Table R1: Ablation for major components of RCT-MNM. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
RCT-MNM	**89.34±4.52**	**89.05±4.23**	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
w/o Region	85.32±4.72	85.16±4.79	78.98±6.14	74.62±6.19	71.73±7.14	69.34±6.01
w/o Spatial	82.04±4.97	82.42±4.28	76.46±6.26	72.10±6.17	69.37±6.28	66.56±7.09
w/o Temporal	85.82±4.46	85.56±4.74	79.11±6.19	75.11±6.46	72.34±6.34	70.27±6.75
w/o AALR	87.71±5.21	87.20±4.76	81.11±6.46	76.81±6.01	73.25±6.90	72.84±6.33
w/o SSL	83.34±5.34	84.06±5.10	78.44±6.78	72.97±7.07	72.96±7.44	68.54±7.26

Table R2: Ablation for removing 4 most (key) and 4 least (minor) important spatial and temporal token. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
RCT-MNM	**89.34±4.52**	**89.05±4.23**	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
w/o key spatial	86.42±5.18	86.37±4.02	79.74±6.13	75.69±6.47	72.66±7.49	70.89±6.89
w/o minor spatial	89.02±4.73	88.82±4.67	81.78±5.91	77.89±6.12	74.87±7.28	73.27±6.01
w/o key temporal	86.44±4.83	86.61±4.49	79.61±6.47	75.07±6.83	72.41±7.18	70.89±6.47
w/o minor temporal	88.92±4.56	88.65±4.87	81.94±6.13	78.11±6.28	74.72±7.33	72.91±6.42

Table R3: Classification performance on MSIT, ECR, SEED and SEED-IV with standard deviations. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded, and the second-best performance is underlined.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
SVM	80.87±8.96	80.10±8.28				
DGCNN	81.72±6.28	82.69±7.74	76.19±10.29	71.10±10.27	68.17±10.46	67.74±9.97
RGNN	82.39±7.17	81.76±6.63	78.32±6.16	73.12±8.84	69.24±8.31	68.38±10.84
TSception			76.02±8.91	74.47±9.04	65.94±10.04	70.08±10.18
EEG Conformer			76.64±10.62	73.94±8.42	72.46±6.92	68.56±9.22
LGGNet	83.89±6.48	84.13±7.79	78.92±9.37	76.86±6.74	73.22±7.57	<u>72.06±8.47</u>
MMM			78.61±10.68	<u>77.95±8.26</u>	73.21±10.74	71.80±7.29
EPNNE			77.92±9.73	76.31±10.01	73.46±7.32	71.17±8.14
ViT	82.88±6.23	83.91±5.19	78.32±8.93	73.06±7.02	72.83±7.57	67.17±8.23
MAE	86.22±4.82	<u>86.67±4.84</u>	<u>81.78±6.99</u>	75.85±6.46	74.48±6.69	70.15±7.12
I-JEPA	<u>87.17±5.14</u>	86.59±4.76	81.30±5.76	77.84±6.25	**75.04±7.29**	71.47±6.74
RCT-MNM	**89.34±4.52**	**89.05±4.23**	**82.10±5.94**	**78.03±6.02**	<u>74.81±6.76</u>	**73.44±6.26**

Table R4: Classification performance on DEAP with standard deviations. DEAP: F1 Score (%). V: Valence; A: Arousal. SD: Subject Dependent; SI: Suject Independent. The best performance is bolded, and the second-best performance is underlined.

Method	DEAP-V:SD	DEAP-V:SI	DEAP-A:SD	DEAP-A:SI
DGCNN	68.08±8.32	60.15±8.48	60.13-8.48	62.92±8.89
RGNN	64.70±9.21	62.28±7.19	58.74±9.22	66.11±8.47
TSception	63.86±7.36	63.86±7.82	65.47±8.84	**68.08±8.24**
EEG Conformer	64.09±7.58	61.76±7.37	62.60±9.29	57.84±9.11
LGGNet	67.14±7.94	64.54±7.03	66.69±8.94	67.44±8.78
MMM	64.96±8.37	65.58±6.82	<u>68.97±7.84</u>	65.64±8.36
EPNNE	65.06±8.01	65.00±6.36	66.17±8.28	63.88±8.59
ViT	65.83±7.27	64.67±7.46	64.42±7.74	63.52±8.29
MAE	66.01±7.04	65.30±6.14	67.16±7.93	65.27±7.85
I-JEPA	<u>68.19±6.72</u>	<u>67.56±6.28</u>	68.29±7.37	66.47±8.18
RCT-MNM	**71.78±6.61**	**68.15±6.19**	**70.60±7.56**	<u>67.94±8.04</u>

Table R5: Subject-dependent classfication perforance on SEED and SEED-IV using 9/6 and 16/8 split. SEED/SEED-IV: F1 Score (%).

Method	SEED:SD	SEED-IV:SD
RCT-MNM	81.29±6.48	73.94±6.92

Table R6: Classification performance on SEED and SEED-IV compared with LaBraM and CBraMod. SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
RCT-MNM	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
LaBraM-Base	80.38±7.27	75.92±7.27	74.78±7.93	71.19±6.49
CBraMod	81.11±6.04	76.73±6.31	74.56±6.24	72.02±7.02

Table R7: Classification performance on MSIT and ECR compared with seegnificant and Du-IN. MSIT/ECR: AUROC (%). The best performance is bolded.

Method	MSIT	ECR
RCT-MNM	**89.34±4.52**	**89.05±4.23**
seegnificant	84.47±5.94	85.02±4.86
Du-IN	87.24±5.43	86.17±5.39

Table R8: Ablation for various masking stragety. RC-Mask: removing cross-time and cross-time-channel targets; RT-Mask: removing cross-channel, cross-time-channel, and cross-time-region targets; CT-Mask: removing cross-region and cross-time-region targets. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
RCT-Mask	**89.34±4.52**	**89.05±4.23**	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
RC-Mask	89.02±4.72	88.95±4.79	81.98±6.14	77.78±6.19	74.78±7.14	72.95±6.01
RT-Mask	88.55±4.97	88.62±4.28	81.53±6.26	77.24±6.17	73.62±6.28	72.33±7.09
CT-Mask	88.32±4.46	88.02±4.74	80.98±6.19	76.89±6.46	72.94±6.34	72.14±6.75
Block-Mask	86.84±5.21	86.12±4.76	79.38±6.46	75.47±6.01	72.23±6.90	70.73±6.33
Random-Mask	87.71±5.34	87.20±5.10	81.11±6.78	76.81±7.07	72.75±7.44	71.36±7.26

Table R9: Ablation for region and channel tokens in RCSA. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI

RC tokens	**89.34±4.52**	**89.05±4.23**	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
R tokens (MMM-like)	87.16±5.22	87.02±4.88	81.48±6.44	76.84±6.69	74.27±7.32	70.98±7.63
C tokens	87.24±4.96	86.72±4.91	81.80.±6.76	77.01±6.83	74.52±7.04	71.16±6.77

Table R10: Ablation for Region Conditioning (RC). MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEED-IV:SI
w/ RC	**89.34±4.52**	**89.05±4.23**	82.10±5.94	**78.03±6.02**	**74.81±6.76**	**73.44±6.26**
w/o RC	88.37±4.84	87.77±4.57	**82.22.±6.08**	77.21±6.74	73.62±7.36	72.08±6.47

Table R11: Ablation for Adaptive Attention-Based Layer Refinement (AALR). LR: Layer Refinement. MSIT/ECR: AUROC(%); SEED/SEED-IV: F1 Score (%). SD: Subject Dependent; SI: Suject Independent. The best performance is bolded, and the second-best performance is underlined.

Method	Blocks	MSIT	ECR	SEED:SD	SEED:SI	SEED-IV:SD	SEE
AALR (Our)	2	**89.34±4.52**	<u>89.05±4.23</u>	**82.10±5.94**	**78.03±6.02**	**74.81±6.76**	<u>73.4</u>
AALR	3	<u>89.28±4.77</u>	**89.37±4.84**	<u>82.01.±6.21</u>	<u>77.76±6.46</u>	<u>74.56±7.19</u>	**73.6**
LR	2	88.46±4.82	88.15±4.87	81.24±6.36	77.33±6.08**	73.97±7.28**	72.5
LR	3	88.49±4.84	88.42±4.48	80.87.±6.01	77.26±6.39	73.82±6.92	72.1
w/o Refinement	1	87.71±4.72	87.20±4.92	81.11±5.77	76.81±6.20	73.24±6.41	71.8

Table R12: Model configurations

Parameter	Encoder	Predictor	Decoder
Temporal/spatial layers	4/4	1/1	1/1
Hidden dimension	64	64	64
Heads	4	4	4
Feed-forward dimension	256	256	256

Table R13: Hyperparameters for pre-training

Hyperparameter	Setting
Epochs	200

Warmup epochs	40
Batch size	256
Dropout	0.3
Optimizer	AdamW, (\(\beta_1\), \(\beta_2\))=(0.9, 0.999)
Scheduler	Cosine Annealing Scheduler
Learning rate	5e-4
Minimal learning rate	1e-6
Weight decay	5e-2
EMA momentum schedule	Linear Scheduler
EMA start momentum	0.9
EMA final momentum	1.0
Context Mask ratio	0.75
Target Mask ratio	0.9/0.9/0.9/0.9/0.9
Patch size	0.5 second

Table R14: Hyperparameters for fine-tuning

Hyperparameter	**Setting**
Epochs	50
Batch size	256
Dropout	0.3
Optimizer	AdamW, (\(\beta_1\), \(\beta_2\))=(0.9, 0.999)
Scheduler	Cosine Annealing Scheduler
Learning rate	5e-4
Minimal learning rate	1e-6
Weight decay	5e-2
Label smooothing	0.1

Table R15: Cross-dataset classification performane on SEED vs. SEED-IV (F1 Score (%)). SEED->SEED-IV: pre-training on SEED and fine-tuning on SEED-IV. SD: Subject Dependent; SI: Suject Independent. The best performance is bolded.

Method	**SEED->SEED-IV:SD**	**SEED-IV->SEED:SI**	**SEED-IV->SEED:SD**	**SEED-IV->SEED:SI**
MMM	76.84±8.74	74.57±8.83	71.92.98±9.53	69.48±9.69
MAE	79.62±7.38	75.04±7.17	73.28.53±8.17	68.74±9.34

I-JEPA	80.07±7.26	75.82.02±6.60	73.74±7.26	69.52±8.87
RCT-MNM	**80.63±6.74**	**76.72±6.82**	**73.86±6.80**	**71.47±7.28**