

Supplementary Materials: SleepMG: Multimodal Generalizable Sleep Staging with Inter-modal Balance of Classification and Domain Discrimination

Anonymous Authors

1 MODEL ARCHITECTURE DESIGN

1.1 FeatureNet

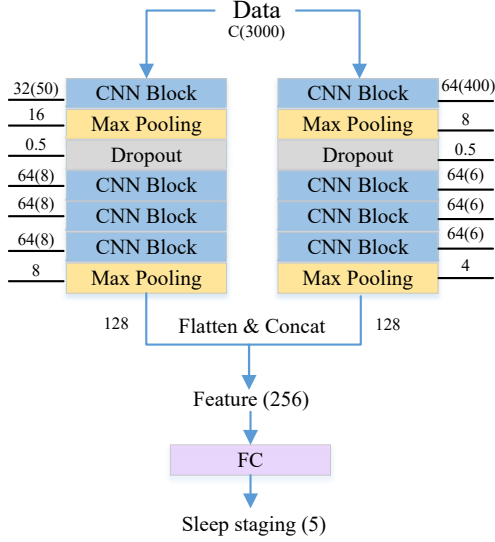


Figure 1: The model architecture of FeatureNet.

Figure 1 illustrates the model architecture of FeatureNet [2, 3], where the 'number(number)' is the 'kernel number(convolution size)' and $C = 10$ channels if FeatureNet is a single baseline method.

1.2 Multimodal sleep staging with FeatureNet

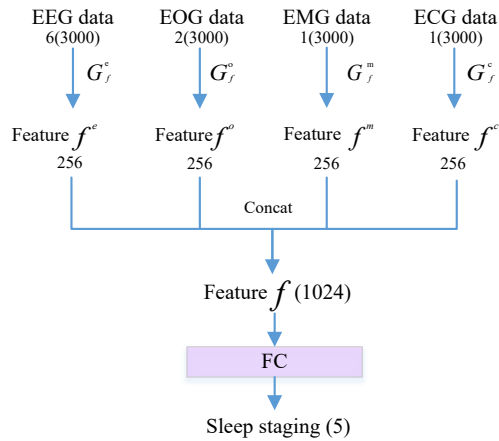


Figure 2: The model architecture of Multimodal sleep staging.

Figure 2 illustrates the model architecture of multimodal sleep staging with FeatureNet. We employ the feature extractor of FeatureNet as the modality-specific feature extractor, followed by the same concatenation and classification with a fully connected model. We add the domain discriminator [1, 3] based on the architecture of Figure 2 for NaiveMG to improve the cross-domain generalization [5].

We employ the same model architecture as NaiveMG for SleepMG. Additionally, we design modality-specific performance assessments [4] in classification and domain discrimination. Leveraging modal performance metrics, we dynamically adjust the gradient of the modality-specific parts of the classifier and domain discriminator, balancing the inter-modal differences in two aspects.

2 CONFUSION MATRIX

2.1 Baseline methods

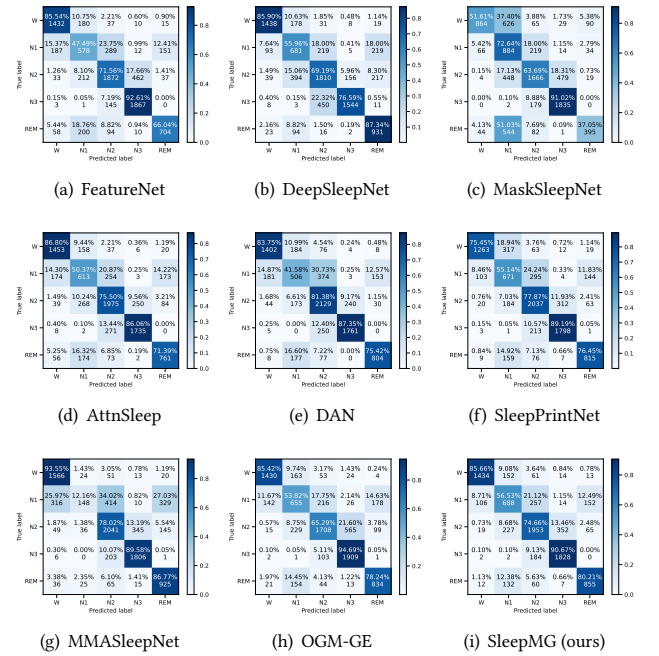


Figure 3: The confusion matrix of the baseline methods on the ISRUC-S3 dataset.

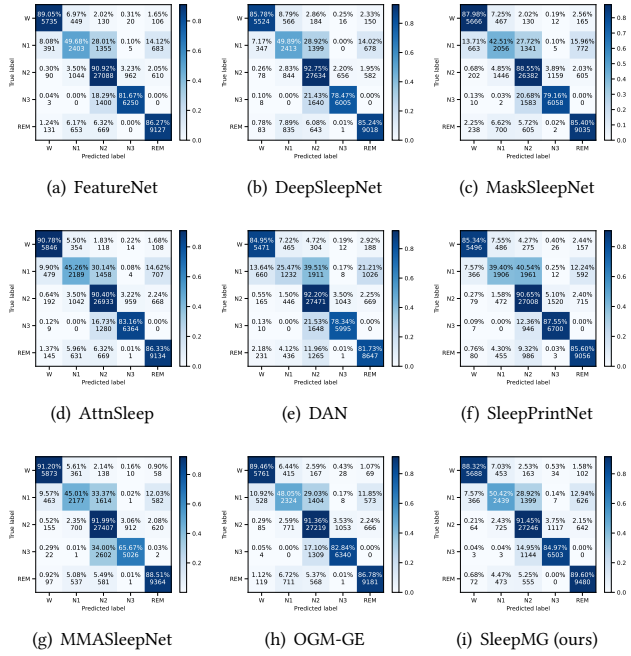


Figure 4: The confusion matrix of the baseline methods on the MASS-SS3 dataset.

2.2 Multimodal and Generalizable modules ablation

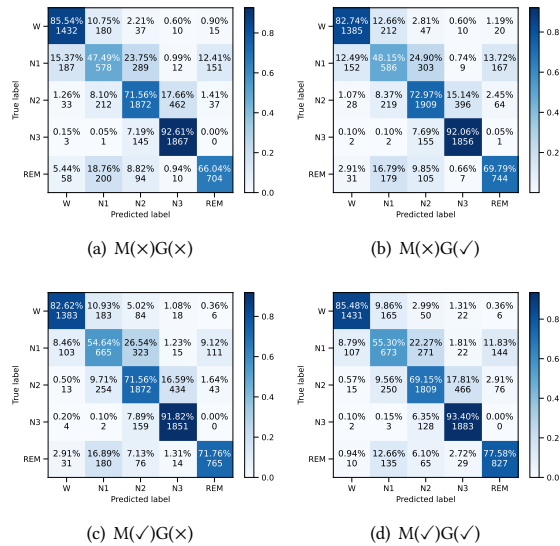


Figure 5: The confusion matrix of the M(ultimodal) and G(eneralizable) modules ablation experiments of NaiveMG(\checkmark) on the ISRUC-S3 dataset.

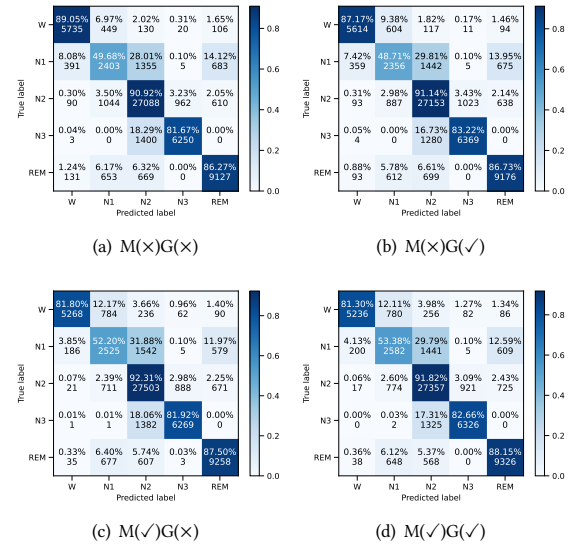


Figure 6: The confusion matrix of the M(ultimodal) and G(eneralizable) modules ablation experiments of NaiveMG(\checkmark) on the MASS-SS3 dataset.

2.3 multimodal feature fusion methods ablation

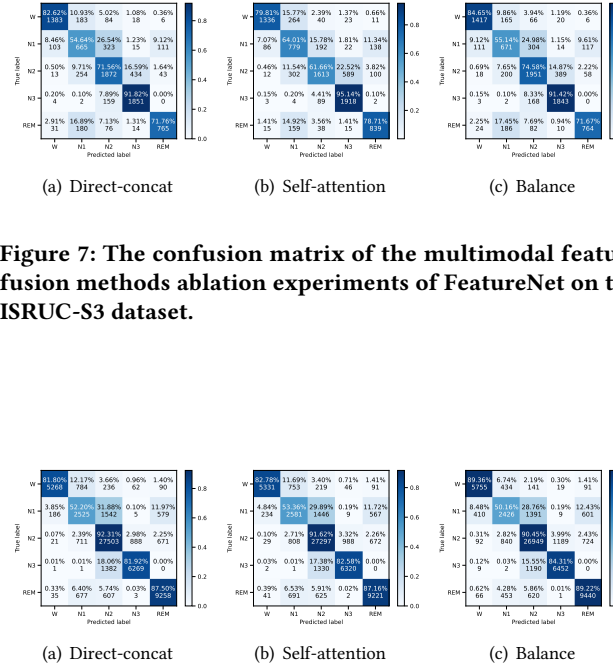


Figure 7: The confusion matrix of the multimodal feature fusion methods ablation experiments of FeatureNet on the ISRUC-S3 dataset.

Figure 8: The confusion matrix of the multimodal feature fusion methods ablation experiments of FeatureNet on the MASS-SS3 dataset.

2.4 balancing different components ablation

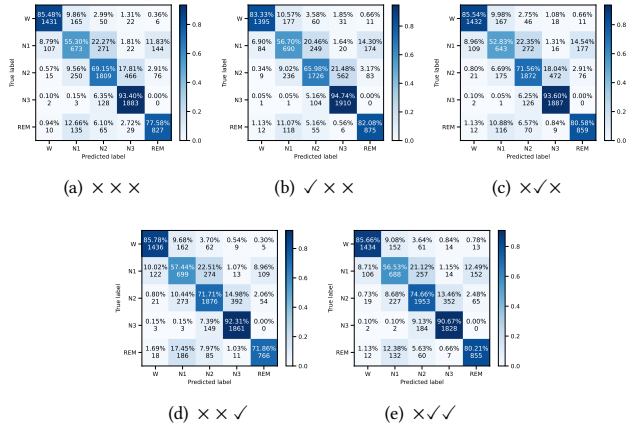


Figure 9: The confusion matrix of balancing different components (G_f^u, G_y^u, G_d^u) of NaiveMG on the ISRUC-S3 dataset.

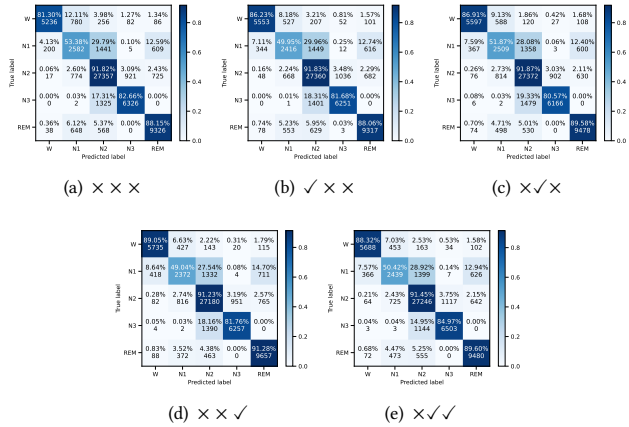


Figure 10: The confusion matrix of balancing different components (G_f^u, G_y^u, G_d^u) of NaiveMG on the MASS-SS3 dataset.

REFERENCES

[1] Yaroslav Ganin, Evgeniya Ustinova, Hana Ajakan, Pascal Germain, Hugo Larochelle, François Laviolette, Mario Marchand, and Victor Lempitsky. 2016. Domain-adversarial training of neural networks. *The journal of machine learning research* 17, 1 (2016), 2096–2030.

[2] Xiaopeng Ji, Yan Li, and Peng Wen. 2022. Jumping Knowledge Based Spatial-temporal Graph Convolutional Networks for Automatic Sleep Stage Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 30 (2022), 1464–1472.

[3] Ziyu Jia, Youfang Lin, Jing Wang, Xiaojun Ning, Yuanlai He, Ronghao Zhou, Yuhan Zhou, and H Lehman Li-wei. 2021. Multi-view spatial-temporal graph

convolutional networks with domain generalization for sleep stage classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 29 (2021), 1977–1986.

[4] Xiaokang Peng, Yake Wei, Andong Deng, Dong Wang, and Di Hu. 2022. Balanced multimodal learning via on-the-fly gradient modulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 8238–8247.

[5] Jindong Wang, Cuiling Lan, Chang Liu, Yidong Ouyang, Tao Qin, Wang Lu, Yiqiang Chen, Wenjun Zeng, and Philip Yu. 2022. Generalizing to unseen domains: A survey on domain generalization. *IEEE Transactions on Knowledge and Data Engineering* 14, 8 (2022), 1–20.