

Figure R1. Emotion recognition and psychiatric state classification performance of supervised and self-supervised learning models. RCT-MNM consistently outperformed the other models across datasets and evaluation settings.

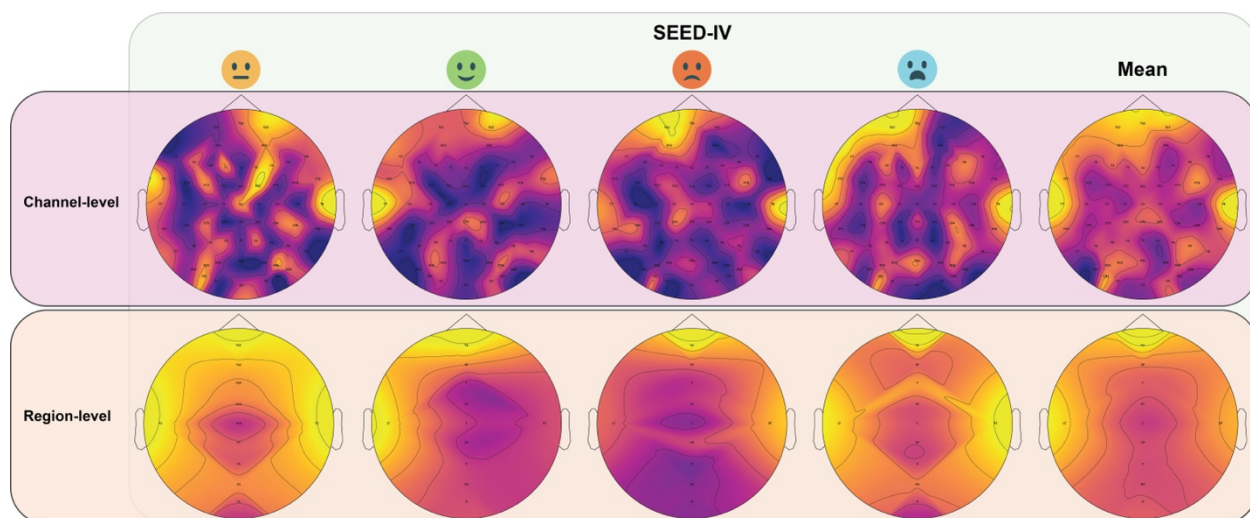


Figure R2. Region- and Channel-Level Class Activation Maps for SEED-IV, showing channel-level fine-grained activations and region-level aggregated representations for a holistic view of emotional processing.

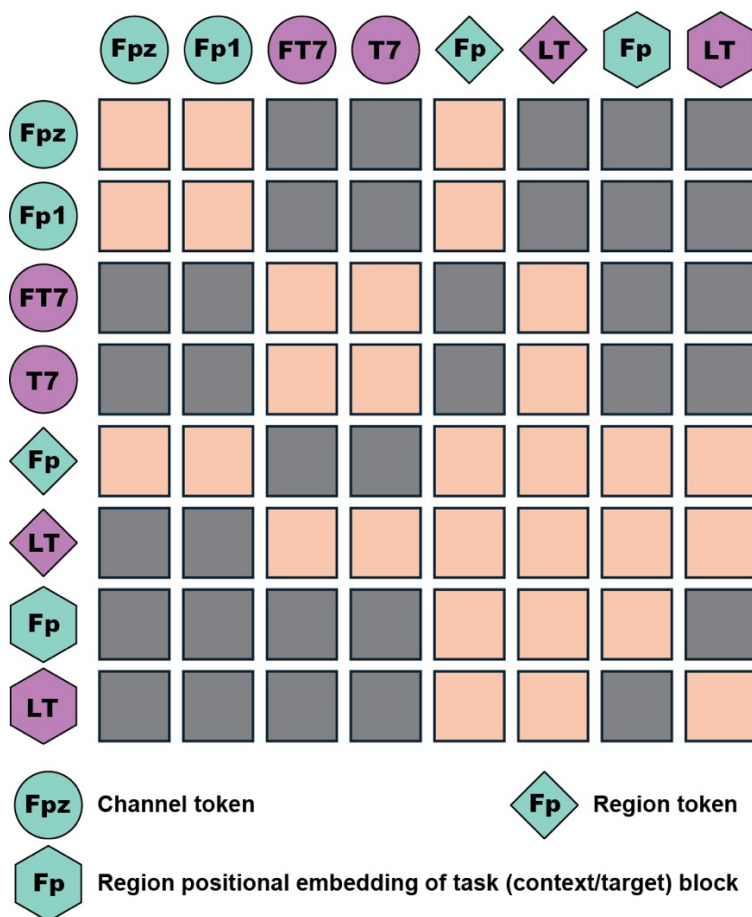


Figure R3. Attention mask in Region-Channel Self-Attention with Region Conditioning. Orange means two tokens attended with each other.