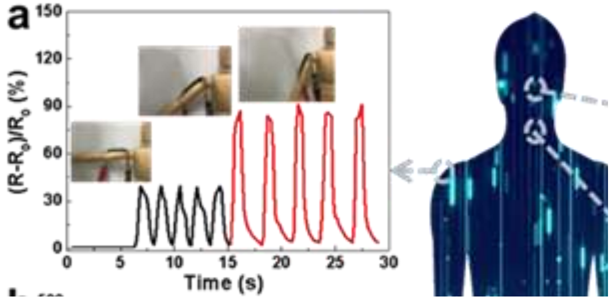
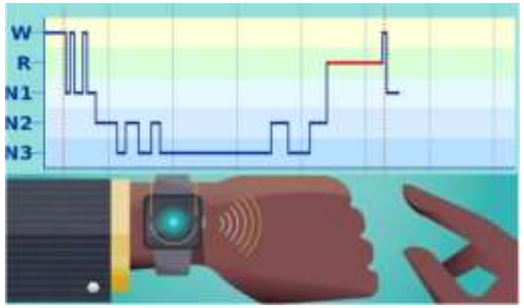
Robust Sleep Staging over Incomplete Multimodal Physiological Signals via Contrastive Imagination

Qi Shen, Junchang Xin, Bingtian Dai, Shudi Zhang , Zhiqiong Wang *

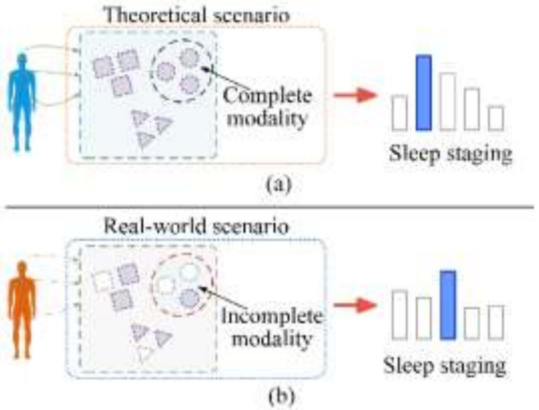
1 Background



Various sensor combinations available



Widespread application of emerging wearable devices



Uncertainty of various factors

Challenges

- **Incomplete multimodal signals**
- **Temporal context modeling**

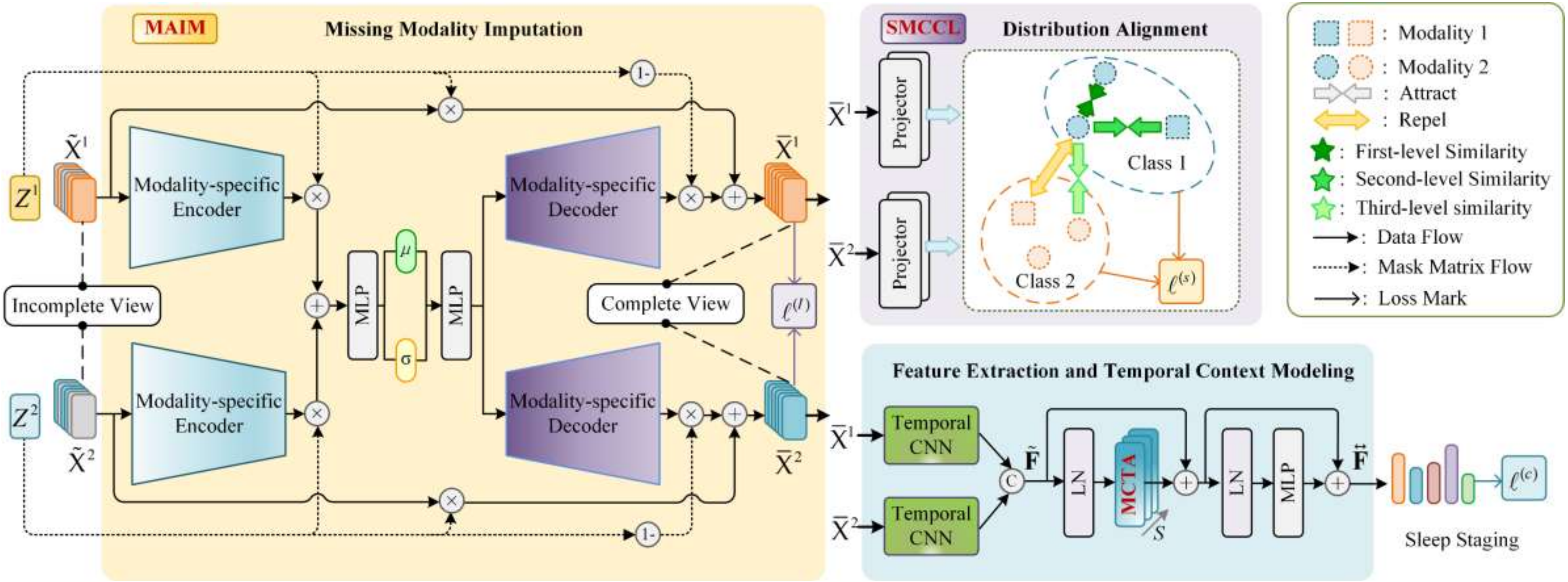
Contributions

- **Modal awareness imagination module (MAIM)**
- **Semantic & modal calibration contrastive learning (SMCCL)**
- **Multi-level cross-branch temporal attention mechanism (MCTA)**

2 Method

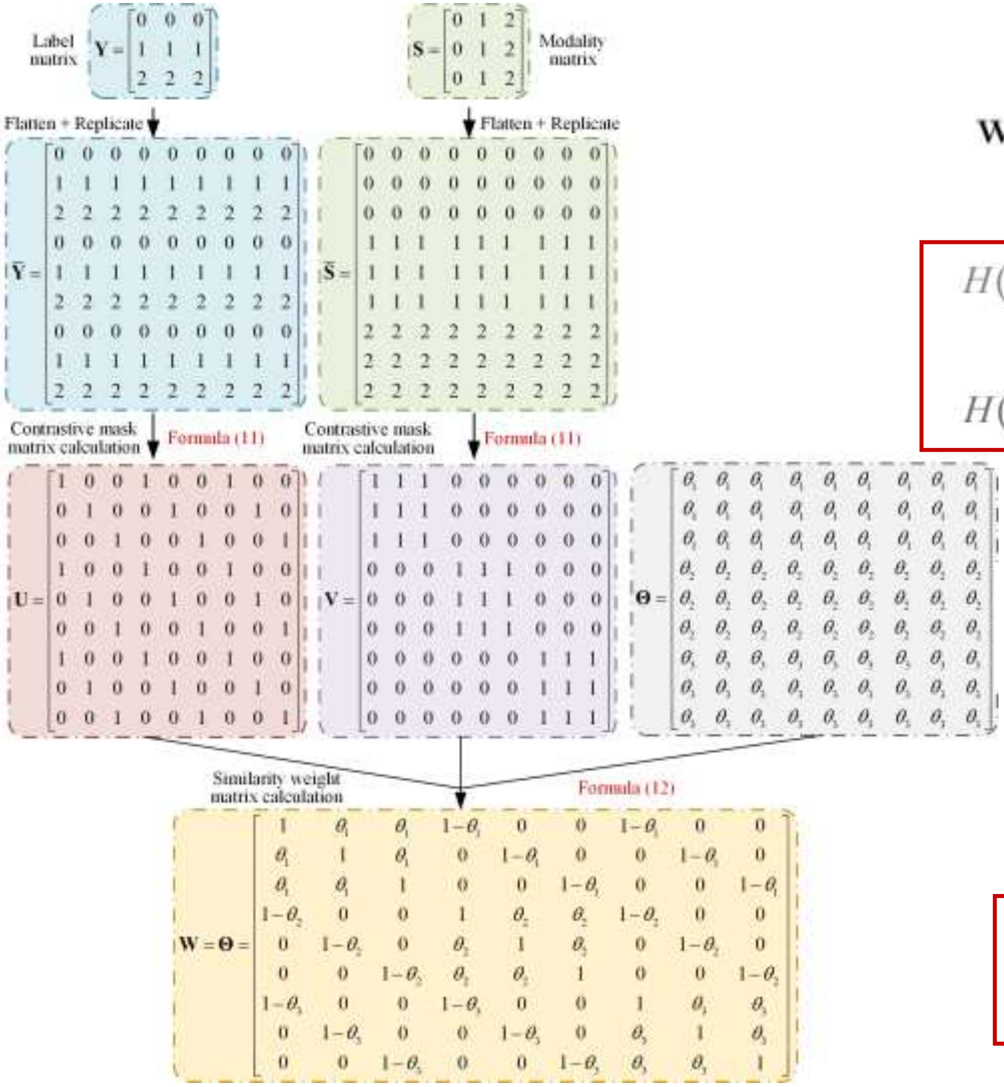
■ Overall framework

Missing rate: $\rho = 1 - \frac{1}{N \cdot M} \sum_{i=1}^{N} \sum_{j=1}^{M} Z_i^j$ At least one mode is kept in one instance



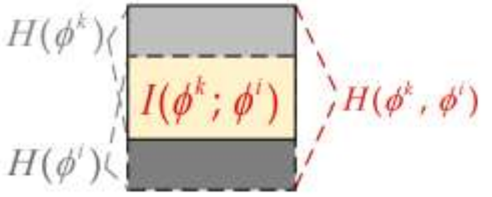
2 Method

■ SMCCL



$$\mathbf{U}=\{\{u_i^j\}_{i=1}^R\}_{j=1}^R,\ u_i^j=\begin{cases}1, & \bar{y}_i^j=\dot{y}_i^j\\ 0, & \bar{y}_i^j\neq\dot{y}_i^j\end{cases}\qquad \mathbf{V}=\{\{v_i^j\}_{i=1}^R\}_{j=1}^R,\ v_i^j=\begin{cases}1, & \bar{s}_i^j=\dot{s}_i^j\\ 0, & \bar{s}_i^j\neq\dot{s}_i^j\end{cases}$$

$$\mathbf{W}=\underbrace{\mathbf{U}\odot\mathbf{V}}_{\text{the 1th level}}+\underbrace{(1\text{-}\mathbf{\Theta})(\mathbf{U}-\mathbf{U}\odot\mathbf{V})}_{\text{the 2th level}}+\underbrace{\mathbf{\Theta}(\mathbf{V}-\mathbf{U}\odot\mathbf{V})}_{\text{the 3th level}}\qquad \theta_k=\frac{1}{M-1}\sum_{i=1}^{M}\mathbb{1}_{i\neq k}\cdot\frac{I(\phi^k;\phi^i)}{H(\phi^k,\phi^i)}$$



$$\frac{I(\phi^k;\phi^i)}{H(\phi^k,\phi^i)}=\frac{H(\phi^k)+H(\phi^i)-H(\phi^k,\phi^i)}{H(\phi^k,\phi^i)}$$
$$=\frac{\int p(x)\ln\frac{1}{p(x)}dx+\int p(y)\ln\frac{1}{p(y)}dy-\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}{\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}$$
$$=\frac{\iint p(x,y)\ln\frac{1}{p(x)}dxdy+\iint p(x,y)\ln\frac{1}{p(y)}dxdy-\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}{\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}$$
$$=\frac{\iint p(x,y)\ln\frac{1}{p(x)p(y)}dxdy-\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}{\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}$$
$$=\frac{\iint p(x,y)\ln\frac{p(x,y)}{p(x)p(y)}dxdy}{\iint p(x,y)\ln\frac{1}{p(x,y)}dxdy}$$
$$=\iint\log_{\frac{1}{p(x,y)}}\left(\frac{p(x,y)}{p(x)p(y)}\right)dxdy$$

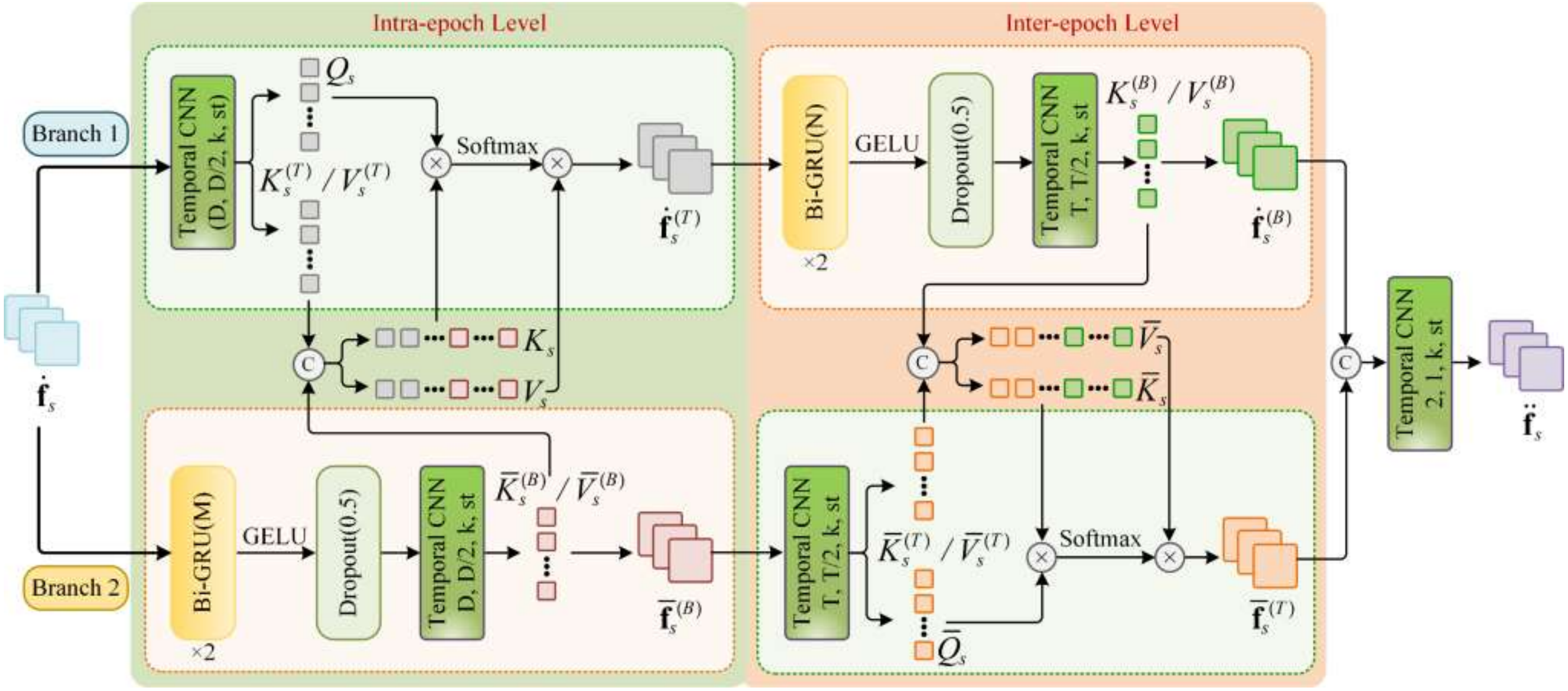
$$\frac{I(\phi^k;\phi^i)}{H(\phi^k,\phi^i)}=\log_{\frac{1}{\mathbf{P}}}\left(\frac{\mathbf{P}}{\mathbf{P}_m\mathbf{P}_n}\right)$$

$$\ell^{(s)}=\frac{-1}{N_{w_i^j>0}-1}\sum_{i=1}^{B\cdot M}\sum_{j=1}^{B\cdot M}\mathbb{1}_{i\neq j}\cdot\mathbb{1}_{w_i^j>0}\cdot w_i^j\cdot\log\frac{\exp(\varphi_i\cdot\varphi_j/\tau)}{\sum_{k=1}^{B\cdot M}\mathbb{1}_{i\neq k}\cdot\exp(\varphi_i\cdot\varphi_k/\tau)}$$

2 Method

■ MCTA

(N*L, C, D)→ (N, L, C*D/S)



3 Experiment

Quantitative results

Table 1: Performance comparison for complete and incomplete modalities in randomly partially missing case. Here "incomplete" means the maximum missing rate.

Datasets	Methods	Complete			Incomplete		
		Acc	MF1	K	Acc	MF1	K
Sleep-EDF-20	FeatConcat	0.825	0.761	0.771	0.497	0.429	0.285
	MultitaskCNN [8]	0.835	0.753	0.775	0.589	0.506	0.449
	SalientSleepNet [23]	**0.872**	**0.827**	**0.827**	0.634	0.565	0.485
	MM-Net [11]	0.867	0.817	0.822	0.570	0.493	0.432
	TransSleep [16]	0.864	0.819	0.821	0.594	0.521	0.457
	XSleepNet [10]	0.864	0.809	0.819	0.623	0.560	0.478
	CIMSleepNet	0.867	0.821	0.824	**0.853**	**0.801**	**0.805**
Sleep-EDF-78	FeatConcat	0.788	0.726	0.717	0.526	0.471	0.392
	MultitaskCNN [8]	0.795	0.727	0.722	0.613	0.535	0.453
	SalientSleepNet [23]	0.843	0.794	0.791	0.722	0.643	0.625
	MM-Net [11]	0.845	0.796	0.794	0.706	0.628	0.597
	TransSleep [16]	0.846	0.797	0.795	0.738	0.654	0.637
	XSleepNet [10]	0.838	0.776	0.779	0.697	0.622	0.583
	CIMSleepNet	**0.849**	**0.799**	**0.797**	**0.830**	**0.772**	**0.775**
SVUH-UCD	FeatConcat	0.745	0.731	0.672	0.502	0.445	0.336
	MultitaskCNN [8]	0.774	0.763	0.705	0.643	0.630	0.533
	TransSleep [16]	0.794	0.782	0.732	0.725	0.698	0.636
	XSleepNet [10]	0.783	0.761	0.725	0.708	0.689	0.615
	CIMSleepNet	**0.801**	**0.794**	**0.751**	**0.788**	**0.777**	**0.726**
MHR	FeatConcat	0.700	0.464	0.237	0.477	0.243	0.011
	MLP [24]	0.723	0.529	0.306	0.610	0.348	0.035
	DeepCNN [9]	**0.759**	**0.615**	**0.421**	0.616	0.354	0.039
	CIMSleepNet	0.729	0.553	0.348	**0.701**	**0.466**	**0.240**

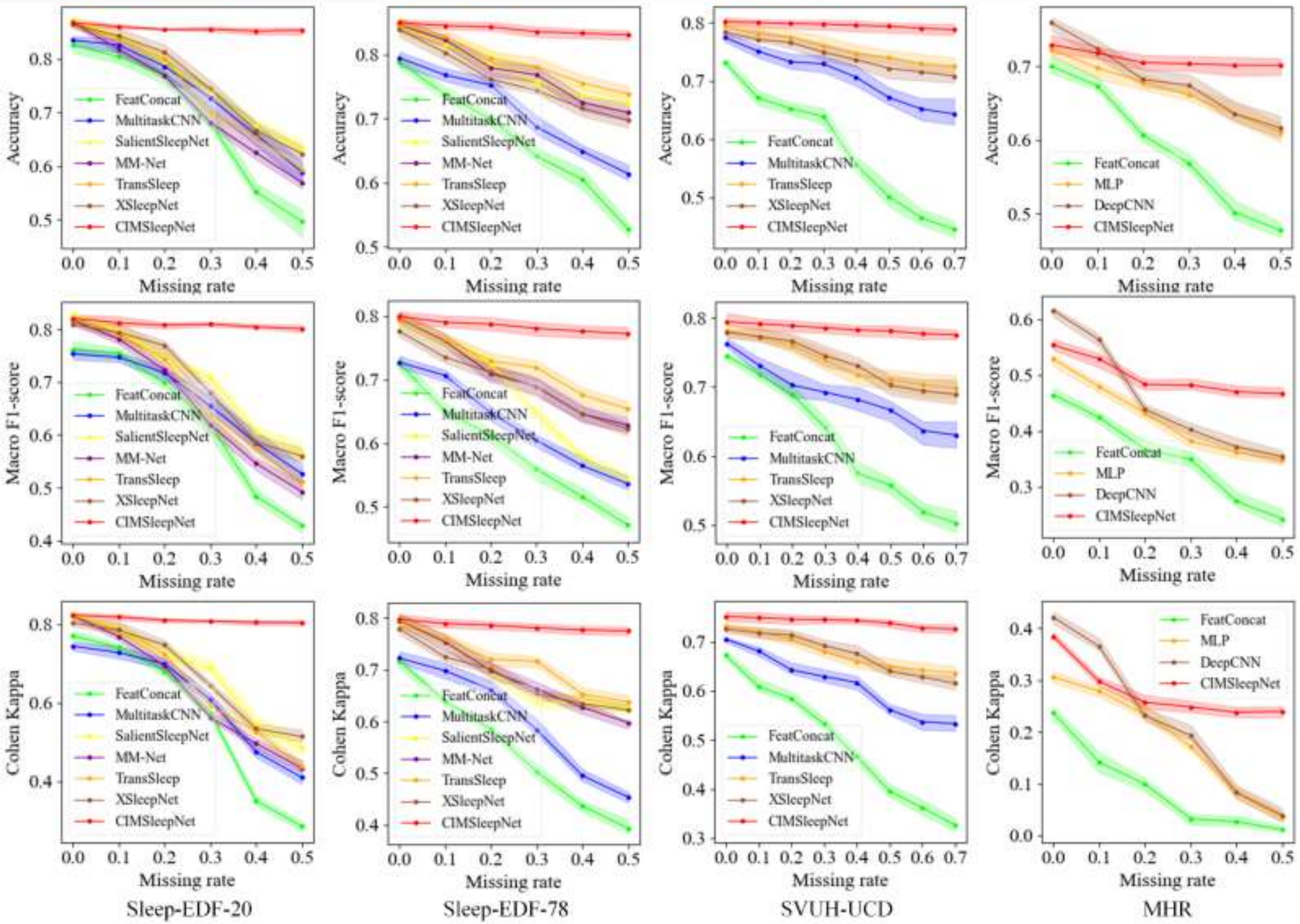
Table 2: Performance comparison in completely missing case.

Test Modalities	Methods	Acc	MF1	K
EEG	CoRe-Sleep [26]	0.882	0.808	0.834
	CIMSleepNet	**0.891**	**0.817**	**0.845**
EOG	CoRe-Sleep [26]	0.853	0.753	0.792
	CIMSleepNet	**0.858**	**0.760**	**0.798**
EEG+EOG	CoRe-Sleep [26]	0.895	0.823	0.853
	CIMSleepNet	**0.903**	**0.828**	**0.862**

3 Experiment

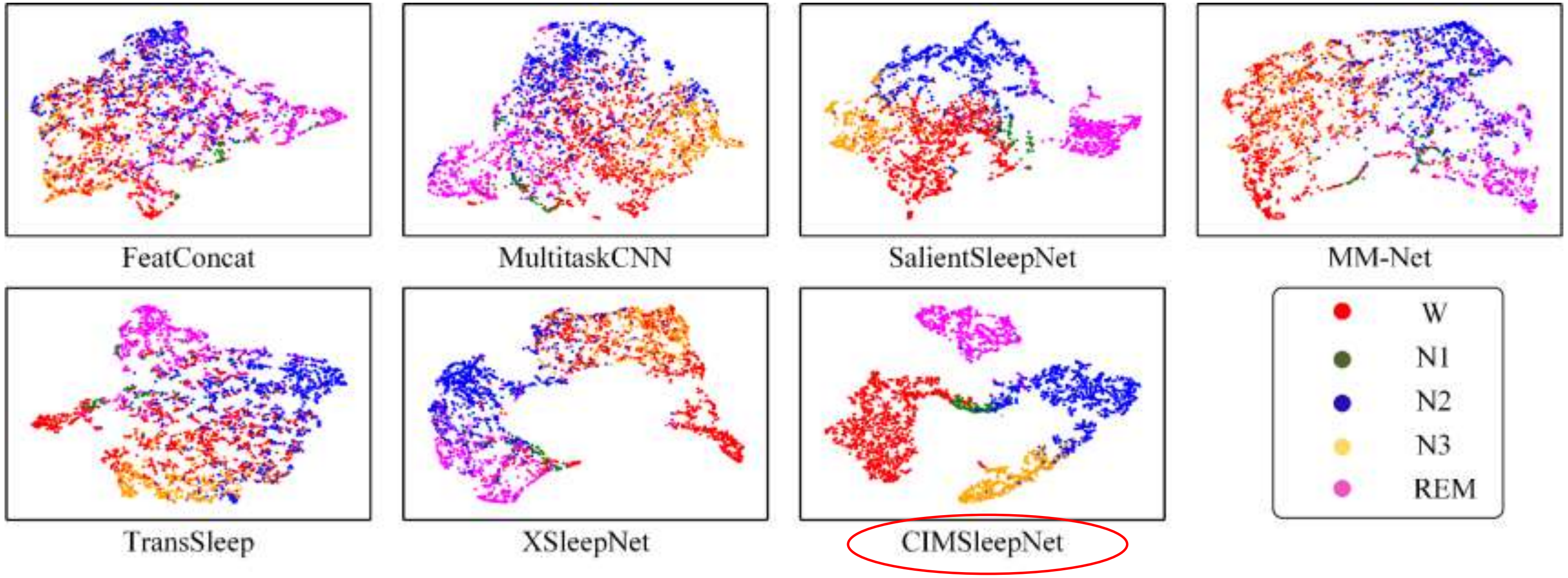
■ Quantitative results

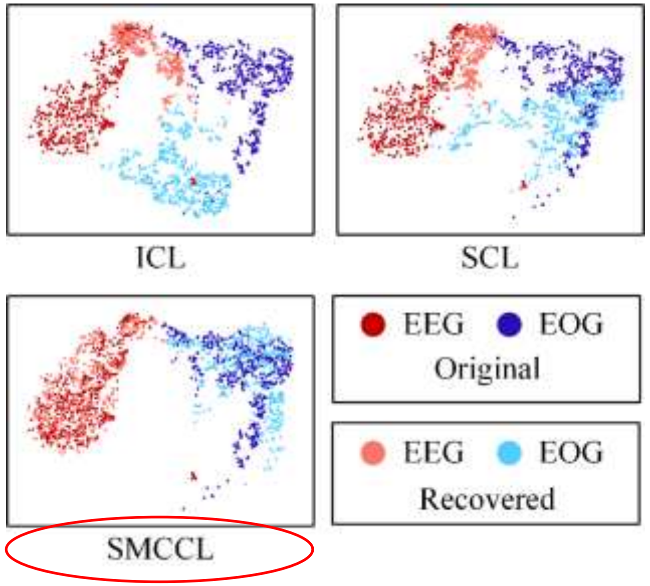
- As the missing rate increases, the performance of other methods begins to decline significantly.
- CIMSleepNet exhibits amore stable trend.



3 Experiment

Qualitative results





3 Experiment

■ Ablation studies

Table 3: Ablation study of CIMSleepNet on Sleep-EDF-20. "✓" indicates the use of this component. MCTA indicates the Transformer equipped with MCTA. The context length of single inference is 25.

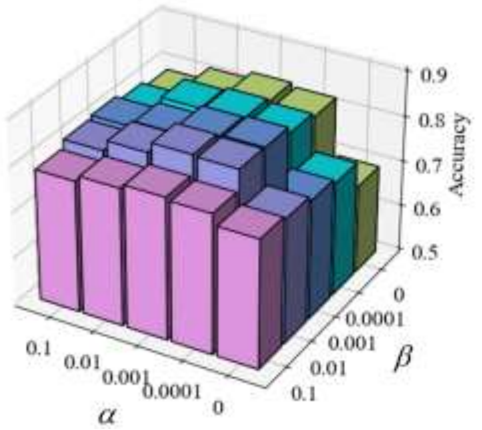
MAIM	SMCCL	MCTA	Acc	MF1	K	Model Size (MB)	GFLOPs
			0.497	0.429	0.285	2.344	0.069
✓			0.771	0.704	0.672	5.767	0.096
	✓		0.786	0.726	0.699	8.458	0.071
		✓	0.694	0.629	0.536	30.272	2.206
✓	✓		0.810	0.756	0.759	4.412	0.097
✓		✓	0.829	0.778	0.777	33.696	2.876
	✓	✓	0.834	0.786	0.784	36.386	2.246
✓	✓	✓	**0.853**	**0.801**	**0.805**	37.678	2.902

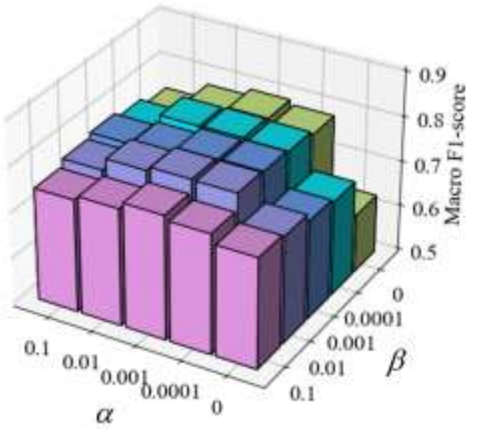
Table 4: Ablation study of Transformer equipped with MCTA on Sleep-EDF-20.

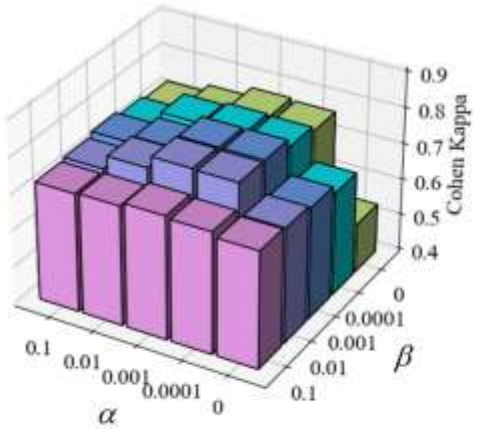
Methods	Acc	MF1	K
Intra-GRU	0.827	0.775	0.772
Inter-GRU	0.835	0.780	0.787
Intra & Inter-GRU	0.839	0.788	0.791
Intra-Transformer	0.813	0.770	0.765
Inter-Transformer	0.837	0.789	0.793
Intra & Inter-Transformer	0.845	0.795	0.797
Transformer with MCTA	**0.853**	**0.801**	**0.805**

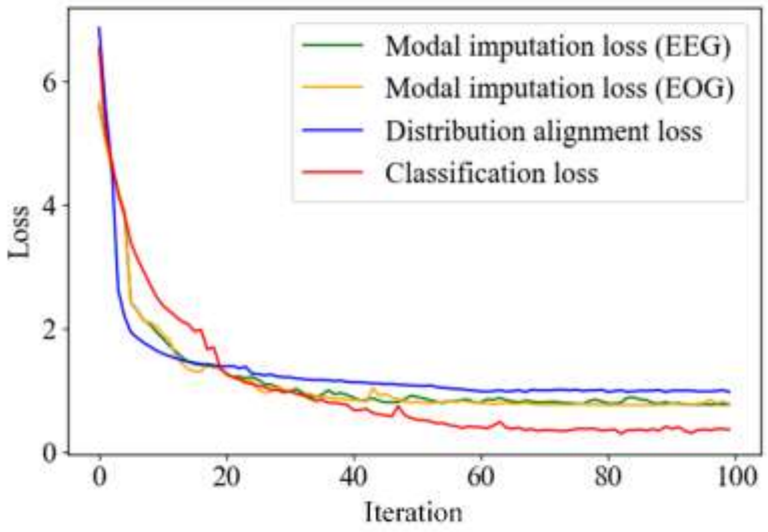
3 Experiment

■ Other analysis









4 Conclusion

- We try to challenge multimodal ASS under incomplete modalities by proposing CIMSleepNet.
- MAIM reconstructs missing modality data by establishing interactions among modalities, which allows for the provision of complete modality data support for subsequent components.
- SMCCL ingeniously leverages semantic information and modal information to subdivide similarity into three levels, thereby simulating real data distribution.
- MCTA mechanism accomplishes comprehensive temporal context modeling, further improving the expressive ability of latent temporal representations.

Thank You For Your Attention!