Supplementary materials of AHFL (Adaptive High-order Fusion Learning for Brain Disorder Detection)

V. DISCUSSION

F. Verification of FBN Adaptive Weighting Results

In order to verify the results of the adaptive weighting operation for different order FBNs, in addition to setting the fusion of high-order FBN and low-order FBN in the ablation experiment, in this subsection, the upper triangular elements of a certain order FBN are used as features to conduct the experiment through a simple MLP layer. Specifically, we used the upper triangular elements of the first five FBNs in ABIDE and REST-MDD as input, including (F^1, F^2, F^3, ..., F^5). The two fully connected layers consist of 128 and 64 neurons, respectively, with the sigmoid function applied in the final layer. The dropout factors for both layers are set to 0.5 to prevent overfitting. The results of verification are shown in Tables XI and XII.

TABLE XI

PERFORMANCE EVALUATION (%) OF DIFFERENT ORDERS IN ASD VS. NC ON THE ABIDE DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD

Order	Accuracy	AUC	Precision	Recall	F1-score
first-order	62.16	68.75	43.75	58.33	50.00
second-order	62.16	67.86	50.00	57.14	53.33
fourth-order	62.86	63.69	51.25	71.43	53.48
fifth-order	59.46	62.50	**56.25**	70.00	52.76
third-order	**64.86**	**67.96**	53.33	**80.00**	**58.10**

TABLE XII

PERFORMANCE EVALUATION (%) OF DIFFERENT ORDERS IN MDD VS. NC ON THE REST-MDD DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD

Order	Accuracy	AUC	Precision	Recall	F1-score
first-order	57.01	64.77	58.65	55.34	61.25
second-order	60.75	61.93	54.39	65.96	59.62
fourth-order	56.07	55.51	63.16	58.06	60.50
fifth-order	54.21	58.25	**53.68**	55.26	57.16
third-order	**61.68**	**67.86**	59.65	**65.38**	**62.39**

To further investigate the impact of convergence on classification performance and systematically evaluate the efficacy of combining different orders of FBNs, we design and conduct an ablation experiment involving FBNs from the 1-th to the n-th order, as detailed in Table XIII and Table XIV, in ABIDE and REST-MDD dataset, respectively. For instance, the 1-th to 5-th orders in the table indicate that the first five orders of FBNs from the sample were used as input features. Following the same procedure as the main experiment, feature extraction was performed using the Sero adaptive weighting and attention fusion operation, followed by brain disorder detection. Due to the phenomenon of convergence, higher-order FBNs constructed based on the 15-th order FBN no longer exhibit significant changes. The experimental results demonstrate that the model achieves optimal performance when using the first 15 orders of FBNs for classification. This phenomenon may be attributed to the AHFL effectively integrates the potentially complementary high-order connectivity information among different orders of FBNs. However, introducing higher-order FBNs not only fails to provide additional information gain but may also lead to the learning of redundant parameters, thereby causing overfitting.

TABLE XIII

PERFORMANCE EVALUATION (%) OF DIFFERENT ORDERS COMBINATION IN ASD VS. NC ON THE ABIDE DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD.

Order	Accuracy	AUC	Precision	Recall	F1-score
1 to 5	70.97	70.91	72.73	63.33	61.54
1 to 10	72.16	75.07	72.86	65.43	65.25
1 to 20	71.27	73.94	73.00	64.67	63.93
1 to 15	**73.34**	**77.82**	**73.24**	**69.25**	**69.29**

TABLE XIV

PERFORMANCE EVALUATION (%) OF DIFFERENT ORDERS COMBINATION IN MDD VS. NC ON THE REST-MDD DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD.

Order	Accuracy	AUC	Precision	Recall	F1-score
1 to 5	61.88	64.02	61.18	59.00	63.03
1 to 10	63.87	65.68	65.97	60.77	67.36
1 to 20	62.50	64.32	63.10	61.63	63.86
1 to 15	**65.36**	**66.04**	**74.51**	**62.30**	**67.86**

G. Sero Operation: Effectiveness Analysis

In our main experiments, we employ the Sero function to adaptively weight high-order FBNs and determine their contributions to classification tasks. To evaluate the influence of various readout operations, we compared Mean, Sum, and Max methods with AHFL. Mean calculates the average, Sum performs summation, and Max selects the maximum value. Tables XV and XVI present the classification results of AHFL using Sero with these readout operations. The results show that Sero achieves the best performance on both datasets, outperforming Mean, Sum, and Max. The effectiveness of the attention-based Sero in AHFL is highlighted, while confirming the differential impact of various-order FBNs on downstream tasks.

TABLE XV

PERFORMANCE EVALUATION (%) OF DIFFERENT READOUT APPROACHES IN ASD VS. NC ON THE ABIDE DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD

Method	Accuracy	AUC	Precision	Recall	F1-score
Mean	70.07	72.62	68.12	65.23	**69.85**
Sum	68.12	70.15	69.33	64.12	66.87
Max	63.77	71.12	65.15	64.26	63.25
Sero (ours)	**73.34**	**77.82**	**73.24**	**69.25**	69.29

TABLE XVI

PERFORMANCE EVALUATION (%) OF DIFFERENT READOUT APPROACHES IN MDD VS. NC ON THE REST-MDD DATASET, WITH THE TOP RESULTS EMPHASIZED IN BOLD

Method	Accuracy	AUC	Precision	Recall	F1-score
Mean	70.07	72.62	68.12	65.23	**69.85**
Sum	68.12	70.15	69.33	64.12	66.87
Max	63.77	71.12	65.15	64.26	63.25
Sero (ours)	**73.34**	**77.82**	**73.24**	**69.25**	69.29

H. Model Convergence Criteria

We visualize the objective function loss in our AHFL to intuitively assess convergence during the training process for ASD vs. NC classification on the ABIDE dataset and MDD vs. NC classification on the Rest-MDD dataset. Specifically, we record the loss of AHFL in the training set at the 50 epoch in Figs. 9 and 10, which can be observed that the overall loss showed a downward trend and finally tended to be stable, showing that our model completed convergence with the progress of training.

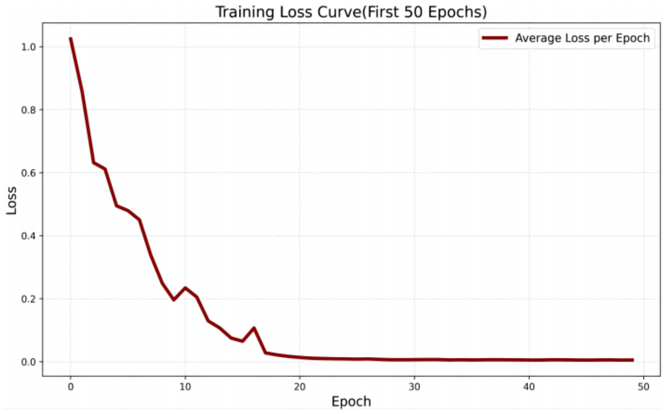

Fig. 9. We show the loss curve corresponding to each epoch in ABIDE dataset during training. Each epoch is composed of 10 batches, we record the average loss of batches in each epoch, which can be clearly seen that the AHFL model converges at about 20 epochs nearly.

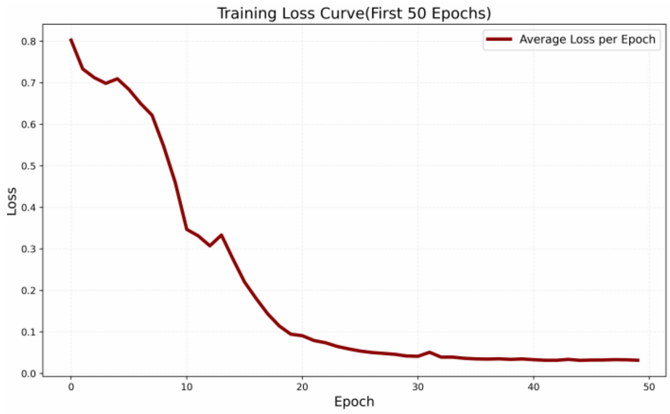

Fig. 10. We show the loss curve corresponding to each epoch in REST-MDD dataset during training. Each epoch is composed of 27 batches, we record the average loss of batches in each epoch, which can be clearly seen that the AHFL model converges at about 25 epochs nearly.