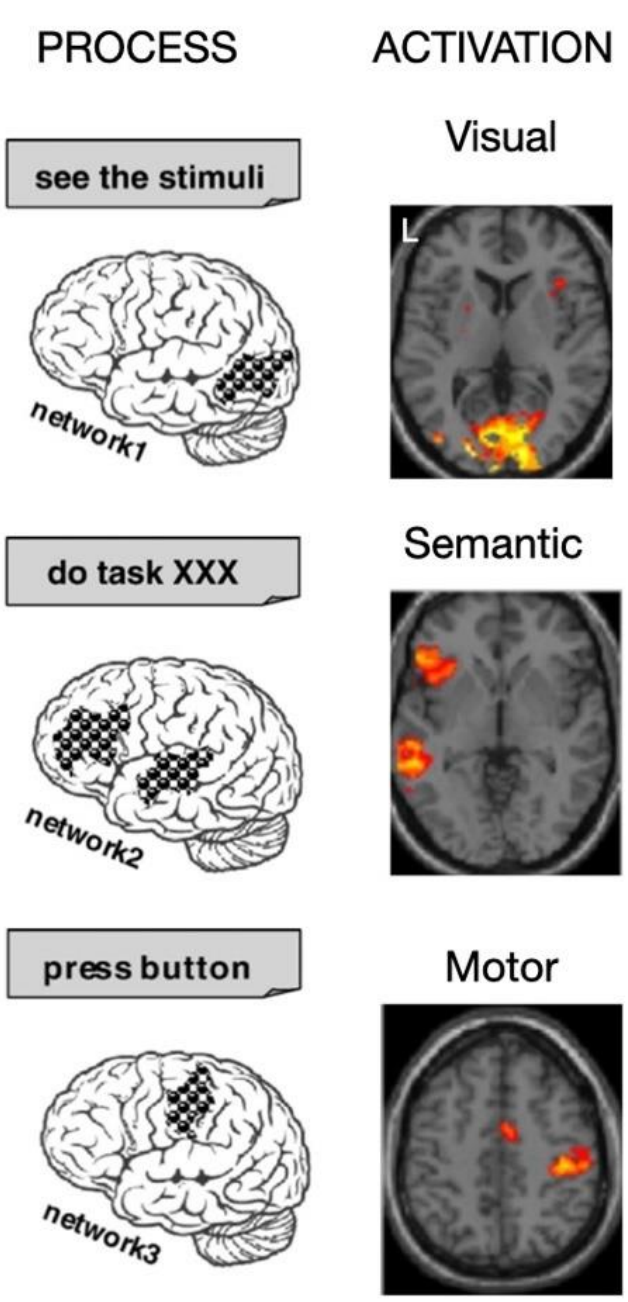


BOLD (blood oxygen-level dependent) signal

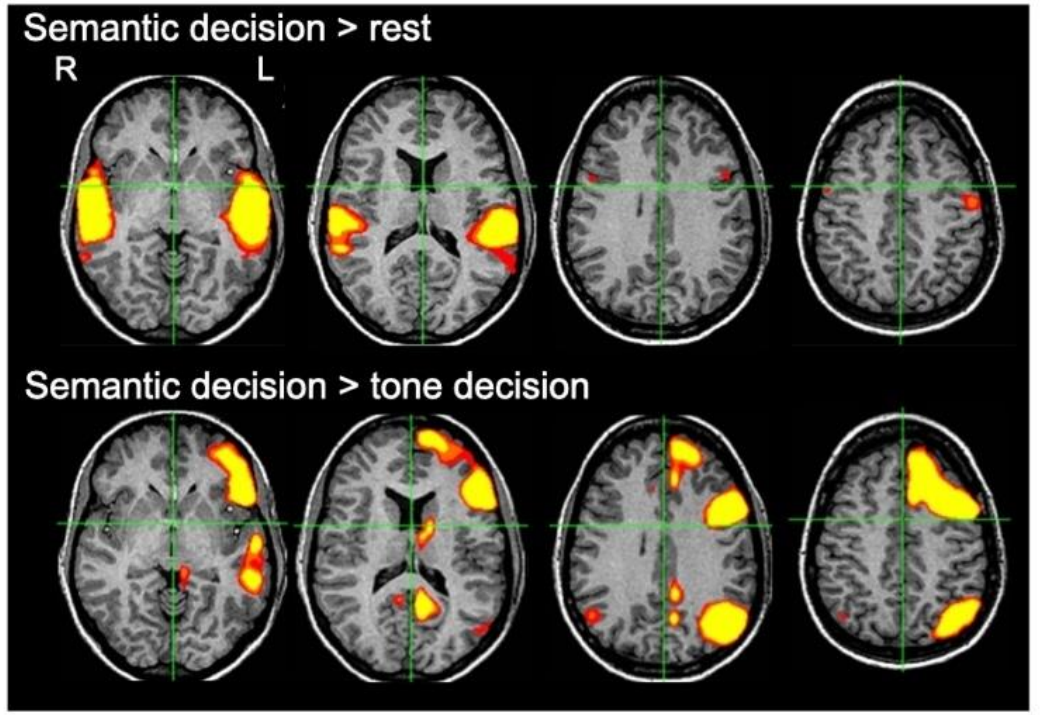




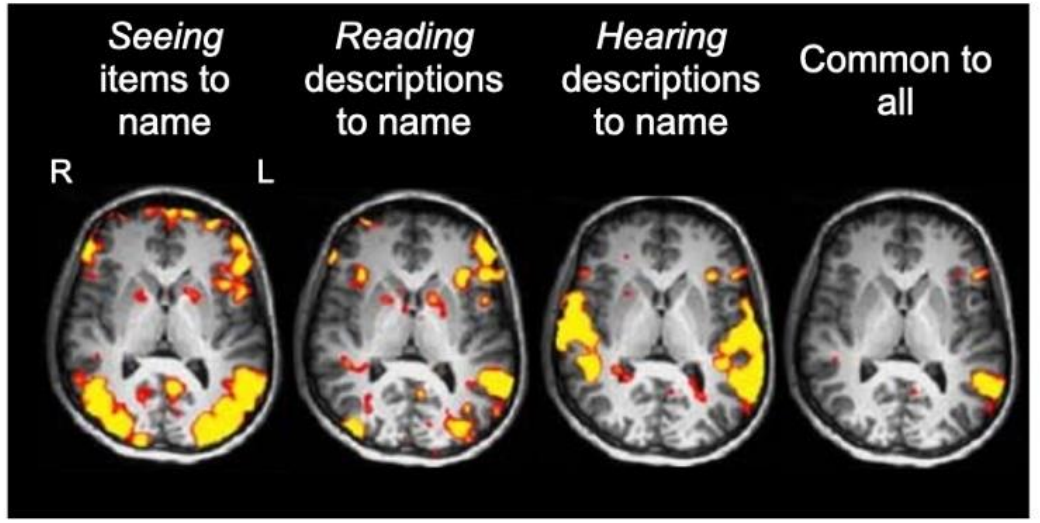
A.


PROCESS
ACTIVATION
see the stimuli
Visual
L
network1
do task XXX
Semantic
network2
press button
Motor
network3

B.

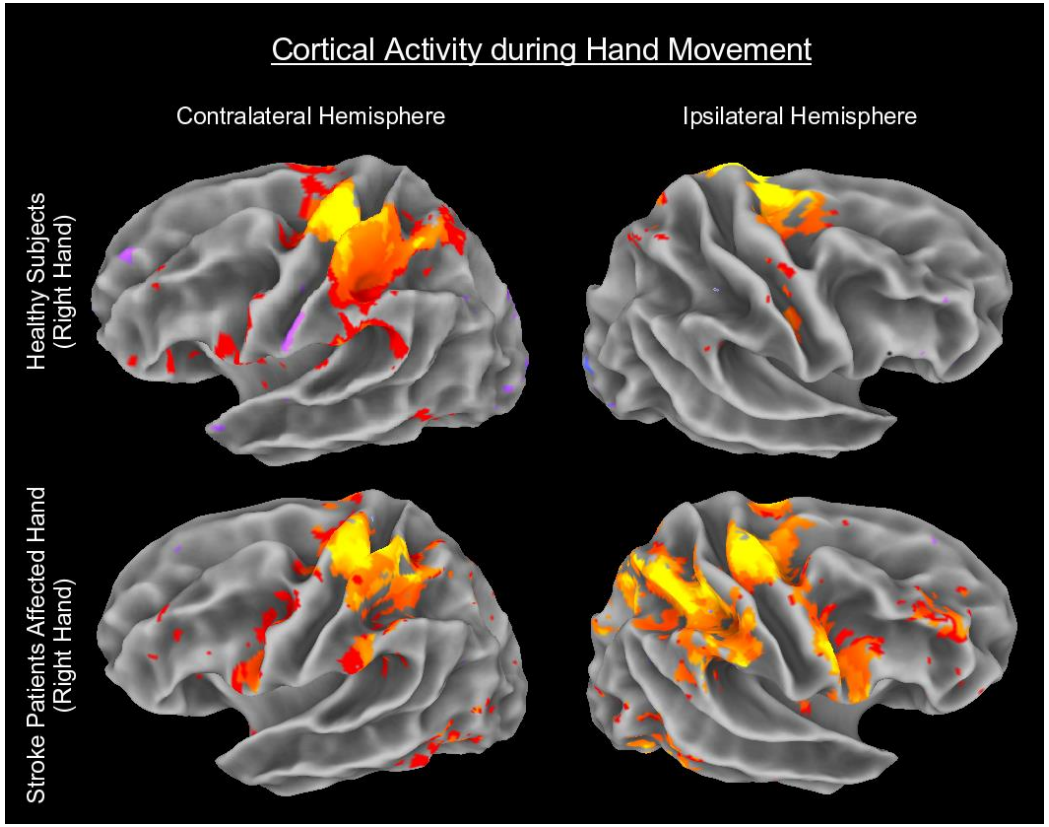

Semantic decision > rest
R
L
Semantic decision > tone decision

C.


Seeing items to name
Reading descriptions to name
Hearing descriptions to name
Common to all
R
L

How fMRI Aids in Brain Disorder Diagnosis

- Each brain region's activity is recorded as a time series.
- Statistical or machine learning methods compare connectivity patterns between healthy and diseased brains.
- **Goal: find biomarkers**, reproducible patterns that indicate specific disorders.
- Used in studies of:
  - Autism Spectrum Disorder (ASD)
  - Depression
  - Alzheimer's Disease
  - Stroke and recovery



Limitations of Current fMRI-GNN Approaches

1. Explainability

- GNNs often act like “black boxes.”
- Clinicians need to know which brain connections drive the diagnosis (biological interpretability).

2. Generalizability

- fMRI data vary across hospitals (different scanners, populations, noise).
- A model trained on one site often fails on another (domain shift problem).

Explainability-Generalizable Graph Neural Network (XG-GNN)

- Designed for fMRI-based diagnosis across multiple sites.
- Combines:
  - **Explainable learning** → finds biologically meaningful brain connections.
  - **Domain generalization** → works across hospitals without retraining.
- Achieved through:
  - A **meta-learning framework** (learning to generalize).
  - **Explainability-generalizable regularizations (Lsp & Lcons)** that enforce sparsity and cross-site consistency.

Explainability-Generalizable Graph Neural Network (XG-GNN)

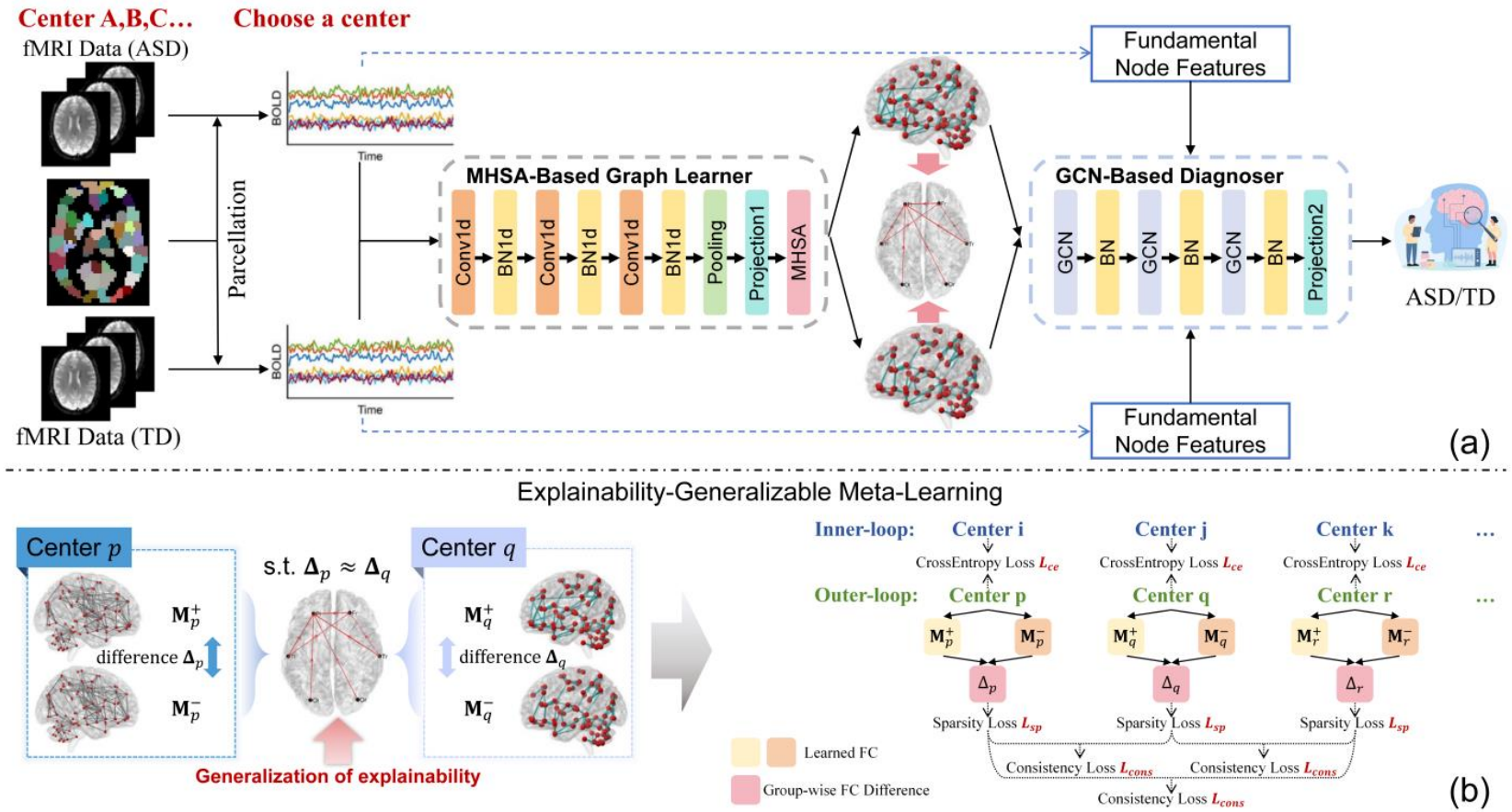


Fig. 1: Illustration of our XG-GNN for multi-site brain disorder diagnosis.

What the model does overall:

The goal is to diagnose brain disorders (e.g. autism) from **fMRI** data recorded at multiple hospitals (domains).

Each subject's fMRI gives you:

- n= number of brain regions of interest (ROIs)
- T= timepoints per region

So each person's fMRI data looks like a matrix

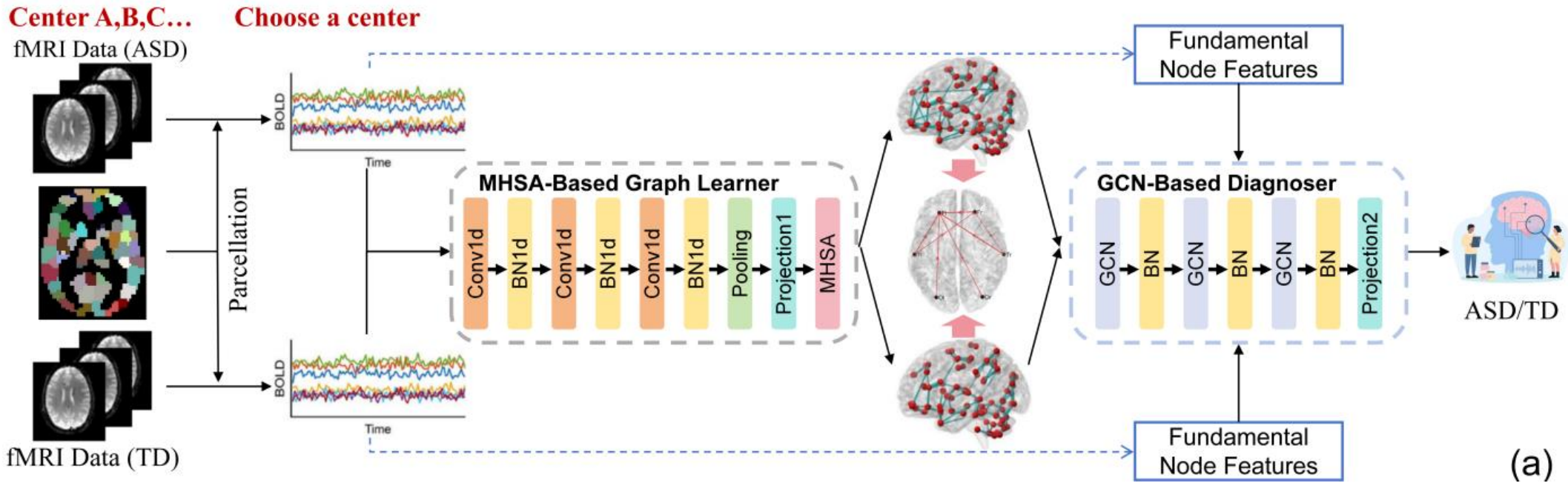
$$X \in \mathbb{R}^{n \times T}$$

Each row = the BOLD signal (activity over time) of one region.

The model must:

1. Build a **functional brain graph** — which regions are connected.
2. Use that graph to **predict** diagnosis (ASD vs. typical).
3. Keep the learned explanations **consistent across hospitals**.

Arquitecture



MHSA-Based Graph Learner → Building the Brain Graph

1. **Input:** $X \in \mathbb{R}^{n \times T}$: fMRI BOLD signals (n = ROIs, T = timepoints).

2. **Temporal feature extraction:** Three Conv1D + BN1D + ReLU layers capture temporal patterns per ROI:

$$F_i = BN1d_i\big(Conv1d_i(X)\big) = BN1d_i(XW_i + b_i)$$

Output $F_i \in \mathbb{R}^{n \times h_i}$

3. **Pooling & projection**: Channel-wise max-pool → 2 linear + softmax → ROI embeddings

$$f \in \mathbb{R}^{n \times h}$$

4. **Multi-Head Self-Attention (MHSA):** Learns cross-region dependencies:

$$Q = W_q f + b_q,\ K = W_k f + b_k,\ V = W_v f + b_v$$

$$\text{MHSA}(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$$

Each region attends to all others; multi-heads capture diverse interactions.

5. **Connectivity matrix**: Merge heads → embedding $A \in \mathbb{R}^{n \times m}$;define

$$M = AA^T$$

Output: nonlinear, subject-specific **functional brain network** M used by the **GCN-based Diagnoser** for classification.

GCN-Based Diagnoser → Classifying Brain Graphs

1 **Input graph** G = (V, E,M)

V: brain regions (nodes)
$M \in \mathbb{R}^{n \times n}$:connectivity matrix learned previously
Initial node features g_0 :cross-ROI **Pearson correlation coefficients**

2 **Network structure**: **Three GCN blocks**, each composed of:

Graph Convolution → Batch Normalization (BN) → LeakyReLU

3 **Node feature update rule**

For block *i* (with parameters U_i)(

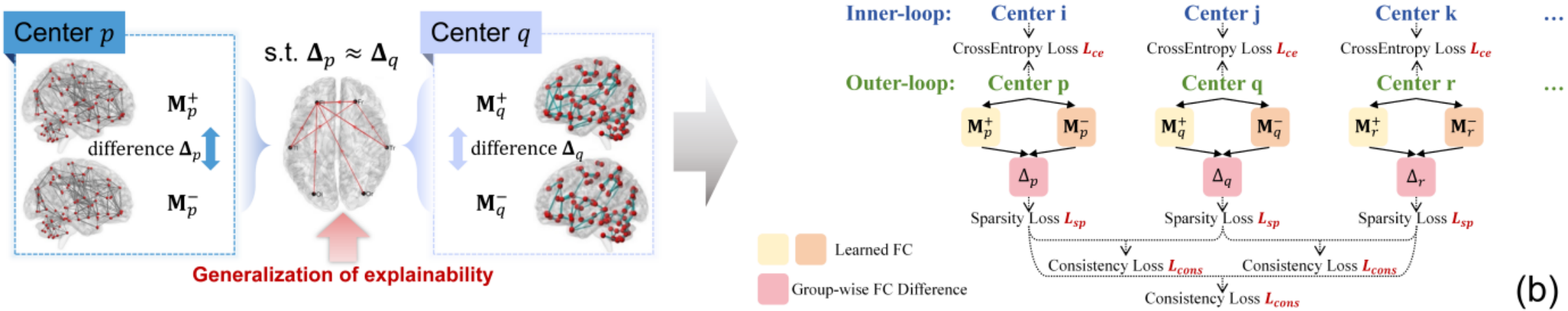
$$g_i = BN\big(GCN(M, g_{i-1})\big) = BN\big(ReLU(M g_{i-1} U_i)\big)$$

Where:
$g_{i-1} \in \mathbb{R}^{n \times r_{i-1}}$:input node features
$g_i \in \mathbb{R}^{n \times r_i}$:updated features
$M g_{i-1} U_i$:propagates information across connected regions.

4 **Output layer**

Flatten node-level features → **graph-level representation**
Feed into **linear projection layers** → final **diagnostic outcome** (e.g., ASD / TD)

Explainability-Generalizable Meta-Learning:



1) Explainability-Generalizable (XG) Regularization:

1 Sparsity of Inter-Group FC Differences (Lsp)

- Functional abnormalities in early disorders (e.g., ASD) are usually **localized**, not widespread.
- Enforce sparsity in FC differences between patients and controls.
- For each site D_k:

$$\Delta_{D_k} = | M^{+}_{D_k} - M^{-}_{D_k} | \in \mathbb{R}^{n \times n}$$

$$L_{sp} = -\| \sum_{k} \Delta_{D_k} \log \Delta_{D_k} \|_2$$

2 Cross-Site Consistency of FC Differences (Lcons)

- True disorder-related FC differences should be stable across centers.
- Encourage inter-site agreement of group-wise patterns:

$$L_{cons} = -\sum_{i,j} \| \frac{| \Delta_{D_i} \cdot \Delta_{D_j} |}{| \Delta_{D_i} | \quad | \Delta_{D_j} |} \|_2$$

- →Promotes domain-agnostic, center-independent explanations.

Combined effect:

$$L_{meta} = L_{ce} + \alpha L_{sp} + \beta L_{cons}$$

Together, L_{sp} +L_{cons} form the **XG regularization**, ensuring the learned biomarkers are **sparse, consistent, and generalizable** across fMRI sites.

2) Bi-Level Meta-Learning Algorithm:

1 Setup

Multi-site dataset: $S = \{D_1, D_2, . , . . D_K\}$)each D_k =one fMRI center).
Simulate unseen domains during training to enhance stability of learned explanation factors.

2 Inner Loop – Classification training

Randomly select subset $S_{inner} \subset S$.
Train model with **cross-entropy loss** $L_{ce}$for diagnosis (ASD vs TD).

$$\theta' = \theta - \gamma \nabla_\theta L_{ce}(S_{inner})$$

3 Outer Loop – Meta-generalization update

Use remaining centers S_{outer}.
Update model with **meta-loss** combining accuracy and explainability terms:

$$L_{meta} = L_{ce} + \alpha L_{sp} + \beta L_{cons}$$
$$\theta = \theta - \gamma \nabla_{\theta'} L_{meta}(S_{outer})$$

4 Outcome

- Inner loop: learns diagnostic performance on known domains.
- Outer loop: enforces **cross-domain generalization** and **stable explanation factors** via XG regularization.
- Mimics “training on some sites, testing on unseen ones” within each iteration.

Experiments

Dataset:

- **ABIDE (Autism Brain Imaging Data Exchange)**
  - Resting-state fMRI from **16 international centers**
  - **416 ASD** and **418 TD (Typical Development)** participants

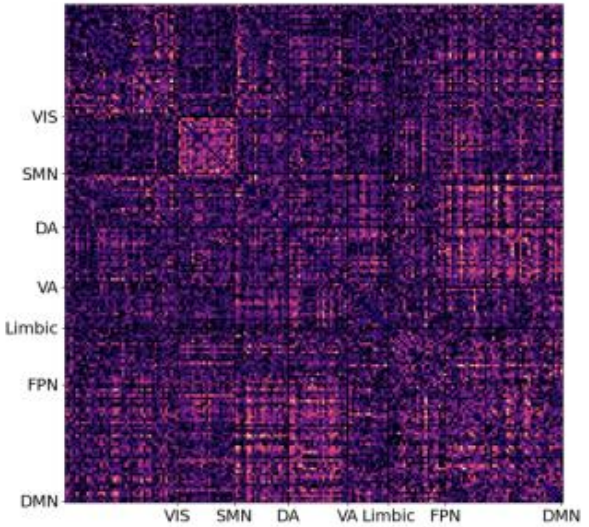
- **Preprocessing:**
  - Brain parcellation using **CC200 atlas** (200 ROIs)
  - Averaged BOLD signals within each ROI → input to model

Table 1: Diagnostic results (mean±std) for different competing methods on the three target domains from the ABIDE dataset.

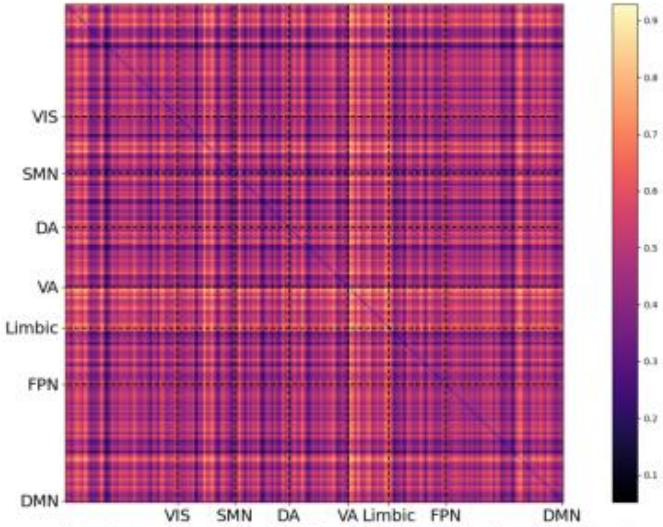
Method	ACC(%)	AUC(%)	F1 Score(%)
BrainnetCNN [9]	65.35 ± 2.20	70.87 ± 1.62	62.04 ± 1.54
IBGNN [2]	65.26 ± 1.61	72.88 ± 2.41	62.01 ± 2.93
Coral[17]	65.46 ± 2.22	70.75 ± 1.99	64.97 ± 4.31
GenM [12]	62.98 ± 4.79	69.40 ± 6.51	57.22 ± 9.31
DuMeta[18]	68.78 ± 5.02	73.41 ± 2.78	69.44 ± 5.33
XG-GNN **(ours)**	$\mathbf{70.66 \pm 1.07}$	$\mathbf{76.32 \pm 0.55}$	$\mathbf{69.78 \pm 3.82}$

Table 2: Ablation studies regarding each key component of our XG-GNN.

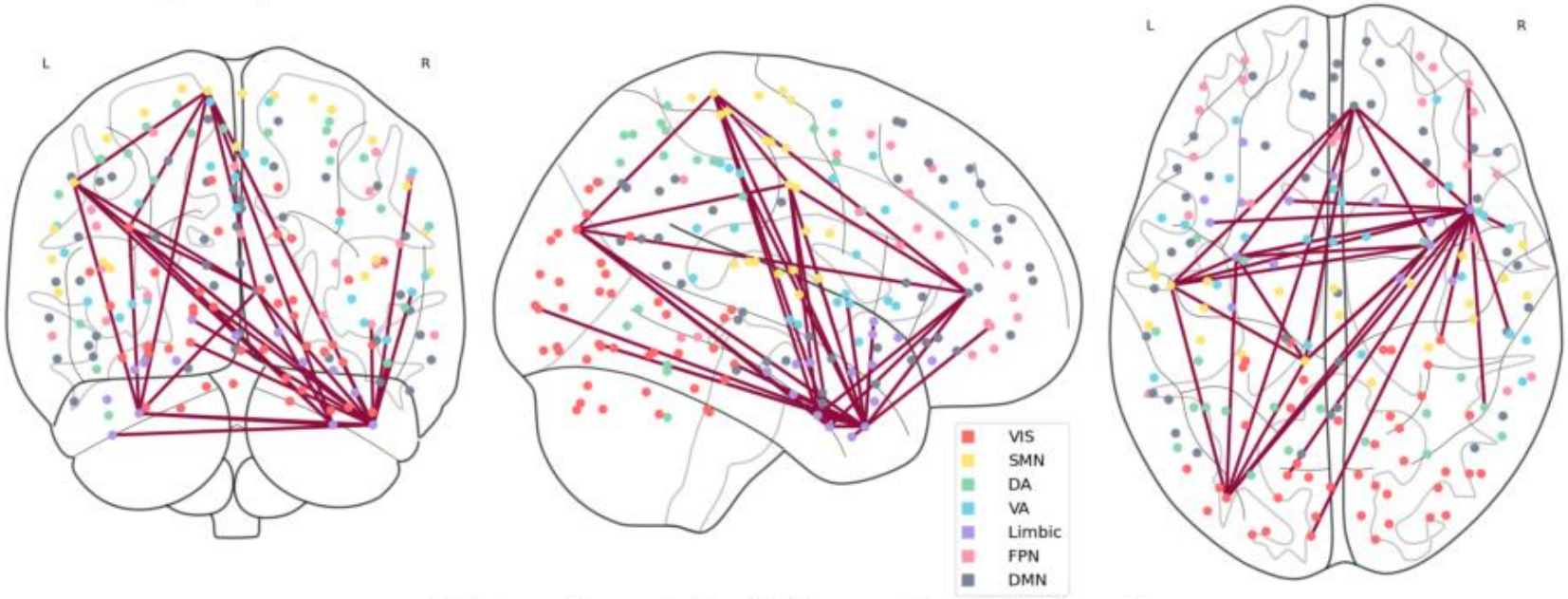
Ablation experiment	Baseline	w/o XG	w/o L_{cons}	w/o L_{sp}	XG-GNN **(ours)**
ACC(%)	64.66	68.80	69.97	69.04	**70.66**
AUC(%)	70.27	75.57	**77.47**	75.94	76.32



(a) Group-wise difference of the linear FCs



(b) Group-wise difference of the learned FCs



(c) Interpretive analysis of differences that can be learned.

Fig. 2: Explanation results on ABIDE dataset.

Conclusions

- They proposed an **explainability-generalizable Graph Neural Network (XG-GNN)** for domain generalization of brain disorder diagnosis across multi-center fMRI data.

- A **meta-learning framework** integrating **specialized regularizations** was developed to learn **task-oriented brain networks** that capture **center-agnostic explanation factors**, enhancing discriminative graph representation learning and diagnostic outcomes.

- Experimental results on the **ABIDE dataset** verified the effectiveness of the proposed XG-GNN in terms of both diagnostic accuracy and explainability.

Summary

Strengths

- **Unified explainability & generalization:** First GNN (XG-GNN) to jointly address both for fMRI diagnosis.
- **Meta-learning with XG regularization:** Improves robustness to unseen centers.
- **Nonlinear graph learning:** MHSA module captures richer brain connectivity than linear FCs.
- **Strong performance:** Outperforms baselines on ABIDE (ACC ≈ 70.7%, AUC ≈ 76.3%).
- **Neuroscientific validity:** Highlights limbic-system patterns consistent with ASD findings.

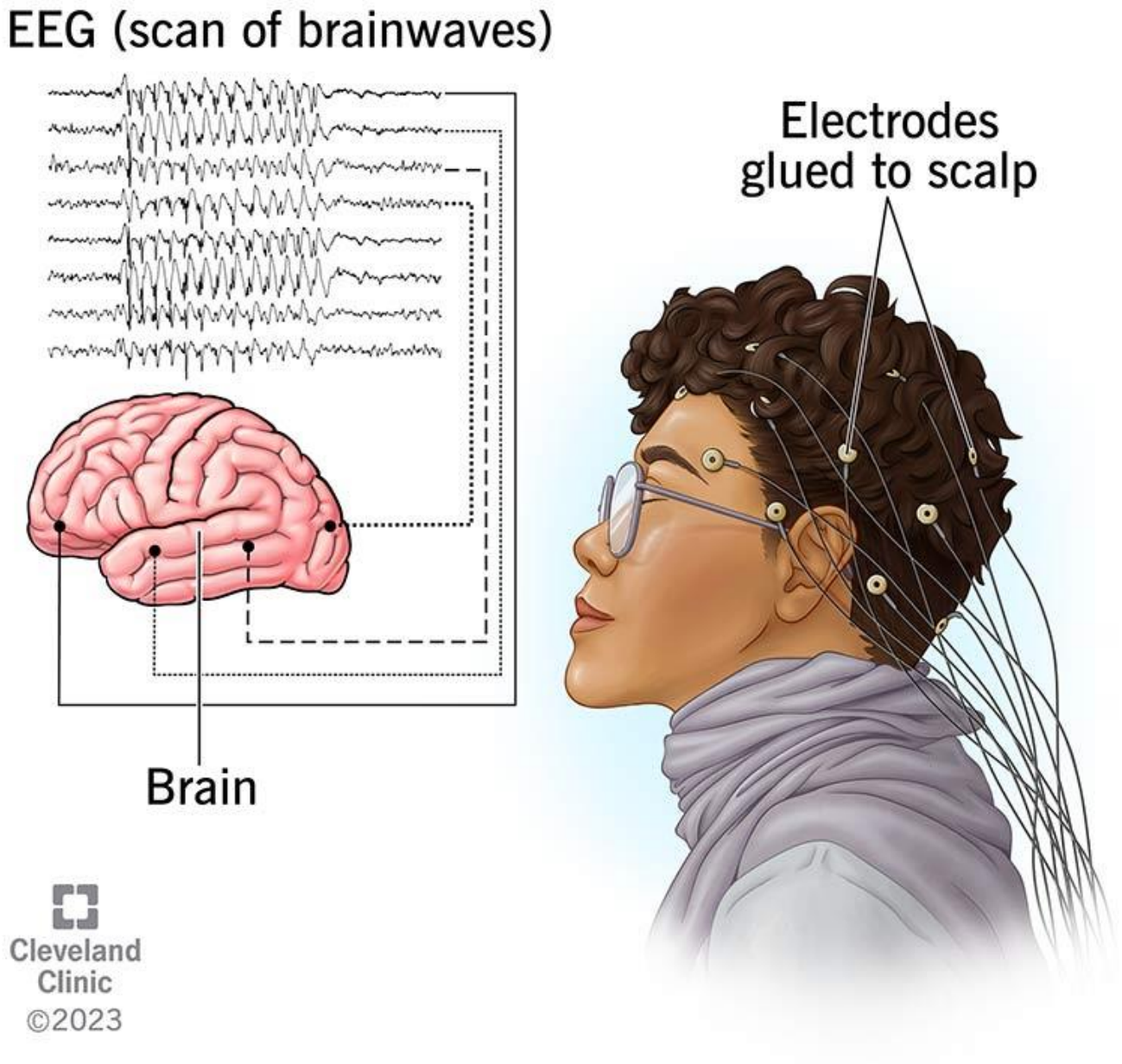
Weaknesses

- **Single dataset (ABIDE):** Limited cross-disorder validation.
- **Training overhead:** Bi-level meta-learning adds complexity.
- **Qualitative explainability:** No quantitative or clinical validation.
- **Residual domain bias:** fMRI preprocessing differences may still affect generalization.
- **Group-level focus:** Limited subject-specific temporal insight.

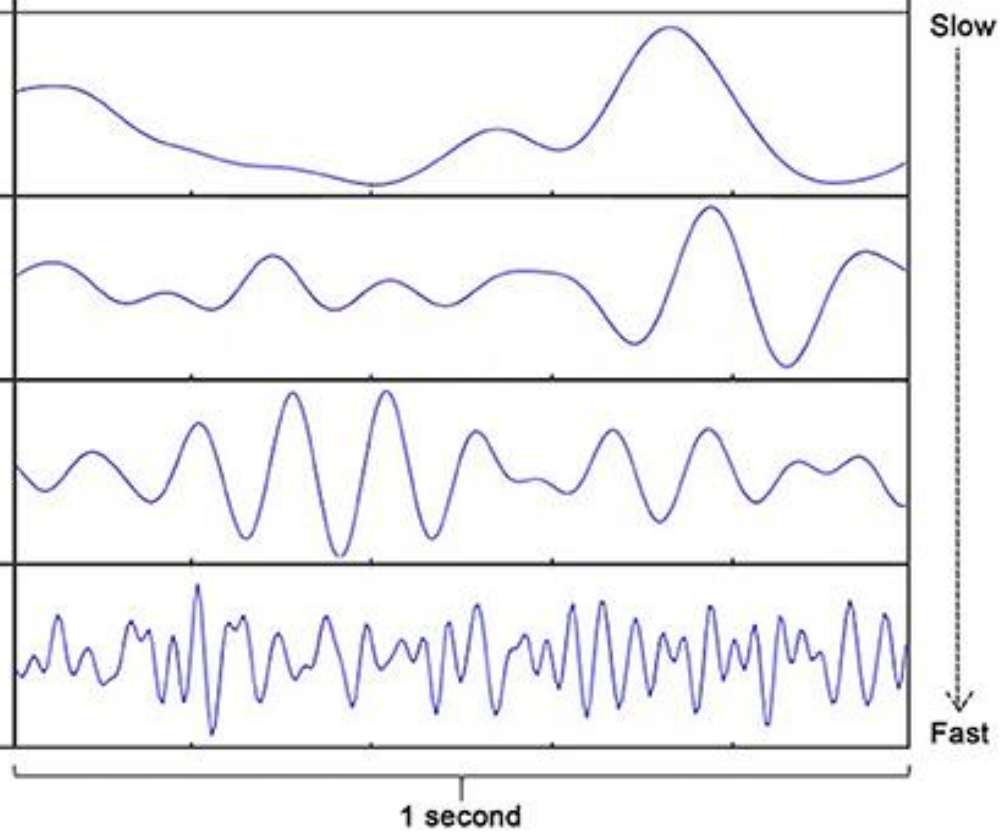
REST: Efficient and Accelerated EEG Seizure Analysis through Residual State Updates

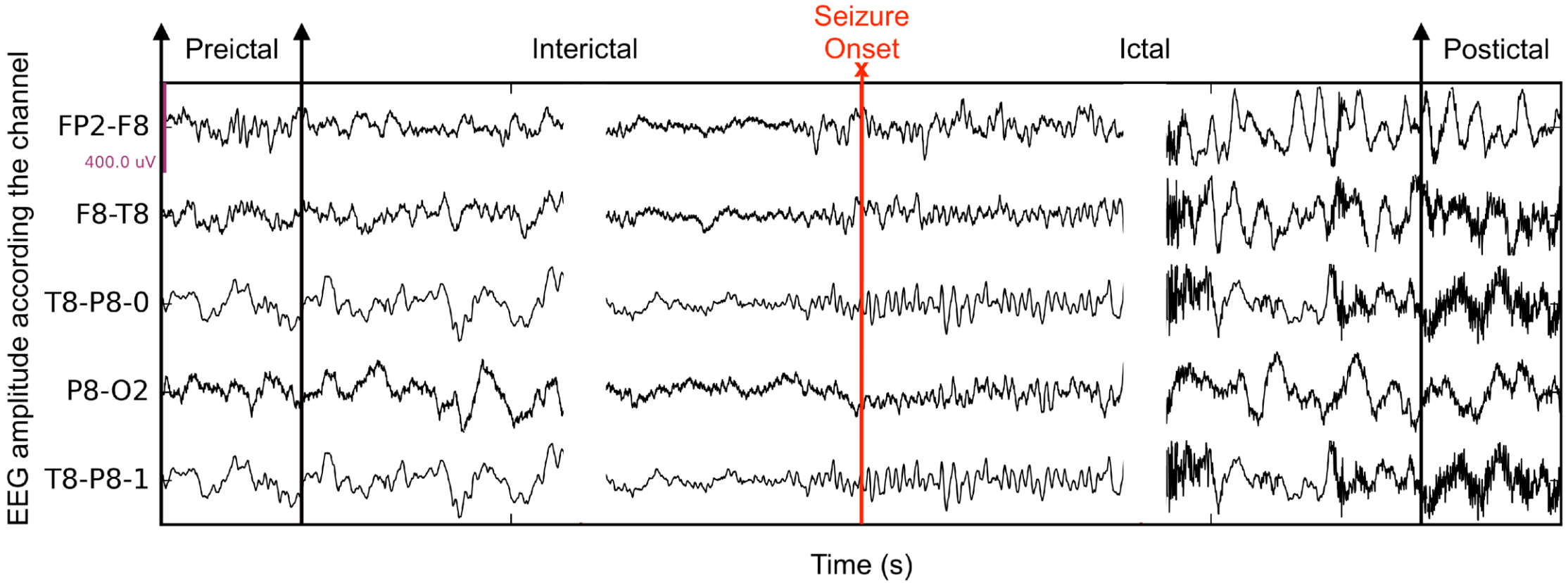
Arshia Afzal [1 2] Grigorios Chrysos [3] Volkan Cevher [* 2] Mahsa Shoaran [* 1]

Electroencephalogram (EEG)



Frequency band	Speed (Hz)	Mental state	Electroencephalography (EEG) recording
Delta	1-4	Deep sleep	
Theta	4-8	Drowsy	
Alpha	8-12	Relaxed	
Beta	12-30	Focused	




Preictal
Interictal
Seizure Onset
Ictal
Postictal
FP2-F8
400.0 uV
F8-T8
T8-P8-0
P8-O2
T8-P8-1
EEG amplitude according the channel
Time (s)

Why a new model for EEG seizure analysis?

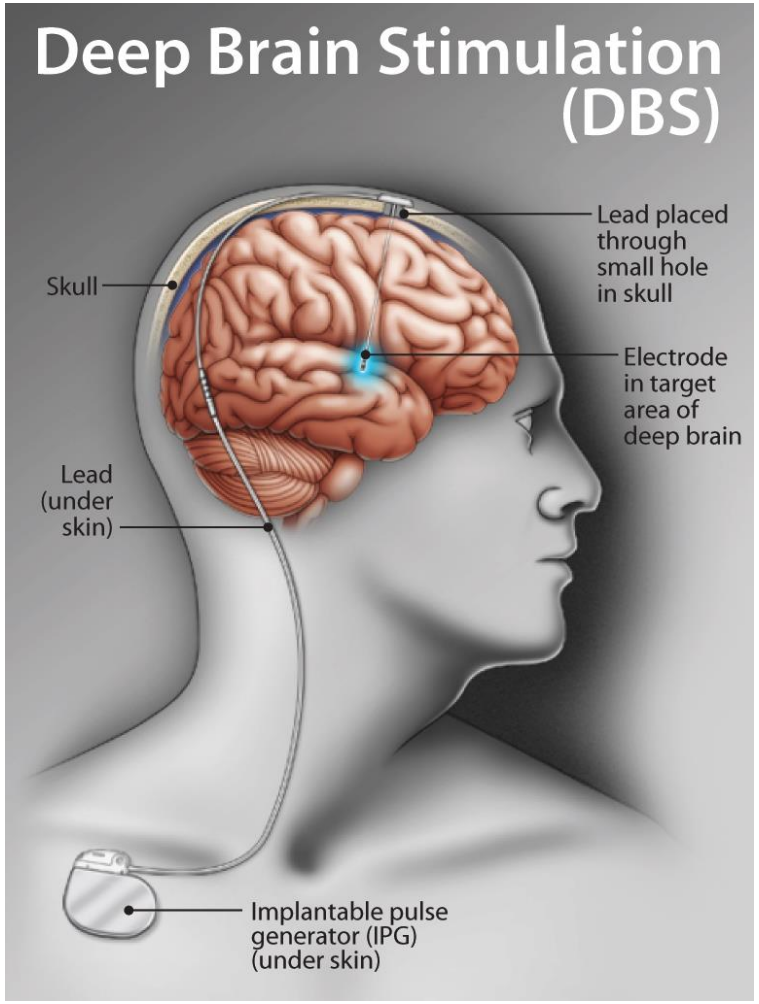
- **Clinical need**
  - Epilepsy affects ~50M people; timely seizure detection is critical for interventions (RNS/DBS, alerts).
  - Experts still spend hours–days manually reviewing EEG

- **Limits of current deep models (CNN/RNN/GNN/Transformers)**
  - **Slow inference** (e.g., gated RNNs, deep stacks, attention over long contexts).
  - **Large memory footprint** (hundreds of K to millions of params) — impractical for edge/implantable devices.
  - Often tuned for **long windows** → poorer responsiveness for real-time detection.
  - Many treat EEG as Euclidean images, **ignoring electrode geometry** and brain-network structure.

Why a new model for EEG seizure analysis?

Requirements for implantable devices:

- **Very low latency** to trigger therapy quickly.
- **Tiny model size** and memory use to fit tight hardware/power budgets.
- **Stable accuracy** across clip lengths for robust real-time operation.



REST (Residual State Updates)

- A **graph-based recurrent model** for **fast, efficient EEG seizure analysis**, capturing both **spatial** (electrode relations) and **temporal** (signal evolution) dynamics **without heavy recurrent gates**.

- Traditional RNNs and LSTMs model time-varying EEG well but rely on **complex gating mechanisms**, making them slow and memory-heavy.

- REST simplifies this by using **residual state updates,** the model keeps a running internal state and adds only small, graph-based corrections at each time step.

Table 1. Comparison of seizure detection and classification methods. **A)** Capturing the non-Euclidean geometry of EEG signals. **B)** Capturing the temporal behavior of EEG signals. **C)** Evaluated for both short and long-term seizure detection. **D)** Runtime efficient **E)** Memory efficient.

Method	A	B	C	D	E
SeizureNet (Asif et al., 2020)	✗	✔	✗	✗	✗
Transformer (Yan et al., 2022a)	✗	✔	✔	✗	✗
EEG-CGS (Ho & Armanfard, 2023)	✔	✔	✗	✗	✗
GGN (Li et al., 2022)	✔	✔	✔	✗	✗
LSTM (Hochreiter & Schmidhuber, 1997)	✗	✔	✗	✗	✗
CNN-LSTM [1] (Ahmedt-Aristizabal et al., 2020)	✗	✔	✗	✗	✗
CNN-LSTM [2] (Thodoroff et al., 2016)	✗	✔	✗	✗	✗
DCRNN (Tang et al., 2021)	✔	✔	✔	✗	✗
REST (Ours)	✔	✔	✔	✔	✔

REST: Efficient and Accelerated EEG Seizure Analysis through Residual State Updates

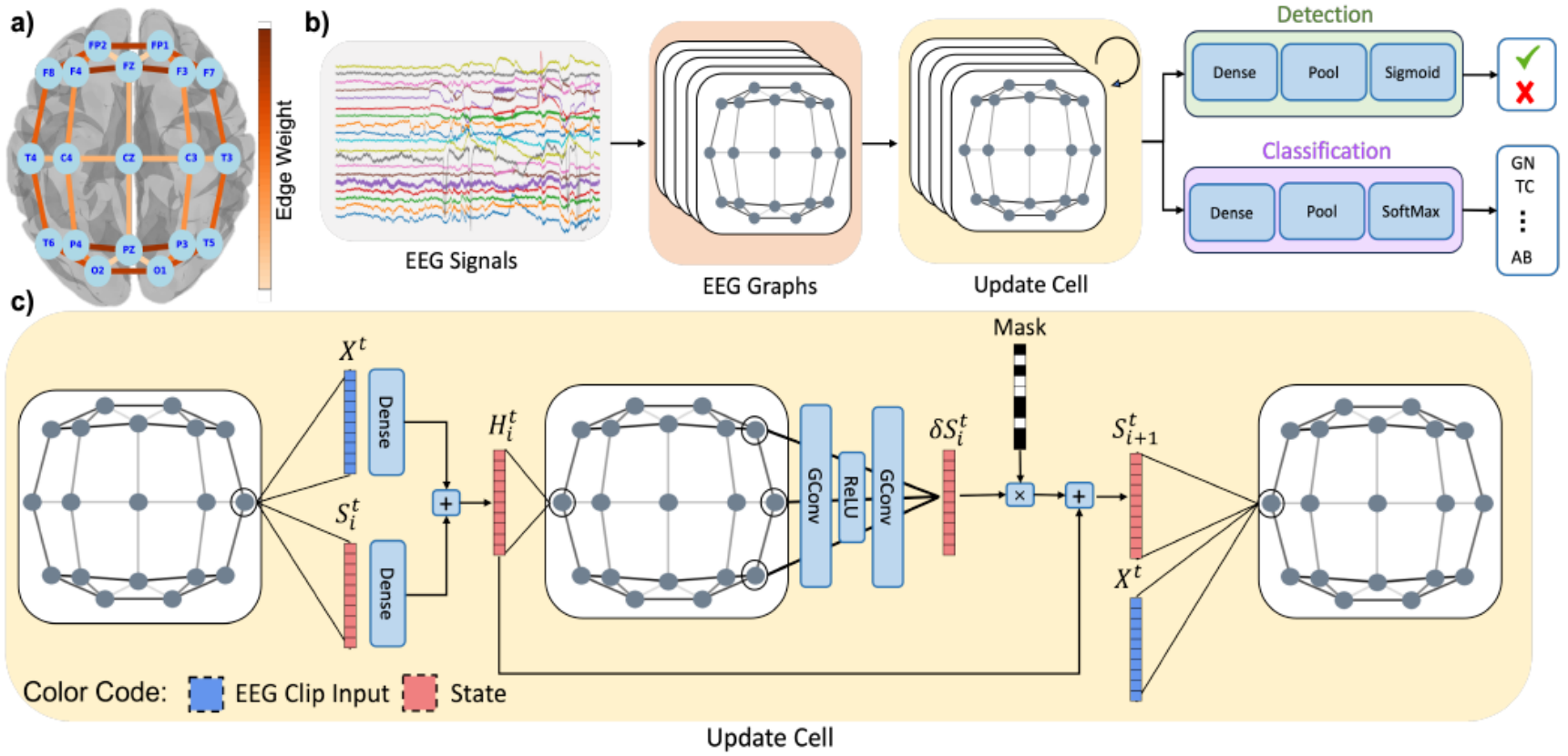


Figure 1. (**a**) EEG electrodes placement based on the 10/20 standard and its constructed distanced based EEG graph. Self edges are not shown for better visualization. (**b**) The REST framework, where raw EEG signals undergo preprocessing and are structured as a graph before feeding as input to the model. Following multiple (or single) updates, the model provides the detection or classification result. (**c**) Single update mechanism of the proposed model. Dense represents the fully connected layer and GConv is the graph convolution. See our web page for more visual results at https://arshiaafzal.github.io/REST/.

Seizure Detection & Classification Problem

Input:

From EEG clip

$$X \in \mathbb{R}^{T \times M \times N}$$

T :time points | N: electrodes | M: features per node

Output label y:

- **Detection:** binary → seizure (1) / non-seizure (0)
- **Classification:** multiclass → $y \in \{0, 1, 2, 3, 4\}$) five seizure types)

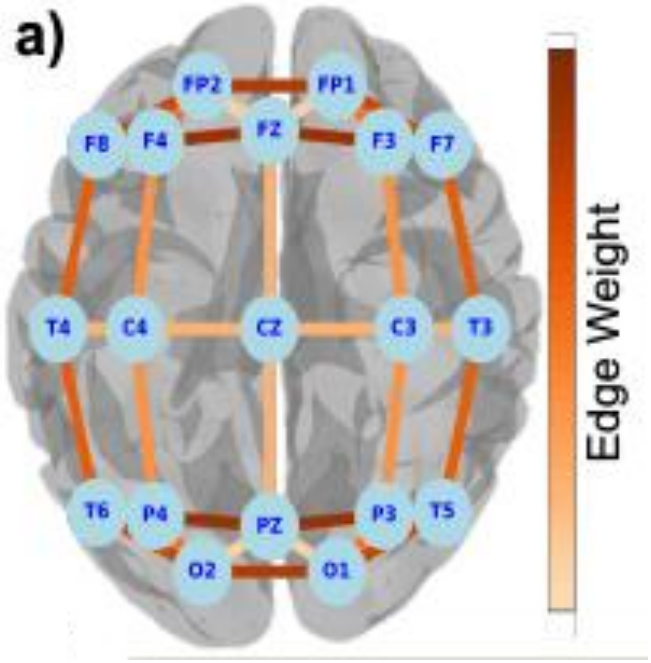
Goal:

Train model $f_\theta(X)$to detect or classify seizures from spatio-temporal EEG graphs efficiently for real-time inference.

EEG Graph Construction

EEG data → **graph** $G = (V, E, A)$

V :nodes = EEG electrodes
E: edges = spatial or functional connections
A: adjacency matrix encoding connection strengths



- Electrodes are positioned using the **standard 10–20 EEG montage**.
- **Adjacency matrix** A built from **pairwise distances** between électrodes. A is **static**
- Node features at each time step come from preprocessed EEG signals (e.g., spectral or temporal features).

REST Update Cell - Residual State Update

Model temporal evolution of EEG signals **without gated RNN units** (LSTM/GRU), keeping computation light and avoiding vanishing gradients.

1 State mapping Input EEG clip at time t: $X_t \in \mathbb{R}^{M \times N}$ Previous state: $S_{t-1} \in \mathbb{R}^{Q \times N}$

Linear mapping to latent space:

$$H_t = WX_t + US_{t-1}$$

where $W \in \mathbb{R}^{Q \times M}$,$U \in \mathbb{R}^{Q \times Q}$.

2 Residual state mechanism: Inspired by residual networks (He et al., 2016):

$$S_t = H_t + \delta S_t$$

The residual term δS_tincrementally refines H_t ,preserving past information while allowing efficient updates.

3 Spatial refinement via Graph Convolution

δS_tcaptures spatial EEG dynamics using a simple graph convolution (Morris et al., 2019):

$$O_t[:, i] = \sigma\left(\Theta_1 H_t[:, i] + \Theta_2 \sum_{j \neq i} a_{ij} H_t[:, j]\right)$$

or compactly

$$\delta S_t = G_\Theta(H_t)$$

Binary Random Mask

Standard dropout is only used during training; REST needs **speed and regularization during inference** too.

During each state update, only a random subset of feature points are activated:

$$S_t = H_t + \delta S_t \odot B$$

where

- $\odot$: element-wise (Hadamard) product
- $B \in \mathbb{R}^{Q \times N}$:**binary mask**, with

$$B_{ij} \sim Bernoulli(p)$$

$\rightarrow B_{ij} = 1$with probability p, else 0

Effect:

- Randomly skips parts of δS_t ,reducing computations.
- Prevents overfitting by forcing the model to rely on multiple pathways.
- **Accelerates inference** since zero-masked updates are ignored.
- Maintains stability under noisy or missing EEG electrodes.

Multiple Update Mechanism

- Deeper RNNs improve performance by extracting richer features.
- But stacking layers → high memory & parameter cost (extra weights, gates, and states).

REST solution:

Perform **multiple residual updates** at each time step using **shared parameters**:

$$H_t^{(i)} = WX_t + US_t^{(i)}$$
$$S_t^{(i+1)} = H_t^{(i)} + \delta S_t^{(i)} \odot B$$

where

- i: current residual update iteration
- $\delta S_t^{(i)} = G_\Theta\left(H_t^{(i)}\right)$: graph convolution update
- B: binary random mask (selective node activation)

Mechanism:

- The **same weights (W, U, Θ)** are reused for all updates.
- The **binary mask B** ensures different parts of the state are refined at each iteration.
- After *I* updates, the final state $S_t^{(I)}$ becomes the initial state for the next time step:

$$S_{t+1}^{(0)} = S_t^{(I)}$$

Benefits:

- Mimics **stacked RNN depth** without extra memory or parameters.
- Improves **temporal representation**, **stability**, and **performance**.
- Maintains **real-time efficiency.**

Experimental Setup: Data & Implementation

Datasets

- **TUSZ (Temple University Seizure Corpus):** 5545 EEG files, 5 seizure types.
- **CHB-MIT:** 24 pediatric patients, 192 seizures.
- Tasks: *Seizure detection* (binary) & *Seizure classification* (five classes).

Preprocessing

- EEG recordings converted into **graphs**:
  - **Nodes:** electrodes
  - **Edges:** spatial distances between electrodes (10–20 layout).
- Signals segmented into time clips $X \in \mathbb{R}^{T \times M \times N}$.

Implementation

- Framework: **PyTorch**
- Hardware: **NVIDIA RTX 3060 GPU**
- **Training setup:** Adam optimizer, early stopping, batch size tuned per dataset.

Experimental Setup: Data & Implementation

Table 2. Summary of TUSZ v.2.0.0 Train and Evaluation sets used in this study. Columns represent (from left to right): 1) total number of EEG files 2) total Number of patients 3) total number of generalized non-specific (GN) 4) tonic-clonic (TC) 5) absence (AB) 6) focal (FN), and 7) complex parietal (CP) seizure types in train and evaluation sets.

	EEG-Files (% Seizures)	Patients (% Seizures)	Seizure Type Numbers (Seizure Type Sessions)				
			GN	TC	AB	FN	CP
Train	4664(5.34%)	579(36%)	335(152)	30(11)	50(15)	1516(496)	279(132)
Evaluation	881(5.82%)	43(79%)	185(54)	57(8)	50(1)	240(98)	108(32)

Table 3. Summary of CHB-MIT Train and Evaluation sets used in this study.

	Patients	Seizures	Recording (hours)
Train	18	154	732
Evaluation	3	19	91
Test	3	19	92.5

Experimental Setup: Baselines

Datasets

- **TUSZ (Temple University Seizure Corpus):** 5545 EEG files, 5 seizure types.
- **CHB-MIT:** 24 pediatric patients, 192 seizures.
- Tasks: *Seizure detection* (binary) & *Seizure classification* (five classes).

Preprocessing

- EEG recordings converted into **graphs**:
  - **Nodes:** electrodes
  - **Edges:** spatial distances between electrodes (10–20 layout).
- Signals segmented into time clips $X \in \mathbb{R}^{T \times M \times N}$.

Implementation

- Framework: **PyTorch**
- Hardware: **NVIDIA RTX 3060 GPU**
- **Training setup:** Adam optimizer, early stopping, batch size tuned per dataset.

Results

Table 4. Summary of models for seizure detection on the TUSZ dataset. AUROC of different models is represented along with their memory demands and inference times.

	Seizure Detection AUROC (%)						Model Efficiency		
Model	4-s	6-s	8-s	10-s	12-s	14-s	Size(MB)	#Param	Inference(ms)
LSTM	$75.5_{\pm 0.3}$	$76.1_{\pm 0.07}$	$80.1_{\pm 0.3}$	$70.43_{\pm 0.02}$	$77.9_{\pm 0.06}$	$74.24_{\pm 0.2}$	2.147	536K	3.254
GRU	$76.1_{\pm 0.02}$	$78.8_{\pm 0.03}$	$73.2_{\pm 0.04}$	$73.5_{\pm 0.02}$	80.1 ± 0.1	$77.9_{\pm 0.04}$	1.61	402K	2.12
ResNet-LSTM	$79.1_{\pm 0.05}$	$80.1_{\pm 0.2}$	$75.6_{\pm 0.07}$	$74.3_{\pm 0.04}$	$78.8_{\pm 0.1}$	$80.0_{\pm 0.08}$	27.6	6.9M	6.78
ResNet-Dilation-LSTM	$80.2_{\pm 0.08}$	$76.5_{\pm 0.12}$	$75.9_{\pm 0.06}$	$73.6_{\pm 0.03}$	$77.4_{\pm 0.15}$	$78.2_{\pm 0.07}$	27.6	6.9M	6.78
CNN-LSTM	$81.3_{\pm 0.1}$	$78.5_{\pm 0.05}$	$76.4_{\pm 0.01}$	$75.4_{\pm 0.05}$	$75.05_{\pm 0.1}$	$74.0_{\pm 0.03}$	22.8	6M	5.624
DCRNN	$79.7_{\pm 0.01}$	$82.1_{\pm 0.04}$	$80.1_{\pm 0.04}$	$80.0_{\pm 0.06}$	$82.5_{\pm 0.1}$	$80.12_{\pm 0.04}$	0.884	126K	9.670
DCRNN w/SS	$\mathbf{83.0_{\pm 0.08}}$	$81.8_{\pm 0.05}$	$\mathbf{82.7_{\pm 0.1}}$	$82.1_{\pm 0.03}$	$85.6_{\pm 0.2}$	$84.0_{\pm 0.01}$	1.319	330K	23.25
Transformer	$83.0_{\pm 0.02}$	$82.1_{\pm 0.03}$	$82.2_{\pm 0.04}$	$\mathbf{85.5_{\pm 0.07}}$	$\mathbf{86.0_{\pm 0.03}}$	$\mathbf{85.1_{\pm 0.02}}$	0.80	120.3K	2.5
REST$_{(DS)}$	$75.3_{\pm 0.2}$	$67.0_{\pm 0.03}$	$72.2_{\pm 0.07}$	$74.1_{\pm 0.1}$	$70.6_{\pm 0.04}$	$70.0_{\pm 0.04}$	**0.037**	**8.4K**	**0.615**
REST$_{(RS)}$	$79.4_{\pm 0.03}$	$81.1_{\pm 0.01}$	$81.0_{\pm 0.08}$	$81.8_{\pm 0.02}$	$80.1_{\pm 0.1}$	$78.1_{\pm 0.4}$	**0.037**	**8.4K**	**0.710**
REST$_{(RM)}$	$82.4_{\pm 0.04}$	$\mathbf{82.2_{\pm 0.05}}$	$\mathbf{82.7_{\pm 0.1}}$	$83.6_{\pm 0.2}$	$83.4_{\pm 0.09}$	$82.0_{\pm 0.1}$	**0.037**	**8.4K**	**1.292**

Results

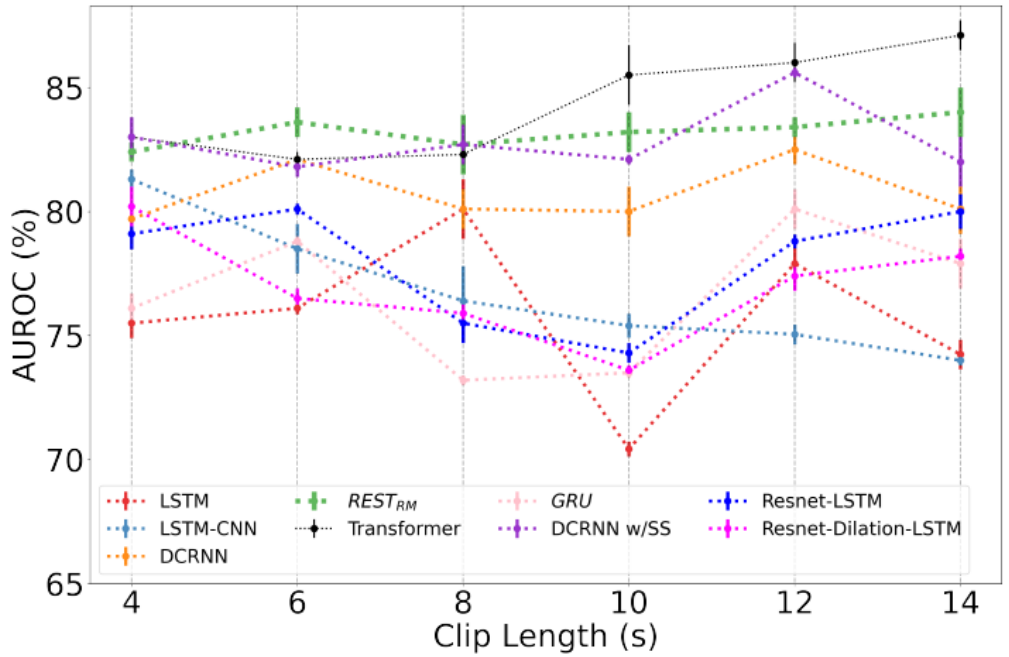


Figure 2. AUROC comparison among various models for seizure detection across different clip lengths on TUSZ dataset. A flatter line indicates more consistent performance, with error bars representing variation across five random seeds. Higher values on the y-axis correspond to increased accuracy. REST$_{(RM)}$ is shown as bold green line to emphasise its stability.

Table 6. Classification Performance, model size and parameter count for different models under the clip length of 10-s.

Model	F1-Score	Size(MB)	Parameter(#)
LSTM	0.39	2.021	512K
GRU	0.44	1.92	553K
ResNet-LSTM	0.58	30.3	7.5M
ResNet-LSTM-Dilation	0.50	30.3	7.5M
CNN-LSTM	0.47	23.9	6M
DCRNN	0.54	0.506	126K
DCRNN w/SS	**0.62**	1.40	332K
Transformer	0.54	0.25	53K
REST$_{(DS)}$	0.51	**0.034**	**8.6K**
REST$_{(RS)}$	0.57	**0.034**	**8.6K**
REST$_{(RM)}$	0.60	**0.034**	**8.6K**

Results

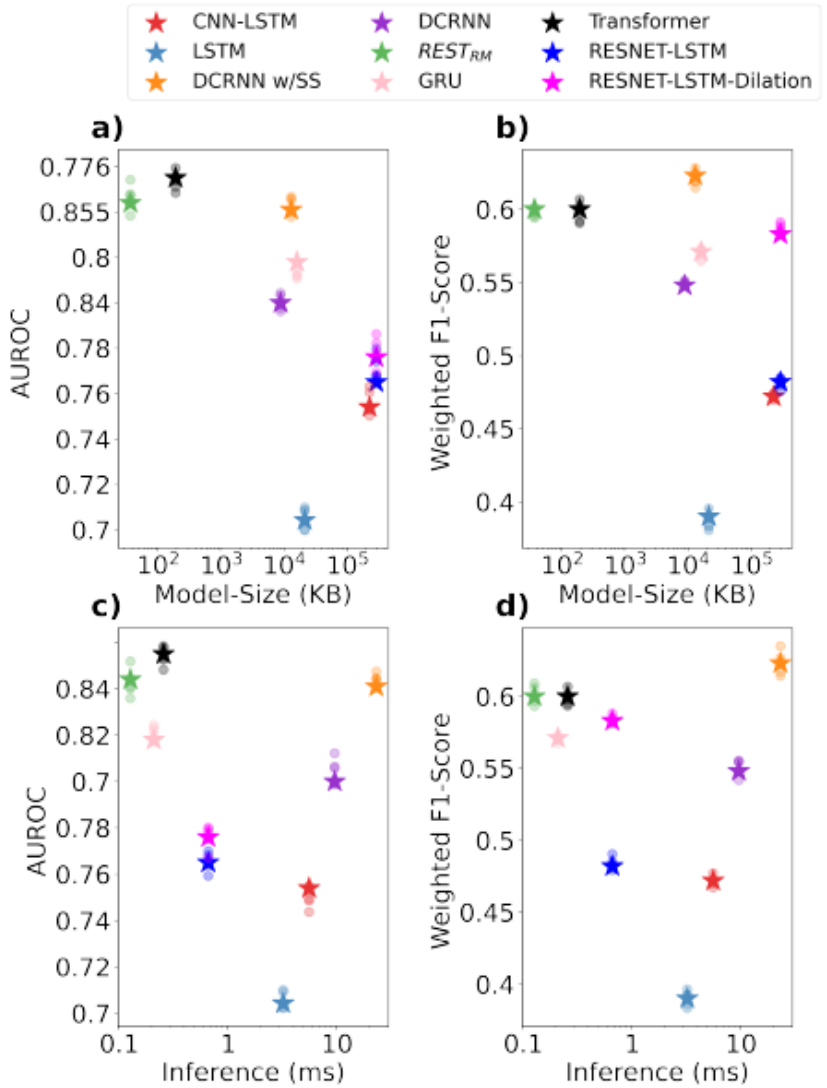


Figure 3. Performance comparison in seizure analysis across models on TUSZ dataset: **a)** Seizure detection AUROC vs. Model size. **b)** Seizure classification weighted F1-score vs. model size. **c)** Seizure detection AUROC vs. inference. **d)** Seizure classification weighted F1-score vs. inference. The •s represents the accuracy on evaluation set for different train/validation splits and ⋆s represent the mean accuracy across different train/validation splits.

Results

Table 5. Summary of models for seizure detection on the CHB-MIT dataset. AUROC of different models is represented along with their memory demands and inference times.

	Seizure Detection AUROC (%)					Model Efficiency		
Model	4-s	6-s	8-s	10-s	12-s	Size(MB)	#Param	Inference(ms)
LSTM	$85.5_{\pm 0.2}$	$84.1_{\pm 0.4}$	$81.0_{\pm 0.2}$	$75.2_{\pm 0.03}$	$73.5_{\pm 0.08}$	2.691	627K	3.56
GRU	$76.1_{\pm 0.3}$	$78.8_{\pm 0.03}$	$73.2_{\pm 0.4}$	$73.5_{\pm 0.01}$	$80.1_{\pm 0.2}$	1.92	553K	2.42
ResNet-LSTM	$77.6_{\pm 0.2}$	$82.1_{\pm 0.14}$	$79.9_{\pm 0.3}$	$76.8_{\pm 0.4}$	$81.4_{\pm 0.17}$	29.1	7.2M	6.84
ResNet-Dilation-LSTM	$78.2_{\pm 0.03}$	$79.8_{\pm 0.1}$	$82.3_{\pm 0.4}$	$77.6_{\pm 0.4}$	$81.2_{\pm 0.1}$	29.1	7.2M	6.84
CNN-LSTM	$86.2_{\pm 0.4}$	$84.9_{\pm 0.2}$	$80.4_{\pm 0.04}$	$80.35_{\pm 0.06}$	$77.6_{\pm 0.3}$	7.6M	30.23	6.432
DCRNN	$88.7_{\pm 0.3}$	$80.0_{\pm 0.02}$	$86.8_{\pm 0.06}$	$88.8_{\pm 0.3}$	$86.5_{\pm 0.3}$	0.591	147K	9.80
Transformer	$80.1_{\pm 0.2}$	$82.3_{\pm 0.6}$	$82.2_{\pm 0.04}$	$85.5_{\pm 0.01}$	$86_{\pm 0.17}$	0.25	52.4K	6.00
REST$_{(DS)}$	$89.1_{\pm 0.2}$	$88.5_{\pm 0.08}$	$90.1_{\pm 0.1}$	$86.3_{\pm 0.03}$	$87.8_{\pm 0.5}$	**0.037**	**9.3K**	**1.314**
REST$_{(RS)}$	$92.3_{\pm 0.1}$	$88.7_{\pm 0.06}$	$\mathbf{92.1_{\pm 0.03}}$	$\mathbf{93.5_{\pm 0.02}}$	$91.5_{\pm 0.02}$	**0.037**	**9.3K**	**1.314**
REST$_{(RM)}$	$\mathbf{96.7_{\pm 0.2}}$	$\mathbf{92.3_{\pm 0.04}}$	$91.4_{\pm 0.1}$	$89.2_{\pm 0.4}$	$\mathbf{91.6_{\pm 0.03}}$	**0.037**	**9.3K**	**1.314**

Conclusions

- **Accuracy**
  - **REST** outperforms all baselines on **CHB-MIT** and matches or exceeds **DCRNN/Transformer** on **TUSZ** (esp. 6–8 s clips).
  - **Multiple random updates** improve stability and consistency across clip lengths.

- **Model size & efficiency**
  - **14× smaller** than the smallest model; **38× fewer parameters** than DCRNN w/ SS, **697× fewer** than CNN-LSTM.
  - Maintains **high AUROC (83.6 %)** and strong F1-score with a compact graph-based design.

- **Speed**
  - **Fastest inference:** 1.29 ms (20× faster than DCRNN w/ SS, 3× faster than LSTM).
  - **< 2 % accuracy loss** while achieving real-time performance.
  - Only model with **AUROC > 90 %** on CHB-MIT and best results for short 4 s clips.

Summary

Strengths

- **Novel architecture:** Introduces REST, a graph-based residual state update mechanism combining GNN and RNN principles without expensive gating.
- **High efficiency:** 9× faster inference (≈ 1.29 ms), 14× smaller (≈ 37 KB) than smallest prior models.
- **Maintains accuracy:** Comparable or superior AUROC and F1 to large models.
- **Strong generalization:** Performs well on two datasets (TUSZ, CHB-MIT) across multiple clip lengths (4–14 s).
- **Hardware relevance:** Designed for real-time, low-power neurostimulation and seizure-alert devices.

Weaknesses

- **Limited scope:** Focused only on epilepsy (EEG), no tests on other disorders or modalities.
- **Evaluation gaps:** Mostly quantitative, lacks clinical validation or interpretability analysis.
- **Deployment untested:** Claims low-power suitability but no on-device benchmarks or energy results.
- **Potential over-simplification:** Fixed EEG graph and residual updates might miss dynamic spatial changes.
- **Comparative bias:** Baselines vary in preprocessing and tuning.