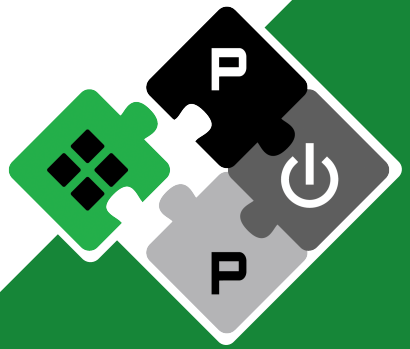
LUNA: Efficient and Topology-Agnostic Foundation Model for EEG Signal Analysis (NeurIPS 2025)

Berkay Döner
Dr. Thorir Mar Ingolfsson
Dr. Yawei Li
Prof. Luca Benini

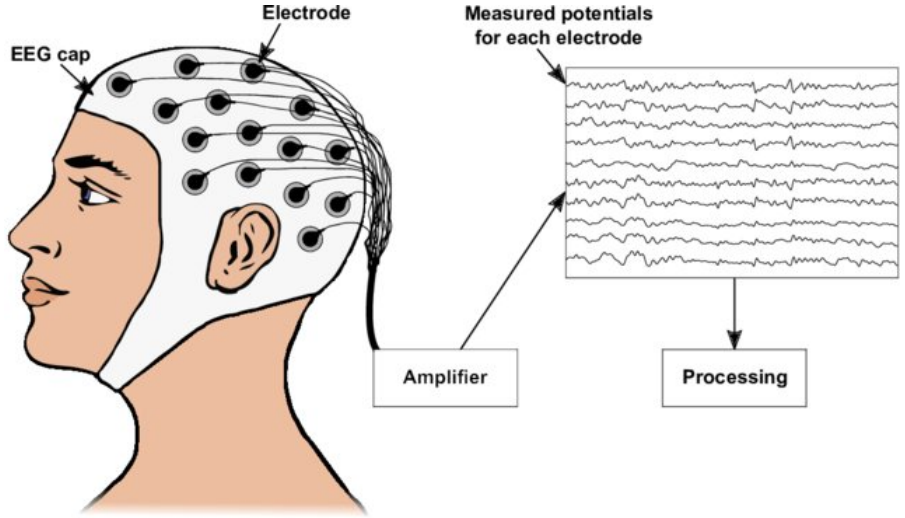


Integrated Systems Laboratory (ETH Zürich)
Embedded Systems Laboratory (EPFL)

@pulp_platform
pulp-platform.org
youtube.com/pulp_platform

Problem: EEG signals are intrinsically hard to process

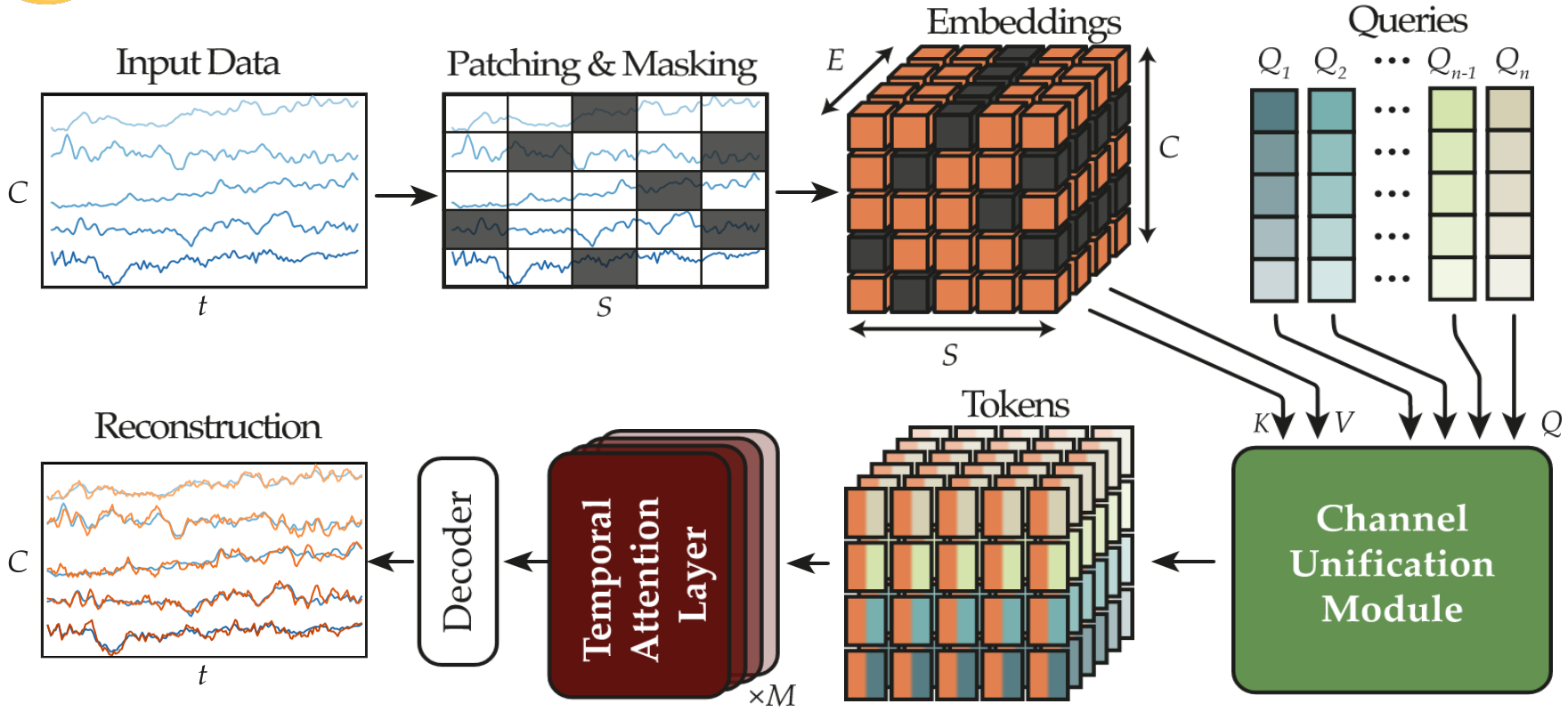
- EEG: Electroencephalography **(EEG) is** a tool to **monitor the brain activity.**



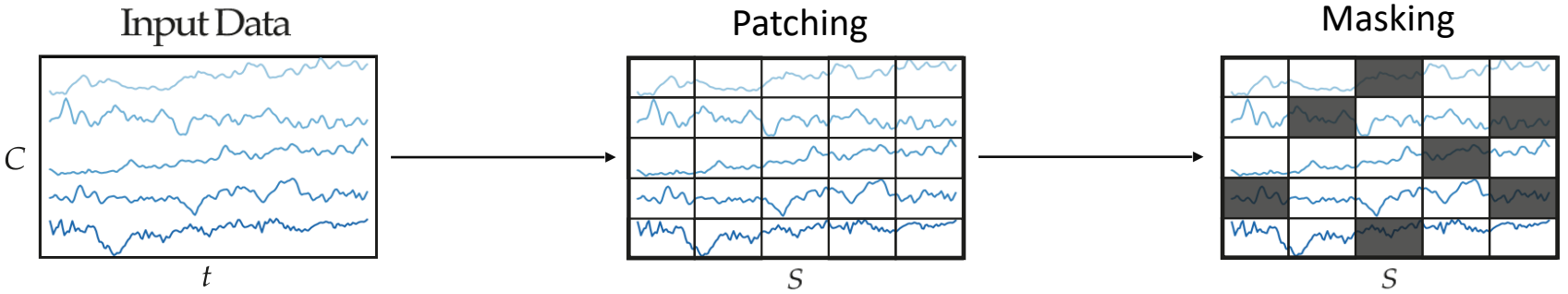
Source: Nagel, Sebastian. Towards a home-use BCI: fast asynchronous control and robust non-control state detection.

- **Challenge: Differing topologies** (number of electrodes and positions)

LUNA: Latent Unified Network Architecture



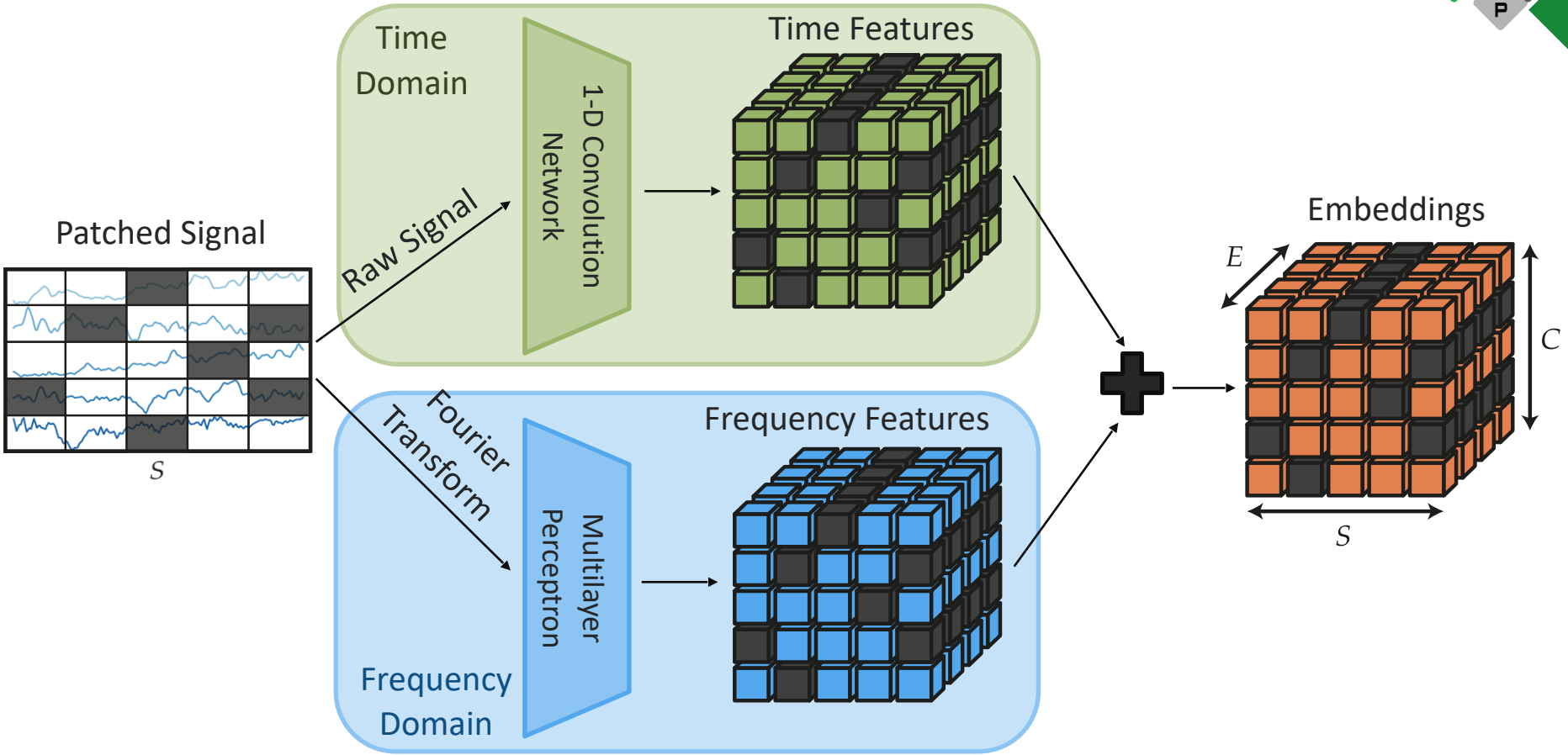
Dividing the input into small ‘tokens’



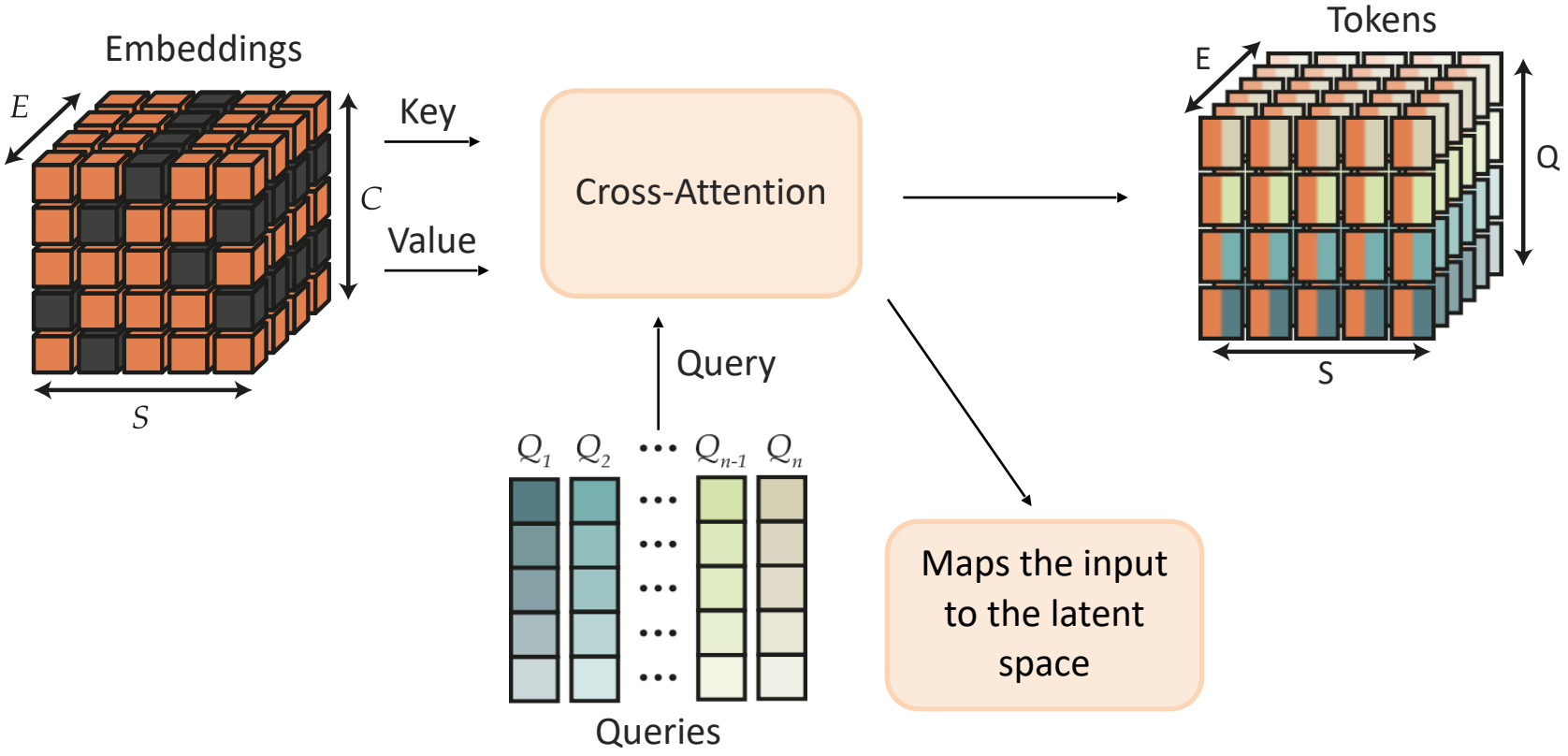
Non-overlapping patches

Masking randomly along channel and temporal axes.

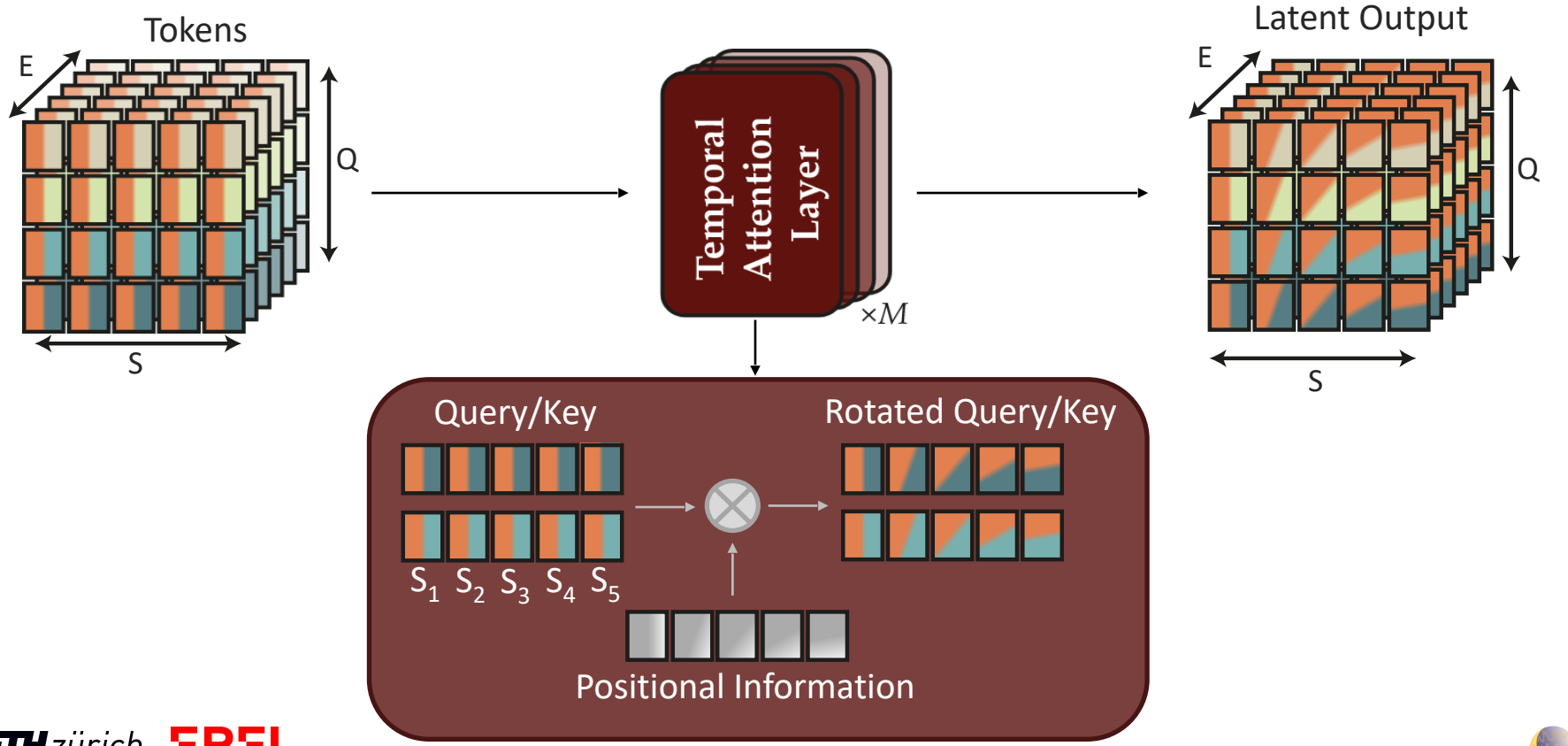
Best of both worlds: time and frequency features



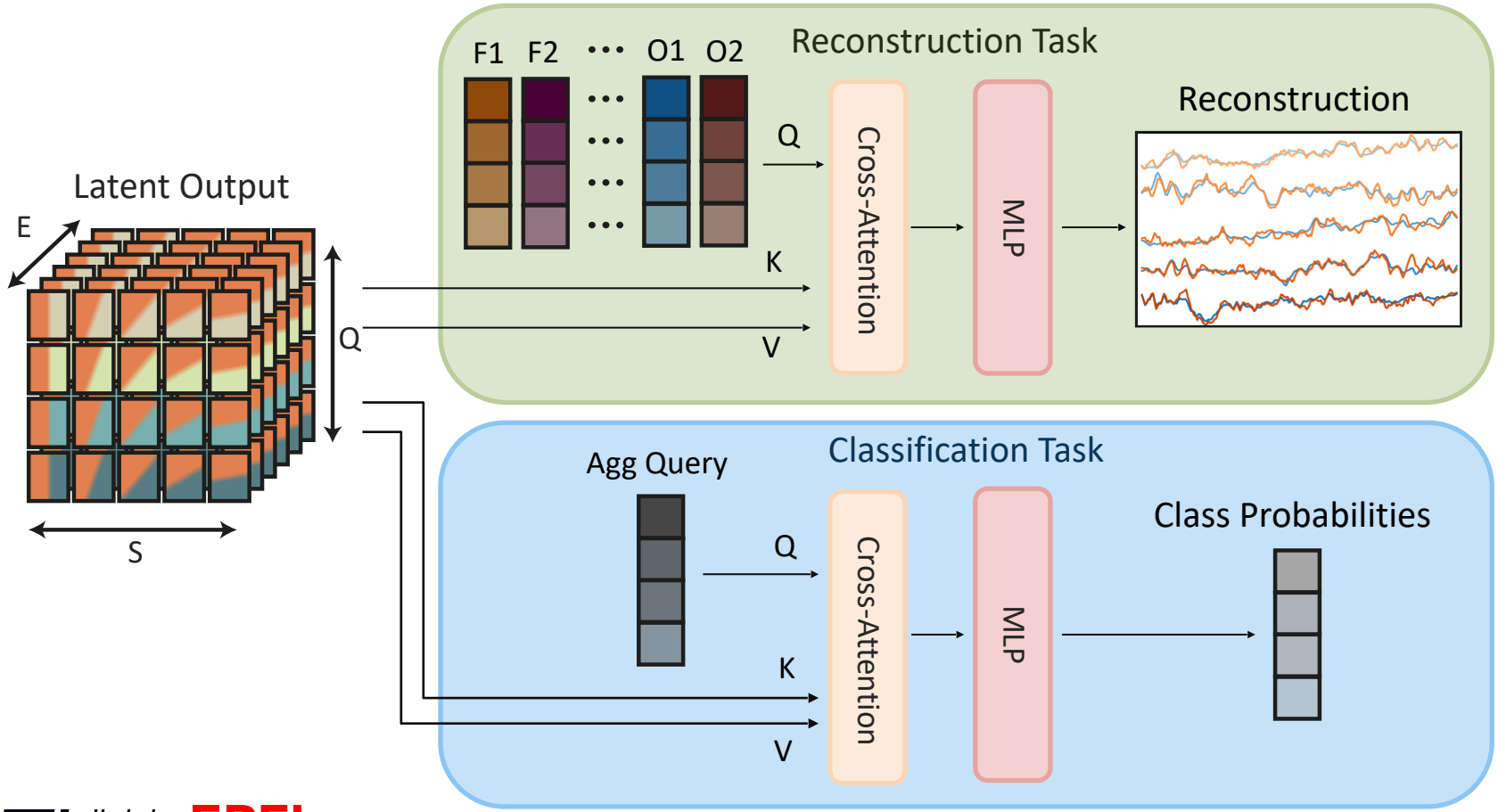
Mapping variable channels to fixed queries



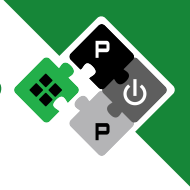
Capturing temporal relationships in the latent space



Task-adaptable decoders: one model, multiple outputs



Scaling for Success: Datasets and Model Architectures



- Pre-training datasets: TUEG and Siena → Around **22k hours** of recordings, **15k subjects** and **different topologies** with 20, 22 or 29 channels.

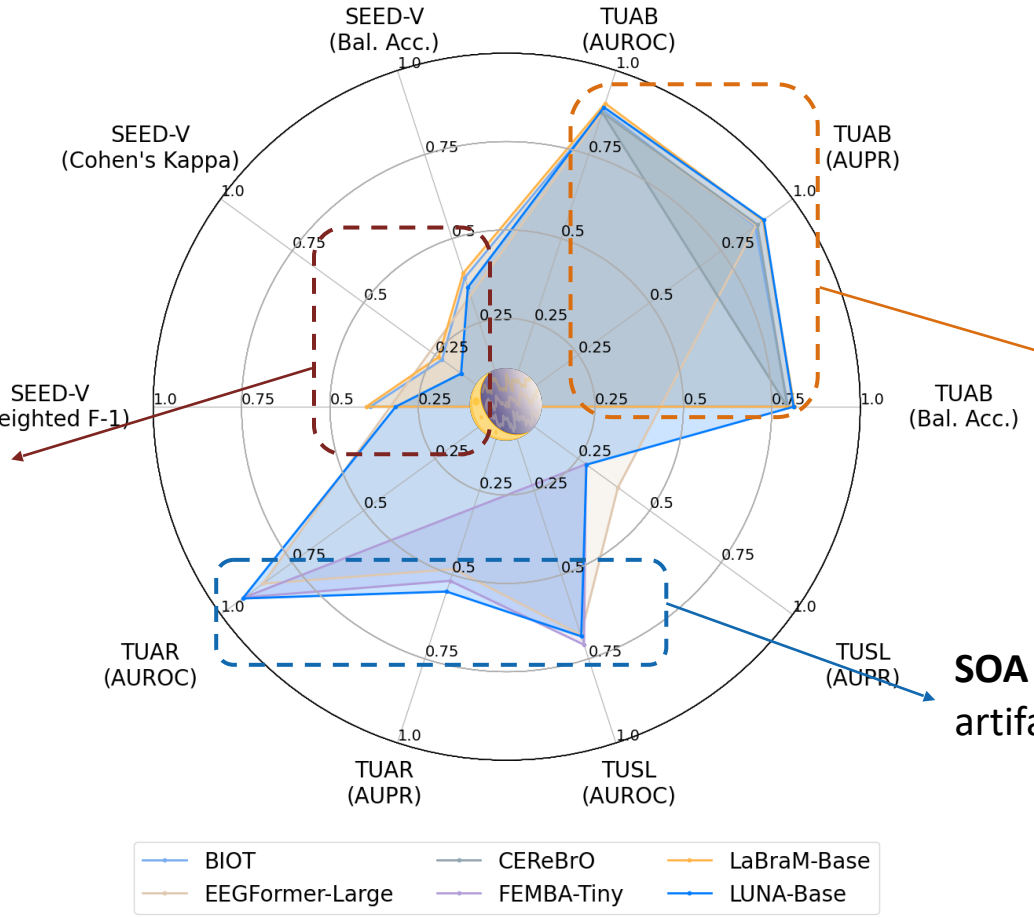
- Downstream tasks: **abnormal signal detection** (TUAB), **artifact detection** (TUAR), **slowing event detection** (TUSL) and **emotion classification** (SEED-V) → **Different topologies** with 22 or 62 channels.

Three model sizes:
Base, Large, Huge

Model Size	#Layers	# Queries	Embedding Size	# Params
Base	8	4	64	7M
Large	10	6	96	43M
Huge	24	8	128	311.4M

Base model shows SOA or comparable performance

Comparing metrics with **models with <10M parameters.**



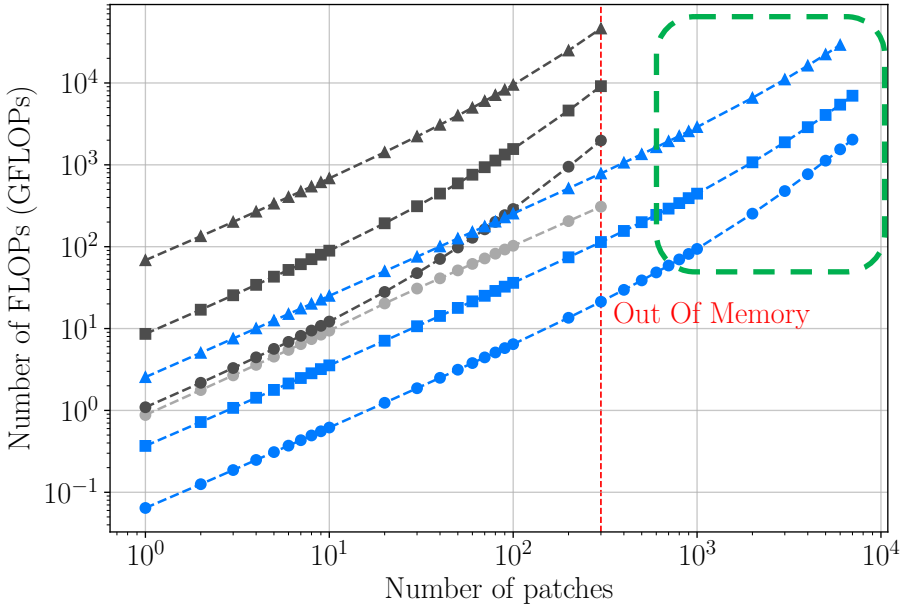
SEED-V dataset shows a limitation of the model: an **unseen** and **dense** (62 channels) topology

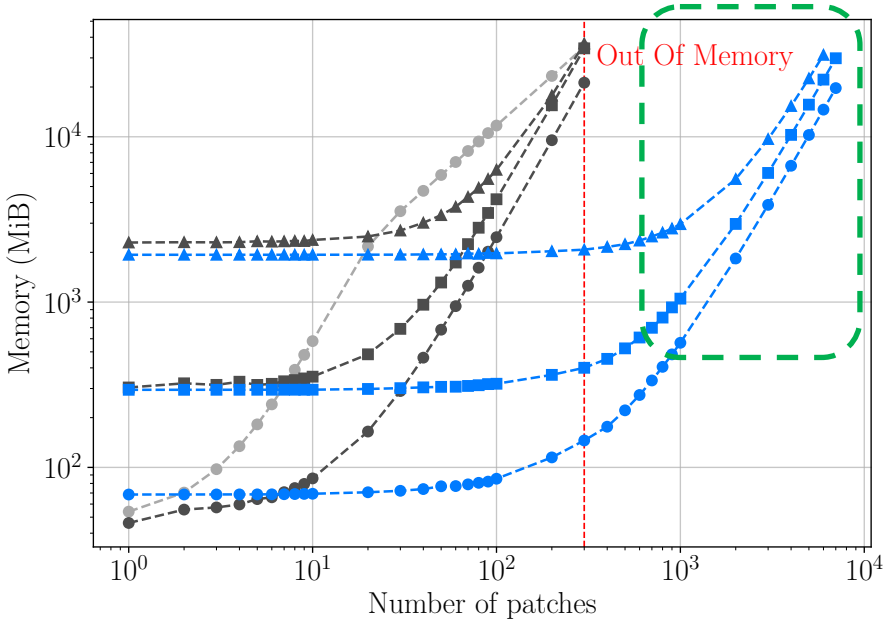
Comparable performance on abnormal signal detection

SOA performance on artifact detection

Patch scaling: efficiency with long sequences

Better scaling than other methods!

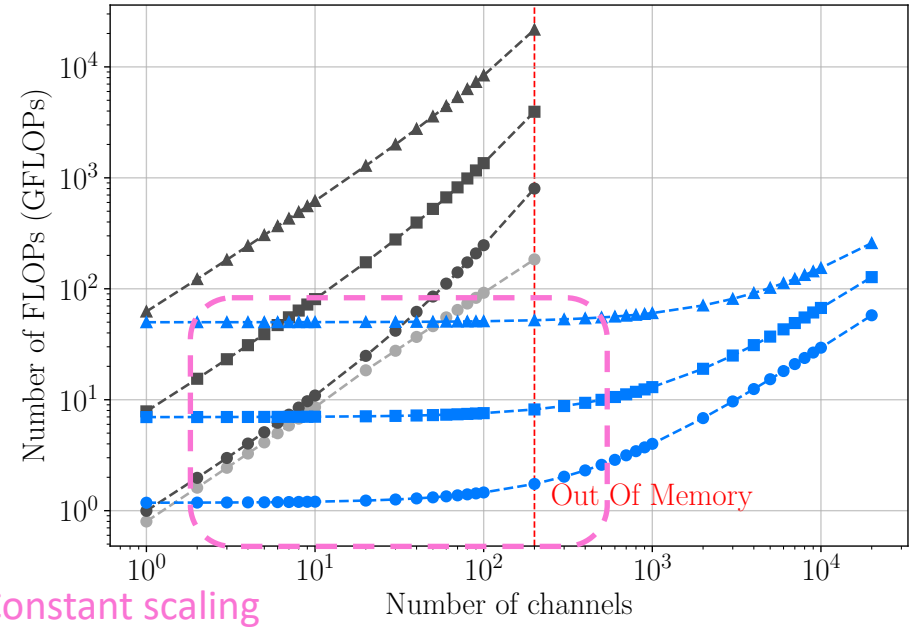


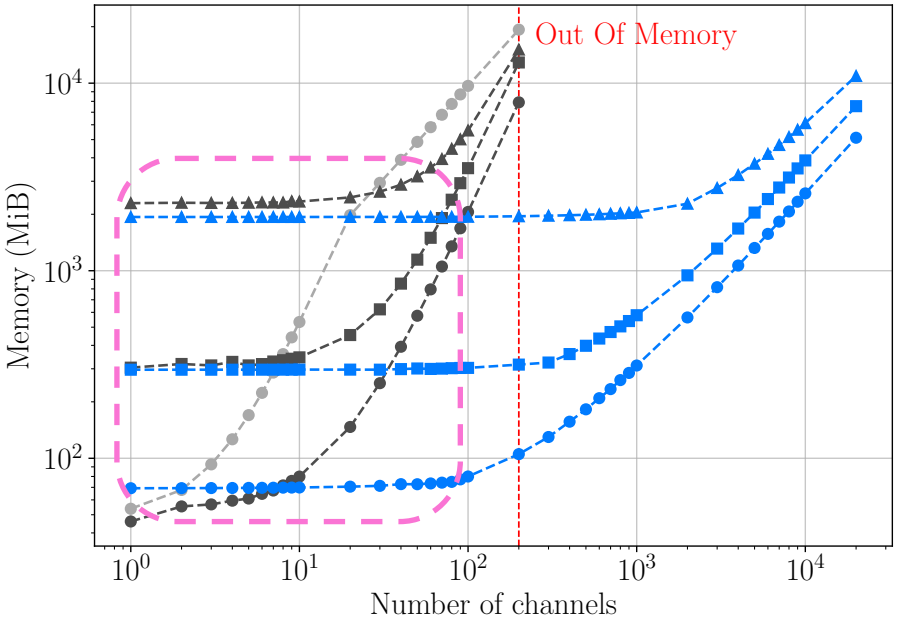


CBraMod LaBraM(Base) LaBraM(Large) LaBraM(Huge)

LUNA(Base) LUNA(Large) LUNA(Huge)

Channel scaling: near-constant, then linear



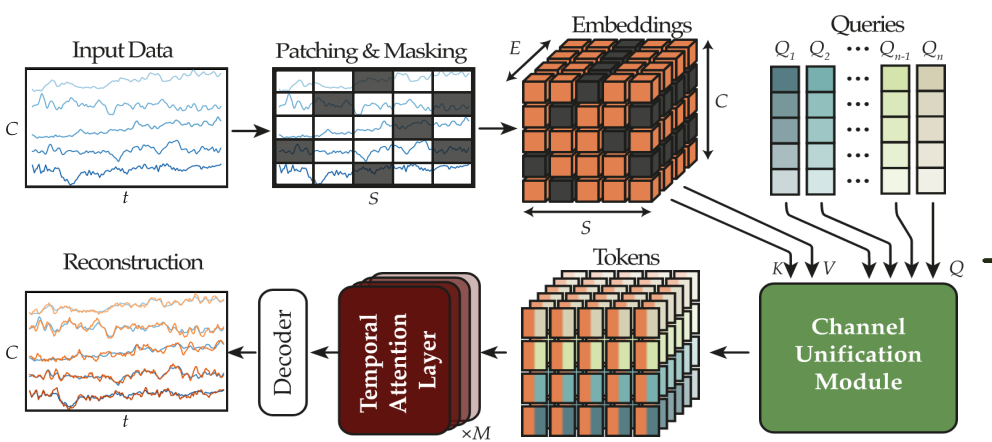


Constant scaling until cross-attention becomes the bottleneck

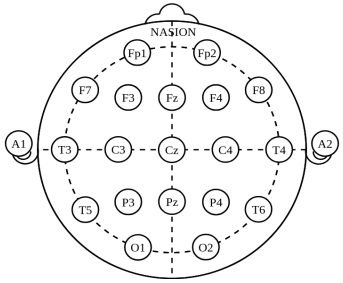
CBraMod
LaBraM(Base)
LaBraM(Large)
LaBraM(Huge)
LUNA(Base)
LUNA(Large)
LUNA(Huge)

Foundation for universal EEG analysis

LUNA



Topology-Invariant EEG Model



SOA results on artifact and slowing event detection.

Efficient analysis of EEG signals