NeuroNetV1: An Adaptive End-to-End Multidomain EEG Classification Network

Supplementary Document

I. Analysis and Discussion on Hyperparameter Tuning

1. Effect of varying Number of Epochs

As shown in Figure 1(a), the classification accuracy showed a clear upward trend as the number of epochs increased, starting from 65.83% at 5 epochs and reaching a peak of 98.75% at 40 epochs. Beyond this point, performance plateaued, with only a marginal improvement at 55 epochs (98.85%) before slightly declining at 60 epochs. The initial increase in accuracy suggests that the model gradually learns complex EEG representations over time, while the saturation beyond 40 epochs implies that further training does not introduce new discriminative information but instead increases the risk of overfitting. This behavior is expected in deep learning models, as excessive training allows the network to memorize training data rather than generalizing well to unseen EEG signals. Consequently, we selected 40 epochs as the optimal choice since it ensures a balance between sufficient training and avoiding unnecessary computational overhead.

2. Effect of varying Learning Rates

The influence of different learning rates on classification accuracy is depicted in Fig. 1(b). The impact of learning rate variation was significant, as it directly influenced the rate of convergence and model stability. Higher learning rates such as $1e^{-3}$ led to unstable learning and frequent fluctuations, which compromised model generalization. In contrast, excessively small values such as $1e^{-6}$ slowed down convergence, causing the model to get stuck in suboptimal minima. The best performance was observed at $1e^{-4}$, achieving an accuracy of 98.75%, as it allowed for fast yet stable convergence. This result aligns with the theoretical understanding that moderate learning rates ensure efficient weight updates without overshooting the loss landscape. Based on this, we adopted $1e^{-4}$ as the default learning rate across all EEG datasets, ensuring consistent and optimized training behavior.

3. Effect of varying Batch Size

The impact of varying batch sizes is shown in Fig. 1(c). Smaller batch sizes (8 and 16) led to high variance in accuracy, likely due to noisy gradient updates that introduced instability. Larger batch sizes (64 and 128) showed slightly lower accuracy, suggesting that reducing stochasticity in updates reduced the model's ability to generalize well. The best performance was observed at batch size 32, which achieved 98.75% accuracy, providing an ideal trade-off between stable weight updates and diverse training samples per batch.

These findings indicate that an intermediate batch size ensures robust feature extraction across multidomain EEG datasets, making it the most efficient choice for NeuroNetV1.

4. Effect of varying Optimizers

The performance of different optimizers is illustrated in Fig. 1(d). Among the tested optimizers, Adam consistently outperformed all other methods, achieving 98.75% accuracy. RMSProp followed closely with 97.97%, while AdaGrad (96.42%), L-BFGS (95.23%), and SGDM (94.92%) lagged behind. Adam's superiority is attributed to its adaptive learning rate strategy, which dynamically adjusts the step size for each parameter, preventing drastic fluctuations and ensuring smooth convergence. SGDM and L-BFGS, on the other hand, suffered from slower adaptation, making them less effective in handling the complex, non-stationary nature of EEG signals. The results confirm that Adam remains the most reliable choice for EEG classification tasks, balancing speed, stability, and convergence efficiency.

5. Effect of different Activation Functions in the Network

As observed in Fig. 1(e), the choice of activation function played a crucial role in determining the model's ability to capture non-linear EEG patterns. ELU yielded the highest accuracy (98.75%), followed by GELU (97.80%) and Leaky ReLU (97.10%). ReLU, while effective, performed slightly worse (96.30%) due to dead neuron issues. Traditional functions like Tanh (94.50%) and Sigmoid (92.10%) exhibited significantly lower performance, as they suffer from saturation problems, leading to vanishing gradients. The superior performance of ELU can be attributed to its ability to maintain a smooth gradient flow, especially for negative inputs, reducing the chances of dead neurons and allowing better feature representation. Given this, ELU was selected as the preferred activation function for NeuroNetV1.

6. Effect of different Loss Functions

The impact of different loss functions is presented in Fig. 1(f). MSE loss yielded the highest accuracy (98.75%), outperforming BCE (97.85%), KL Divergence (97.23%), and Huber Loss (96.87%). The effectiveness of MSE suggests that minimizing squared errors enhances the model's ability to distinguish EEG patterns across multiple domains. BCE, while often used in classification, was less effective due to its sensitivity to imbalanced data, whereas KL Divergence and Huber Loss showed slightly lower accuracy due to their weaker adaptation to multidomain EEG feature variations. Thus, MSE was adopted as the optimal loss function, ensuring stable and accurate classification.

7. Effect of varying Kernel Size

The role of kernel size in different convolutional layers is analyzed in Fig. 1(g). The effect of kernel size was analyzed across the various convolutional layers in NeuroNetV1, with specific focus on dilated convolution (Block 1), ConvAT (Block 2), and depth-wise convolution. In Block 1, which uses dilated convolutions, a kernel size of 8 yielded the best performance (98.75%). This can be attributed to the fact that smaller kernels allow the model to capture local temporal dependencies more efficiently, especially in time-series data like EEG, where fine-grained variations within short time windows are crucial for identifying event-related synchronization and desynchronization patterns. The dilation factor further increases the receptive field, allowing the model to analyze these short-term dependencies over a wider context without requiring larger kernels, thereby reducing the computational burden.

In Block 2, the ConvAT block, which combines convolution and attention mechanisms, performed optimally with a kernel size of 32. Larger kernels are essential here because ConvAT needs to capture long-range contextual dependencies across the entire signal. This kernel size allows the model to integrate information from broader temporal spans, enhancing the self-attention mechanism to focus on the most relevant parts of the signal, which are typically spread across larger windows. By using a larger kernel, the model can better capture global patterns while maintaining higher computational efficiency, crucial for handling the non-stationary nature of EEG signals.

Lastly, the depth-wise convolution (Block 2) showed the best results with a kernel size of 16. Depth-wise convolutions operate on each input channel independently, making it important to strike a balance between capturing local spatial features and retaining important inter-channel relationships. A kernel size of 16 optimally extracts spatially invariant features while preserving enough context to identify inter-frequency dependencies across the EEG bands. This size is large enough to capture meaningful patterns but not so large as to introduce unnecessary computational complexity, which could disrupt the model's ability to generalize across domains.

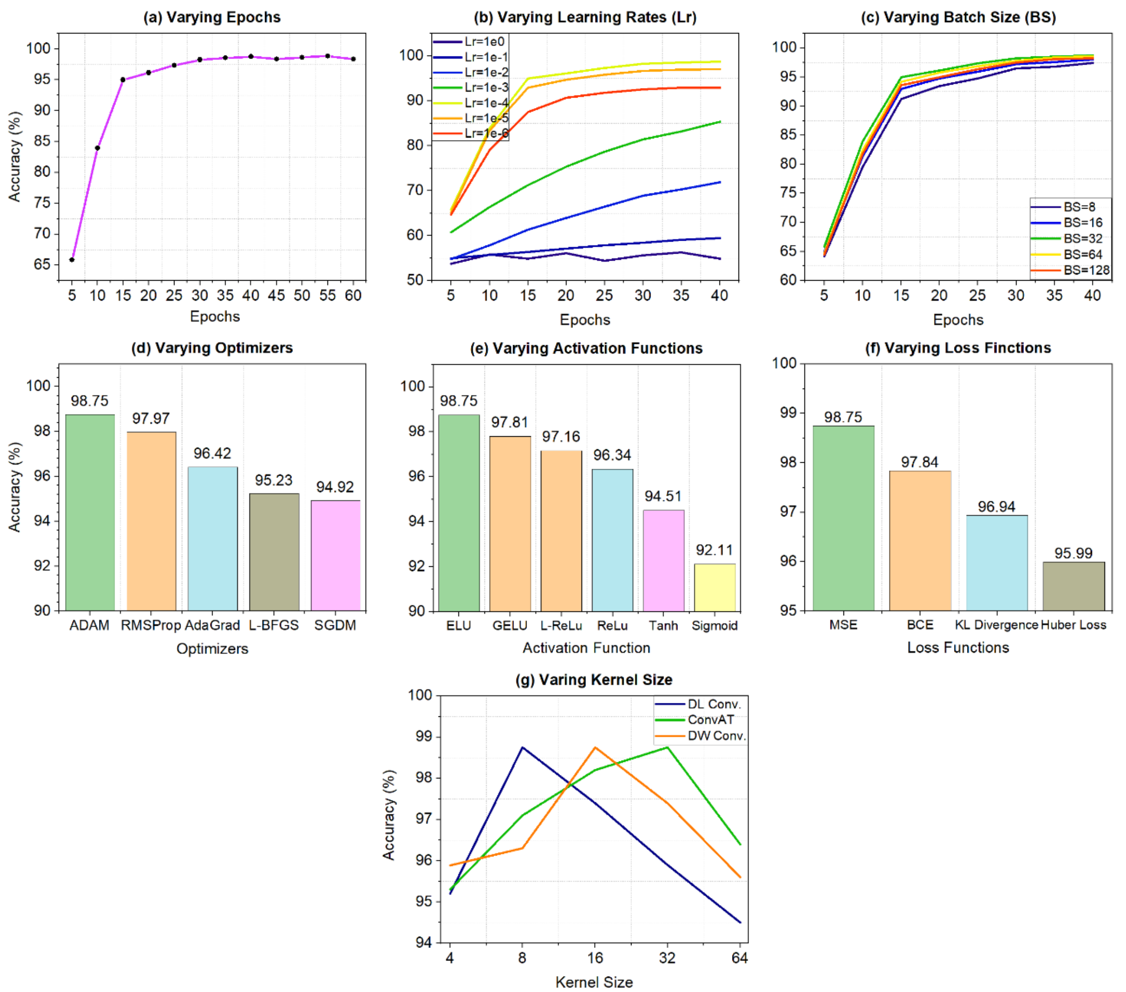

Figure 1: Hyperparameter tuning results for NeuroNetV1 on the MI-I dataset. (a) Number of epochs, (b) learning rate, (c) batch size, (d) optimizers, (e) activation functions, (f) loss functions, and (g) kernel size in different blocks.

II. Subject-wise 5-Fold Average Cross Validation Results for Individual Datasets using NeuroNetV1

This following tables offer a comprehensive analysis of the subject-wise results across all EEG datasets. The performance metrics included are accuracy, sensitivity, specificity, precision, $f1$-score, and Cohen's Kappa coefficient. These results are derived by averaging the 5-fold outcomes for each subject. The detailed results are presented as follows:

Table 1: Classification Performance of NeuroNetV1 on Motor Imagery EEG Dataset MI-I, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	98.80	98.80	98.88	98.82	98.80	97.59
2	99.11	99.11	99.11	99.11	99.11	98.21
3	97.62	97.62	97.62	97.73	97.62	95.24
4	98.21	98.21	98.45	98.28	98.22	96.42
5	100.00	100.00	100.00	100.00	100.00	100.00
Mean ± Std	98.75 ± 0.90	98.75 ± 0.90	98.81 ± 0.87	98.79 ± 0.86	98.75 ± 0.90	97.49 ± 1.81

Table 2: Classification Performance of NeuroNetV1 on Motor Imagery EEG Dataset MI-II, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	93.65	93.65	96.58	93.70	93.66	89.85
2	94.20	94.20	96.95	94.23	94.21	90.73
3	94.75	94.75	97.14	94.78	94.76	91.61
4	93.65	93.65	96.21	93.64	93.63	89.80
5	90.06	90.06	95.96	90.77	90.16	84.39
6	94.20	94.20	96.39	94.22	94.20	90.70
7	91.44	91.44	95.29	91.53	91.45	86.33
8	89.23	89.23	94.35	89.56	89.30	82.93
9	94.20	94.20	96.42	94.36	94.21	90.70
10	89.78	89.78	94.52	89.99	89.82	83.74
11	91.44	91.44	95.28	91.60	91.46	86.34
12	94.20	94.20	96.39	94.19	94.20	90.69
13	92.82	92.82	95.56	92.81	92.81	88.47
14	93.92	93.92	97.42	94.19	93.95	90.38
15	95.30	95.30	97.32	95.32	95.31	92.48
16	94.48	94.48	97.99	94.83	94.53	91.26
17	90.88	90.88	96.05	91.33	90.98	85.59
18	90.61	90.61	95.39	90.83	90.66	85.09
19	95.58	95.58	97.78	95.64	95.59	92.95
20	95.86	95.86	98.07	95.92	95.87	93.39

21	91.99	91.99	96.00	92.23	92.03	87.27
22	93.09	93.09	96.02	93.12	93.10	88.96
23	90.33	90.33	95.67	90.76	90.41	84.74
24	93.92	93.92	96.85	94.08	93.94	90.32
25	93.65	93.65	96.77	93.76	93.67	89.88
26	95.30	95.30	97.32	95.33	95.31	92.49
27	91.99	91.99	95.66	92.09	92.02	87.20
28	92.27	92.27	96.12	92.41	92.30	87.69
29	90.61	90.61	95.39	90.89	90.66	85.12
30	95.03	95.03	97.43	95.12	95.05	92.07
31	89.50	89.50	94.65	89.72	89.54	83.34
32	89.23	89.23	95.32	89.85	89.34	83.05
33	92.82	92.82	96.70	93.04	92.86	88.61
34	95.03	95.03	97.04	95.03	95.02	92.02
35	95.58	95.58	98.35	95.77	95.61	92.97
36	94.48	94.48	97.04	94.54	94.49	91.19
37	95.86	95.86	98.07	95.96	95.88	93.40
38	94.20	94.20	96.94	94.23	94.21	90.70
39	95.86	95.86	98.07	95.89	95.87	93.37
40	95.03	95.03	97.97	95.19	95.06	92.10
41	90.06	90.06	95.02	90.35	90.10	84.27
42	94.20	94.20	96.76	94.27	94.21	90.75
43	95.03	95.03	97.79	95.22	95.04	92.12
44	95.03	95.03	96.66	95.02	95.01	91.99
45	95.30	95.30	98.07	95.44	95.34	92.53
46	93.92	93.92	97.80	94.37	93.96	90.43
47	91.99	91.99	95.86	92.11	92.02	87.24
48	90.61	90.61	95.39	90.83	90.66	85.06
49	94.48	94.48	97.05	94.52	94.49	91.17
50	90.33	90.33	94.94	90.50	90.38	84.61
51	89.50	89.50	95.02	89.90	89.57	83.41
52	95.03	95.03	97.60	95.06	95.04	92.04
53	92.54	92.54	96.02	92.70	92.55	88.12
54	92.82	92.82	97.06	93.14	92.87	88.64
55	90.88	90.88	95.11	91.02	90.92	85.46
56	94.75	94.75	97.13	94.80	94.77	91.61
57	91.44	91.44	95.63	91.72	91.49	86.38
58	90.06	90.06	95.01	90.14	90.09	84.10
59	91.16	91.16	95.76	91.50	91.24	85.98
60	94.20	94.20	96.95	94.27	94.22	90.75
61	93.65	93.65	96.95	93.83	93.67	89.92
62	91.44	91.44	96.04	91.77	91.49	86.45
63	89.50	89.50	94.84	89.88	89.58	83.37

64	95.30	95.30	97.51	95.37	95.31	92.52
65	92.54	92.54	95.84	92.64	92.55	88.11
66	95.03	95.03	97.42	95.04	95.03	92.04
67	92.82	92.82	96.49	92.95	92.84	88.56
68	95.58	95.58	97.60	95.60	95.59	92.93
69	91.44	91.44	95.85	91.63	91.49	86.39
70	91.71	91.71	95.19	91.75	91.72	86.76
71	91.44	91.44	95.10	91.58	91.45	86.36
72	95.03	95.03	97.23	95.06	95.03	92.06
73	93.65	93.65	96.96	93.73	93.66	89.87
74	90.06	90.06	94.65	90.19	90.07	84.14
75	91.16	91.16	95.80	91.50	91.24	86.00
76	93.92	93.92	96.49	93.92	93.92	90.25
77	93.65	93.65	97.55	93.98	93.72	89.94
78	90.06	90.06	95.58	90.64	90.14	84.35
79	89.23	89.23	94.19	89.39	89.27	82.86
80	90.33	90.33	94.93	90.56	90.36	84.67
81	94.20	94.20	96.59	94.20	94.20	90.69
82	92.82	92.82	96.50	92.96	92.85	88.57
83	93.92	93.92	97.41	94.21	93.98	90.36
84	89.78	89.78	95.86	90.48	89.90	83.93
85	89.78	89.78	94.74	89.99	89.82	83.76
86	92.82	92.82	95.76	92.93	92.79	88.47
87	95.03	95.03	97.77	95.19	95.06	92.09
88	91.97	91.97	96.29	92.27	92.03	87.28
89	95.77	95.77	98.21	95.92	95.80	93.27
90	91.16	91.16	96.15	91.63	91.24	86.05
91	94.20	94.20	96.40	94.24	94.20	90.71
92	94.04	94.04	97.32	94.19	94.08	90.52
93	91.44	91.44	94.89	91.54	91.39	86.23
94	90.33	90.33	94.74	90.54	90.36	84.64
95	94.48	94.48	97.61	94.59	94.51	91.20
96	92.27	92.27	96.13	92.44	92.31	87.69
97	93.37	93.37	96.86	93.53	93.40	89.46
98	91.44	91.44	95.11	91.62	91.47	86.38
99	93.09	93.09	96.40	93.18	93.12	88.99
100	90.91	90.91	95.30	91.09	90.94	85.53
101	94.48	94.48	97.24	94.57	94.49	91.19
102	95.03	95.03	97.22	95.09	95.05	92.04
103	91.16	91.16	95.00	91.15	91.15	85.82
104	94.13	94.13	96.73	94.19	94.15	90.64
105	93.65	93.65	97.33	94.00	93.67	89.97
106	91.79	91.79	95.69	91.91	91.83	86.89

107	92.54	92.54	96.22	92.65	92.56	88.12
108	92.27	92.27	96.32	92.52	92.29	87.75
109	95.03	95.03	97.61	95.16	95.03	92.11
Mean ± Std	92.84 ± 1.95	92.84 ± 1.95	96.37 ± 1.04	93.01 ± 1.87	92.87 ± 1.93	88.61 ± 3.08

Table 3: Classification Performance of NeuroNetV1 on Motor Imagery EEG Dataset MI-III, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	96.00	96.00	96.00	96.02	96.00	92.00
2	91.50	91.50	91.50	91.50	91.50	83.00
3	96.50	96.50	96.50	96.54	96.50	93.00
4	94.50	94.50	94.50	94.54	94.50	89.00
5	97.00	97.00	97.00	97.00	97.00	94.00
6	95.00	95.00	95.00	95.07	95.00	90.00
7	92.92	92.92	92.92	92.94	92.92	85.83
8	93.50	93.50	93.50	93.54	93.50	87.00
9	94.17	94.17	94.17	94.36	94.16	88.33
10	97.00	97.00	97.00	97.02	97.00	94.00
11	92.50	92.50	92.50	92.50	92.50	85.00
12	96.00	96.00	96.00	96.07	96.00	92.00
13	91.50	91.50	91.50	91.50	91.50	83.00
14	91.50	91.50	91.50	91.54	91.50	83.00
15	95.00	95.00	95.00	95.02	95.00	90.00
16	93.00	93.00	93.00	93.16	92.99	86.00
17	94.50	94.50	94.50	94.54	94.50	89.00
18	96.00	96.00	96.00	96.00	96.00	92.00
19	93.50	93.50	93.50	93.50	93.50	87.00
20	95.00	95.00	95.00	95.00	95.00	90.00
21	92.50	92.50	92.50	92.50	92.50	85.00
22	94.00	94.00	94.00	94.07	94.00	88.00
23	93.00	93.00	93.00	93.07	93.00	86.00
24	92.00	92.00	92.00	92.07	92.00	84.00
25	94.50	94.50	94.50	94.54	94.50	89.00
26	91.50	91.50	91.50	91.50	91.50	83.00
27	96.50	96.50	96.50	96.54	96.50	93.00
28	97.50	97.50	97.50	97.50	97.50	95.00
29	92.00	92.00	92.00	92.02	92.00	84.00
30	94.00	94.00	94.00	94.00	94.00	88.00
31	96.00	96.00	96.00	96.00	96.00	92.00
32	92.50	92.50	92.50	92.50	92.50	85.00
33	92.50	92.50	92.50	92.71	92.49	85.00
34	94.00	94.00	94.00	94.07	94.00	88.00
35	92.00	92.00	92.00	92.27	91.99	84.00

36	93.00	93.00	93.00	93.00	93.00	86.00
37	93.50	93.50	93.50	93.61	93.50	87.00
38	96.50	96.50	96.50	96.62	96.50	93.00
39	96.00	96.00	96.00	96.02	96.00	92.00
40	91.50	91.50	91.50	91.60	91.49	83.00
41	91.50	91.50	91.50	91.50	91.50	83.00
42	93.00	93.00	93.00	93.02	93.00	86.00
43	92.50	92.50	92.50	92.50	92.50	85.00
44	93.00	93.00	93.00	93.02	93.00	86.00
45	92.00	92.00	92.00	92.02	92.00	84.00
46	92.92	92.92	92.92	93.16	92.91	85.83
47	96.50	96.50	96.50	96.50	96.50	93.00
48	97.50	97.50	97.50	97.50	97.50	95.00
49	95.50	95.50	95.50	95.50	95.50	91.00
50	92.00	92.00	92.00	92.00	92.00	84.00
51	92.50	92.50	92.50	92.50	92.50	85.00
52	97.50	97.50	97.50	97.50	97.50	95.00
Mean ± Std	94.00 ± 1.90	94.00 ± 1.90	94.00 ± 1.90	94.05 ± 1.89	94.00 ± 1.90	88.00 ± 3.81

Table 4: Classification Performance of NeuroNetV1 on Motor Imagery EEG Dataset MI-IV, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	94.79	94.79	98.26	94.81	94.78	93.06
2	94.44	94.44	98.15	94.46	94.43	92.59
3	94.10	94.10	98.03	94.18	94.11	92.13
4	97.92	97.92	99.31	97.93	97.92	97.22
5	98.26	98.26	99.42	98.30	98.27	97.69
6	94.44	94.44	98.15	94.47	94.45	92.59
7	97.57	97.57	99.19	97.60	97.57	96.76
8	98.26	98.26	99.42	98.29	98.26	97.69
9	99.31	99.31	99.77	99.31	99.31	99.07
Mean ± Std	96.57 ± 2.07	96.57 ± 2.07	98.86 ± 0.69	96.59 ± 2.06	96.57 ± 2.07	95.42 ± 2.76

Table 5: Classification Performance of NeuroNetV1 on Mental Imagery EEG Dataset MeI-V, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	98.60	98.60	99.37	98.63	98.61	97.89
2	97.76	97.76	98.94	97.78	97.76	96.61
3	98.17	98.17	99.08	98.18	98.17	97.26
Mean ± Std	98.18 ± 0.42	98.18 ± 0.42	99.13 ± 0.22	98.19 ± 0.43	98.18 ± 0.42	97.25 ± 0.64

Table 6: Classification Performance of NeuroNetV1 on P300 EEG Dataset P300-VI, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	95.88	95.88	96.62	96.05	95.91	91.12
2	95.29	95.29	95.79	96.91	95.77	71.75
3	96.18	96.18	94.18	96.85	96.38	80.55
4	96.76	96.76	97.50	96.90	96.79	92.88
5	97.35	97.35	97.09	97.36	97.35	94.63
6	96.76	96.76	96.78	96.77	96.77	93.52
7	95.29	95.29	95.25	95.35	95.31	89.95
8	97.94	97.94	97.84	97.95	97.94	95.63
9	96.47	96.47	95.61	96.47	96.47	92.08
10	96.76	96.76	96.37	96.89	96.80	91.13
11	95.59	95.59	95.90	95.76	95.63	89.80
12	95.29	95.29	99.58	97.01	95.76	75.29
13	97.06	97.06	96.36	97.66	97.23	82.28
14	97.06	97.06	96.09	97.05	97.05	93.51
15	96.18	96.18	97.24	96.42	96.22	91.07
16	97.35	97.35	97.06	97.43	97.35	94.67
Mean ± Std	96.45 ± 0.82	96.45 ± 0.82	96.58 ± 1.22	96.80 ± 0.68	96.54 ± 0.74	88.74 ± 7.25

Table 7: Classification Performance of NeuroNetV1 on Slow Cortical Potential EEG Dataset SCP-VII, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	95.90	95.90	95.87	95.96	95.89	91.79
2	92.50	92.50	92.50	92.50	92.50	85.00
Mean ± Std	94.20 ± 2.40	94.20 ± 2.40	94.18 ± 2.38	94.23 ± 2.44	94.20 ± 2.40	88.39 ± 4.80

Table 8: Classification Performance of NeuroNetV1 on Emotions EEG Dataset Emot-VIII, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	97.99	97.99	99.00	97.99	97.99	96.98
2	94.31	94.31	97.16	94.31	94.31	91.45
3	93.72	93.72	96.89	93.79	93.72	90.57
4	96.81	96.81	98.39	96.81	96.81	95.21
5	97.35	97.35	98.67	97.35	97.35	96.02
6	98.09	98.09	99.04	98.09	98.09	97.13
7	95.14	95.14	97.58	95.14	95.14	92.71
8	93.86	93.86	96.90	93.89	93.86	90.79
9	92.73	92.73	96.36	92.73	92.73	89.10
10	93.27	93.27	96.65	93.28	93.28	89.91
11	94.55	94.55	97.25	94.56	94.55	91.82
12	95.24	95.24	97.62	95.25	95.24	92.85
13	93.47	93.47	96.74	93.47	93.47	90.20

14	98.23	98.23	99.11	98.24	98.23	97.35
15	98.87	98.87	99.44	98.88	98.87	98.31
Mean ± Std	95.58 ± 2.10	95.58 ± 2.10	97.79 ± 1.05	95.59 ± 2.10	95.58 ± 2.10	93.36 ± 3.15

Table 9: Classification Performance of NeuroNetV1 on Sleep Stage Dataset Sleep-IX, Averaged Over 5 Folds for Each Subject

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	95.86	95.86	98.92	95.94	95.88	94.58
2	96.69	96.69	99.31	96.81	96.72	95.02
3	91.78	91.78	98.08	92.23	91.88	88.58
4	99.16	99.16	99.78	99.17	99.16	98.78
5	98.30	98.30	99.53	98.36	98.31	97.58
6	95.83	95.83	98.93	95.94	95.85	94.24
7	97.08	97.08	99.36	97.18	97.10	95.96
8	94.09	94.09	98.45	94.19	94.11	92.23
9	95.23	95.23	98.87	95.48	95.30	93.78
10	98.02	98.02	99.65	98.10	98.04	97.13
11	94.93	94.93	98.65	95.42	95.06	91.73
12	97.69	97.69	99.56	97.74	97.71	96.50
13	95.07	95.07	98.58	95.17	95.09	93.40
14	83.45	83.45	96.02	85.23	83.92	76.82
15	90.80	90.80	97.77	91.84	91.10	87.47
16	98.79	98.79	99.68	98.80	98.79	98.35
17	94.87	94.87	98.84	95.16	94.94	92.96
18	98.71	98.71	99.64	98.73	98.72	98.22
19	93.78	93.78	98.46	93.98	93.81	91.77
20	94.15	94.15	98.72	94.63	94.26	91.75
Mean ± Std	95.21 ± 3.59	95.21 ± 3.59	98.84 ± 0.87	95.51 ± 3.20	95.29 ± 3.49	93.34 ± 4.99

Table 10: Classification Performance of NeuroNetV1 on Schizophrenia EEG Dataset SZ-XI, Averaged Over 5 Folds for Each Subject. The data for all 81 subjects was combined to form the training and testing set. Here we present the classification accuracies for three different cases.

Cases	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)
1	98.51	98.51	98.58	98.52	98.51	96.91
2	99.13	99.13	99.15	99.13	99.13	98.21
3	99.33	99.33	99.34	99.33	99.33	98.62
Mean ± Std	98.99 ± 0.43	98.99 ± 0.43	99.02 ± 0.40	98.99 ± 0.43	98.99 ± 0.43	97.91 ± 0.89

Table 11: Classification Performance of NeuroNetV1 on Alcoholism EEG Dataset AC-XI, Averaged Over 5 Folds for Each Subject. The data for all 122 subjects was combined to form the training and testing set.

Subjects	Accuracy (%)	Sensitivity (%)	Specificity (%)	Precision (%)	F1-score (%)	Kappa (%)

1-122 combined	98.29	98.29	98.27	98.30	98.29	96.34

Table 12: Classification Performance of NeuroNetV1 on Alzheimer EEG Dataset AD-XI, Averaged Over 5 Folds for Each Subject. The data for all 88 subjects was combined to form the training and testing set.

Subjects	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**F1-score (%)**	**Kappa (%)**
1-88 combined	89.18	89.18	89.13	89.28	89.21	77.88

Table 13: Classification Performance of NeuroNetV1 on Epilepsy EEG Dataset EP-XI, Averaged Over 5 Folds for Each Subject

Subjects	**Accuracy (%)**	**Sensitivity (%)**	**Specificity (%)**	**Precision (%)**	**F1-score (%)**	**Kappa (%)**
1	98.82	98.82	98.82	98.82	98.82	97.63
2	98.72	98.72	98.72	98.72	98.72	97.43
3	99.15	99.15	99.15	99.15	99.15	98.30
4	98.78	98.78	98.78	98.78	98.78	97.57
5	99.38	99.38	99.38	99.38	99.38	98.77
Mean ± Std	98.97 ± 0.29	98.97 ± 0.29	98.97 ± 0.29	98.97 ± 0.29	98.97 ± 0.29	97.94 ± 0.57

III. Multiple Comparison Correction: Benjamini-Hochberg Results

Multiple hypothesis testing increases the risk of false-positive results, necessitating statistical correction techniques to control for inflated Type I errors. In this study, we applied the Benjamini-Hochberg (BH) correction to the Wilcoxon signed-rank test p-values to ensure that our statistical significance findings remain robust against multiple comparisons. The BH correction method is particularly suitable for our analysis, as it controls the false discovery rate (FDR) while maintaining higher statistical power compared to conservative approaches such as the Bonferroni correction. To assess the impact of multiple comparison correction, we computed the BH-adjusted p-values for all statistical comparisons conducted in the manuscript. The results are illustrated in Fig. 2, which provides a direct comparison between the original Wilcoxon p-values (blue bars) and the BH-corrected p-values (yellow bars) for all models analyzed in this study.

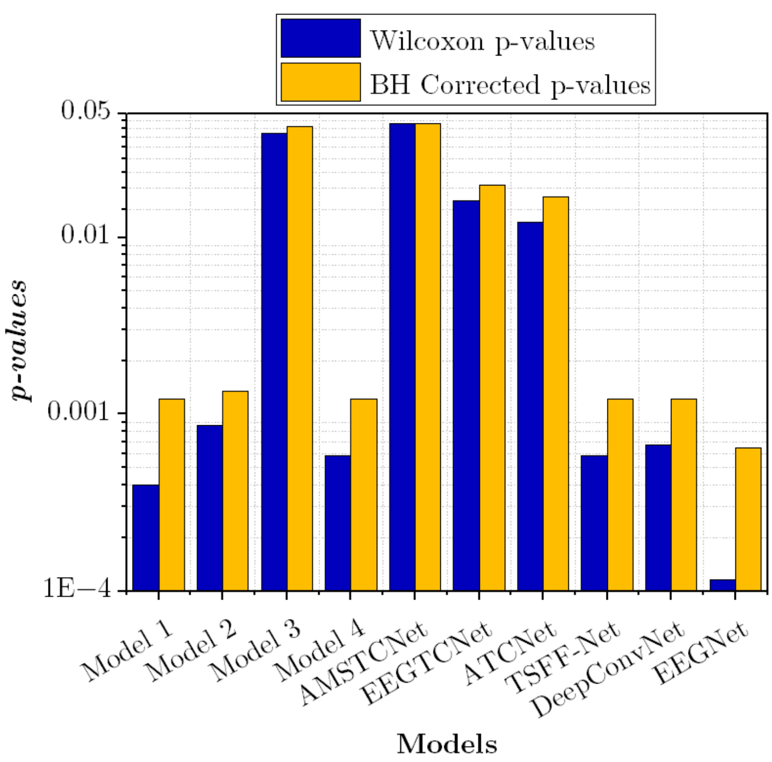

Figure 2: Comparison of Wilcoxon signed-rank test p-values (blue) and Benjamini-Hochberg (BH) corrected p-values between NeuroNetV1 and different models.

After applying the BH correction, the p-values slightly increased across all models compared to their original Wilcoxon values. However, all corrected p-values remained below the 0.05 threshold, confirming that the reported statistical significance of NeuroNetV1's performance is not an artifact of multiple comparisons. This result reinforces the validity of our findings and suggests that the improvements demonstrated by NeuroNetV1 are statistically robust. The extent of correction varied across models, with Model 3 and AMSTCNet exhibiting a more noticeable increase in p-values post-correction. This variation is expected, as the BH procedure ranks p-values and applies corrections based on their relative significance levels. Despite the adjustments, all models retained statistically significant differences, further confirming the reliability of our comparative analysis. Conversely, models such as DeepConvNet, EEGNet, and TSFF-Net showed only minor increases in p-values, indicating that their statistical significance was already strong and less affected by FDR correction. This approach ensures that our statistical analysis remains both rigorous and interpretable while avoiding unnecessary modifications to the main manuscript.

IV. Difference Between ConvAT, Conformer and ATCNet

Figure 2 below illustrates convolutional attention mechanisms across three scenarios: ***a)*** Mimicking the *ConFormer architecture*, where features are first extracted from the EEG signals and then passed to the attention network to compute attention weights for different features. ***b)*** Mimicking the *ATCnet architecture*, where the attention mechanism is applied to the EEG features first, and then convolutional kernels are used to extract temporal features. ***c)*** Our *ConvAT mechanism*, where convolution is integrated directly within the attention process, enabling the model to extract both local and global contextually aware features.

From these architectures, we can identify two major differences in how the mechanisms function:

i. *Traditional Attention in ConFormer and ATCnet:* These models use a conventional attention mechanism that relies on separate weight matrices for queries, keys, and values (denoted as W_q, W_k, W_v). The dimensions of these matrices are $W_q = d \times d_q$, $W_k = d \times d_k$, $W_v = d \times d_v$, accordingly, where d is the embedding size, and d_q, d_k is and d_v represent the query, key, and value dimensions, respectively.

ii. *ConvAT Block:* In contrast, the ConvAT block modifies the traditional attention mechanism by replacing the weight matrices with convolutional kernels. The kernel sizes in ConvAT are designed to match the 1-dimensional time sequence of the input ($T_s \times 1$), allowing the model to focus on local and global dependencies directly in the attention process.

Below, we outline the specific innovations that set ConvAT apart from prior convolution-attention designs:

i. ***Reduction in Weight Matrix Complexity:*** The replacement of traditional attention weight matrices (W_q, W_k, W_v) with convolution kernels simplifies the model architecture, potentially reducing the computational burden and improving efficiency. The number of trainable parameters in traditional attention model are $3d(d_q + d_k + d_v)$ (combining the weight matrices of query, key and value). Whereas, the number of trainable parameters in ConvAT are $3T_s$. This drastically reduces the number of parameters while keeping the soul of attention process intact. A more comprehensive comparison for model parameters and computational efficiency is given in ***Section 4.5 of the manuscript***.

ii. ***Improved Temporal Contextualization:*** Weight matrix multiplication in ConFormer and ATCNet applies a uniform transformation across all time steps, treating each segment equally in terms of attention. It does not allow for nuanced, dynamic focus on short-term vs long-term temporal patterns. In contrast, ConvAT uses convolutional kernels, which are inherently more sensitive to local features and can prioritize temporal patterns at different scales. This results in more effective temporal contextualization, especially for EEG signals with both short-term and long-range dependencies.

iii. ***Flexibility in Temporal Representation:*** The convolutional operation in ConvAT enables the model to learn dynamic time-specific features. The local context captured by convolutions is crucial for accurately understanding time-sensitive phenomena in EEG signals, such as event-related potentials (ERPs), rhythmic brain oscillations and disorder-oriented signals. This is not as naturally possible with fixed weight matrices in traditional attention mechanisms, where long-range dependencies are emphasized more than the short-term local features that are often the most relevant in EEG classification.

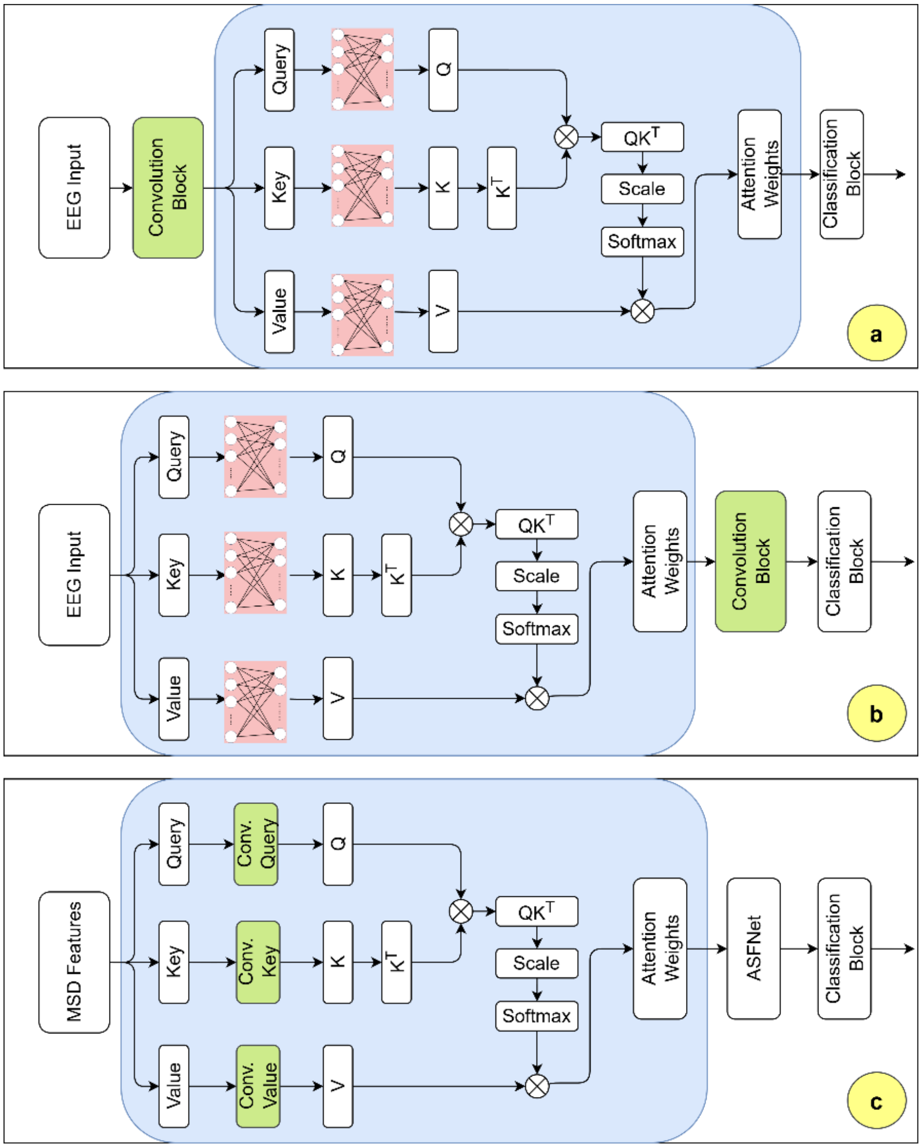

Figure 3: Coupling patterns between convolution and self-attention in EEG models. a) ConFormer-style (Conv→Attn); b) ATCNet-style (Attn→Conv); c) Proposed ConvAT (Conv ⊂ Attn). The Green region indicates convolution operations (either stand-alone blocks or kernels inserted in attention), red denotes fully connected projection weight matrices used in conventional attention, and lastly, blue region showcases self-attention workspace.

V. Extended Comparisons of NeuroNetV1 for Individual Datasets

Table 14: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the MI-I Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	Shalu et.al [1]	2019	DCNN	99.35
2	**Our Study**	**2025**	**NeuroNetV1**	**98.75**
3	Sadiq et al [2]	2019	MEWT	97
4	Taheri et.al [3]	2020	CT+ Fourier EEMD+ CSP+DCNN	96.34
5	Wijaya et al. [4]	2021	LRFS+TSD	95.21
6	Xianglong et.al [5]	2025	Multi-Domain Feature Rotation and Stacking Ensemble	92.92
7	Yu et.al [6]	2015	SFBCSP	92.05
8	Ming et.al [7]	2024	STGAT-CS	91.50
9	Liang et.al [8]	2023	PCC+GCN	89.14
10	Miao et.al [9]	2020	Adaptive Multi-Domain Feature Optimization	87.80
11	Miao et.al [10]	2021	SFT-3D CNN	86.60

Table 15: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the MI-II Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	Hamidi et.al [11]	2025	Transformer+ GCN	97.43
2	Hou et al. [12]	2022	Attention-based BiLSTM-GCN	95.48
3	Huang et.al [13]	2023	RP-BCNNs	94.07
4	Hou et al. [14]	2022	GCNs-net	93
5	**Our Study**	**2025**	**NeuroNetV1**	**92.84**
6	Huang et al. [15]	2023	Convolutional Sliding window Attention Network (CSANet)	92.36
7	Fan et.al [16]	2023	3D-convolutional neural networks	89.86
8	Dose et.al [17]	2018	CNNs	87.98
9	Chowdhury et.al [18]	2023	EEGNet Fusion V2	87.80
10	Moaveninejad et al. [19]	2024	Using Fractal Dimension as a discriminative feature + machine learning	86
11	Lin et al. [20]	2024	NGC-STCSA SaSFS	84.49

Table 16: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the MI-III Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	Wu et.al [21]	2023	Compact CNN	96.75
2	**Our Study**	**2025**	**NeuroNetV1**	**94**
3	Grear et al. [22]	2021	DR+ICA+SVM	93
4	Fan et.al [16]	2023	Compact 3D-CNN	91.91
5	Huang et al. [23]	2022	EFD-CNN	89.97
6	Chowdhury et.al [24]	2024	AIDC-CN	89.47
7	Zhang et.al [25]	2024	GPL	84.22
8	Yu et.al [26]	2021	IEFD	83.84
9	Kumar et.al [27]	2019	LSTM_CSP	82.22
10	Park et.al [28]	2023	3D-EEGNet	81.31
11	Zheng et.al [29]	2021	Adaptive layer+ fully connected layer	76

Table 17: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the MI-IV Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	Phadikar et.al [100]	2023	Transforming EEG signal into a new domain, weight vector of autoencoder, unsupervised neural network	97
2	**Our Study**	**2025**	**NeuroNetV1**	**96.57**
3	Wang et.al [30]	2024	ERDFIS-2DSCG	89.89
4	Cai et.al [31]	2024	MT-MBCNN	89.30
5	Zhang et.al [32]	2023	CLRNet	89
6	Yang et.al [33]	2024	MSFCNNet	87.16
7	Lian et.al [34]	2024	An end-to-end deep neural network	85.10
8	Song et.al [35]	2023	FBCSP+ Transformer	84.16
9	Gu et.al [36]	2025	LSTM+ Transformer	83.02
10	Zhao et.al [37]	2024	CTNet	83
11	Zheng et.al [29]	2021	Adaptive layer+ fully connected layer	82
12	Zhao et.al [38]	2025	Multi-branch temporal convolutional network	81.47

Table 18: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the MeI-V Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	Sadiq et. al [44]	2020	SDI feature extraction	99.33
2	**Our Study**	**2025**	**NeuroNetV1**	**98.18**
3	Li et.al [39]	2023	MABLES	94.87
4	Yu et.al [40]	2022	CABLES	94
5	Huang et al. [41]	2022	EFD-CNN	93.81
6	Manoharan et.al [42]	2022	DWT+SVM+ANN	92
7	Sadiq et.al [43]	2020	Matrix determinant feature extraction	91.80
8	Yu et.al [45]	2021	IEFD+ Welch PSD+FFNN	88.08
9	Siuly et al. [46]	2017	PCA based RF Model	83.27
10	Tiwari et.al [47]	2022	MIDNN	82.48
11	Hashim et.al [48]	2021	LS-SVM	81.67

Table 19: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the P300-VI Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)
1	**Our Study**	**2025**	**NeuroNetV1**	**96.45**
2	Rabeya et.al [49]	2024	SVM	80.48
3	Bhattacharyya et.al [50]	2017	An online transferable BCI system	74
4	Tong et.al [51]	2016	A new approach of fusing multiple-channel features from temporal, spectral, and spatial domains through two times of dimensionality reduction based on neural network	78.18
5	Sowndhararajan et.al [52]	2018	xDAWN	78

Table 20: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the SCP-VII Dataset.

Sr. #	Authored By	Publication Year	Methodology Name	Average Accuracy (%)

1	Nazila et.al [53]	2023	FAM+SVM	99.83
2	Annaby et. al [54]	2019	Digraph Fourier transforms	96.58
3	Paranjape et.al [55]	2019	SVM+KNN	95
4	Hou et. al [56]	2018	V-SVM	94.50
5	**Our Study**	**2025**	**NeuroNetV1**	**94.20**
6	L. Duan et.al [57]	2016	PCA+LDA	94.20
7	Meena et.al [58]	2018	A preprocessing block for signal denoising of slow cortical potential (SCP)	94.10
8	Göksu et.al [59]	2018	Log Energy Entropy of wavelet packet analysis	92.80
9	Hou et. al [60]	2019	Concave convex feature	92.50
10	Duan et.al [61]	2017	KHELM	92
11	Yazici et.al [62]	2015	Time domain features + Nonlinear classifier	91.10

Table 21: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the Emot-VIII Dataset.

Sr. #	**Authored By**	**Publication Year**	**Methodology Name**	**Average Accuracy (%)**
1	Zhang et.al [63]	2021	SparseDGCNN	98.53
2	Pusarla et.al [64]	2022	LMD	98.00
3	Hou et. al [65]	2024	MECAM	98
4	Gu et.al [66]	2023	DGGN	97.28
5	Li et al. [67]	2023	GMSS	96.48
6	**Our Study**	**2025**	**NeuroNetV1**	**95.58**
7	Zhong et.al [68]	2022	RGNN	94.24
8	Li et al. [69]	2021	BiHDM	93.12
9	Li et al. [70]	2018	BiDANN	92.38
10	Wang et.al [71]	2020	DGCNN	92.27
11	Li et.al [72]	2022	Reduced channel + PSD feature	89.63
12	Zhang et.al [73]	2019	STRNN	89.50
13	Zheng et.al [74]	2015	DBN	86.08
14	Zheng et.al [75]	2017	GSCCA	82.96

Table 22: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the Sleep-IX Dataset.

Sr. #	**Authored By**	**Publication Year**	**Methodology Name**	**Average Accuracy (%)**
1	Liu et.al [76]	2024	GCN+ Transfomer	97.10
2	Xiao et.al [77]	2024	SPTESleepNet	96.60
3	**Our Study**	**2025**	**NeuroNetV1**	**95.21**
4	Duan et.al [78]	2025	MMS-SleepNet	92.90
5	Mao et.al [79]	2024	MVFSleepNet	90
6	Phan et.al [80]	2023	L-SeqSleepnet	88.60
7	Cong et.al [81]	2024	BiTS-SleepNet	88.50
8	Tsoi et.al [82]	2024	Positive-only approach	87.10
9	She et.al [83]	2024	CBLSNet	86.40
10	Shen et.al [84]	2023	LGSleepNet	86.00
11	Singh et.al [85]	2024	Efficient AttnSleep	85.80

Table 23: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the SZ-X Dataset.

Sr. #	**Authored By**	**Publication Year**	**Methodology Name**	**Average Accuracy (%)**
1	Siuly et. al [101]	2023	DeepResNet	99.23
2	**Our Study**	**2025**	**NeuroNetV1**	**98.99**
3	Li et.al [39]	2023	MABLES	95.28
4	Yu et.al [86]	2022	CABLES	92
5	Siuly et.al [87]	2020	EMD+IMF	89.59
6	Bose et.al [88]	2017	A modified odd ball-paradigms	88.50
7	Faizal et. al [89]	2023	CNN	86.93

Table 24: Extended Performance Evaluation of NeuroNetV1 Against Other Classification Methods on the EP-XIII Dataset.

Sr. #	**Authored By**	**Publication Year**	**Methodology Name**	**Average Accuracy (%)**
1	Parija et.al [90]	2024	CWCA-OVMD-CSAE-KELM	99
2	**Our Study**	**2025**	**NeuroNetV1**	**98.97**
3	Kumar et. al [91]	2017	Approximate Entropy, Reiny Entropy, Sample Entropy, Non-nested generalized exemplars classifier	98
4	Sharma et. al [92]	2017	TQWT, Entropy, LS-SVM	95
5	Gupta et. al [93]	2017	FAWT, entropy, LS-SVM	94.41
6	Sharma et. al [94]	2017	WFB, Entropy, LS-SVM	94.2
7	Sriraam et. al [95]	2017	statistical, frequency based, entropy, FD, SVM	92.1
8	Das et. al [96]	2016	EMD–DWT, entropy, KNN	89.40
9	Acharya et. a. [97]	2019	bi-spectrum, DFA, entropies, FD, Hjorth parameters, Hurst exponent Kolmogorov complexity, LLE, LZC, LS-SVM	87.93
10	Sharma et.al [98]	2015	Average of Entropies over IMF + LS-SVM	87
11	Bhattacharyya et. al [99]	2017	TQWT, entropy, LS-SVM	84.67

References:

[1] S. Chaudhary, S. Taran, V. Bajaj, and A. Sengur, ‘Convolutional neural network based approach towards motor imagery tasks EEG signals classification’, *IEEE Sensors Journal*, vol. 19, no. 12, pp. 4494–4500, 2019.

[2] M. T. Sadiq *et al.*, ‘Motor imagery EEG signals decoding by multivariate empirical wavelet transform-based framework for robust brain--computer interfaces’, *IEEE access*, vol. 7, pp. 171431–171451, 2019.

[3] S. Taheri, M. Ezoji, and S. M. Sakhaei, ‘Convolutional neural network based features for motor imagery EEG signals classification in brain--computer interface system’, *SN Applied Sciences*, vol. 2, pp. 1–12, 2020.

[4] A. Wijaya, T. B. Adji, and N. A. Setiawan, ‘Logistic Regression based Feature Selection and Two-Stage Detection for EEG based Motor Imagery Classification’, *International Journal of Intelligent Engineering & Systems*, vol. 14, no. 1, 2021.

[5] X. Zhu, M. Meng, Z. Yan, and Z. Luo, 'Motor Imagery EEG Classification Based on Multi-Domain Feature Rotation and Stacking Ensemble', *Brain Sciences*, vol. 15, no. 1, p. 50, 2025.

[6] Y. Zhang, G. Zhou, J. Jin, X. Wang, and A. Cichocki, 'Optimizing spatial patterns with sparse filter bands for motor-imagery based brain--computer interface', *Journal of neuroscience methods*, vol. 255, pp. 85–91, 2015.

[7] M. Meng, B. Xu, Y. Ma, Y. Gao, and Z. Luo, 'STGAT-CS: spatio-temporal-graph attention network based channel selection for MI-based BCI', *Cognitive Neurodynamics*, pp. 1–16, 2024.

[8] W. Liang, J. Jin, I. Daly, H. Sun, X. Wang, and A. Cichocki, 'Novel channel selection model based on graph convolutional network for motor imagery', *Cognitive Neurodynamics*, vol. 17, no. 5, pp. 1283–1296, 2023.

[9] M. Miao, W. Zhang, W. Hu, and R. Wang, 'An adaptive multi-domain feature joint optimization framework based on composite kernels and ant colony optimization for motor imagery EEG classification', *Biomedical Signal Processing and Control*, vol. 61, p. 101994, 2020.

[10] M. Miao, W. Hu, and W. Zhang, 'A spatial-frequency-temporal 3D convolutional neural network for motor imagery EEG signal classification', *Signal, Image and Video Processing*, vol. 15, no. 8, pp. 1797–1804, 2021.

[11] A. Hamidi and K. Kiani, 'Motor Imagery EEG signals classification using a Transformer-GCN approach', *Applied Soft Computing*, vol. 170, p. 112686, 2025.

[12] Y. Hou *et al.*, 'Deep feature mining via the attention-based bidirectional long short term memory graph convolutional neural network for human motor imagery recognition', *Frontiers in Bioengineering and Biotechnology*, vol. 9, p. 706229, 2022.

[13] W. Huang, G. Yan, W. Chang, Y. Zhang, and Y. Yuan, 'EEG-based classification combining Bayesian convolutional neural networks with recurrence plot for motor movement/imagery', *Pattern Recognition*, vol. 144, p. 109838, 2023.

[14] Y. Hou *et al.*, 'GCNs-net: a graph convolutional neural network approach for decoding time-resolved eeg motor imagery signals', *IEEE Transactions on Neural Networks and Learning Systems*, 2022.

[15] Y. Huang *et al.*, 'An improved model using convolutional sliding window-attention network for motor imagery EEG classification', *Frontiers in Neuroscience*, vol. 17, p. 1204385, 2023.

[16] C. Fan, B. Yang, X. Li, and P. Zan, 'Temporal-frequency-phase feature classification using 3D-convolutional neural networks for motor imagery and movement', *Frontiers in Neuroscience*, vol. 17, p. 1250991, 2023.

[17] H. Dose, J. S. Møller, H. K. Iversen, and S. Puthusserypady, 'An end-to-end deep learning approach to MI-EEG signal classification for BCIs', *Expert Systems with Applications*, vol. 114, pp. 532–542, 2018.

[18] R. R. Chowdhury, Y. Muhammad, and U. Adeel, ‘Enhancing cross-subject motor imagery classification in EEG-based brain--computer interfaces by using multi-branch CNN’, *Sensors*, vol. 23, no. 18, p. 7908, 2023.

[19] S. Moaveninejad *et al.*, ‘Fractal Dimension as a discriminative feature for high accuracy classification in motor imagery EEG-based brain-computer interface’, *Computer Methods and Programs in Biomedicine*, vol. 244, p. 107944, 2024.

[20] R. Lin *et al.*, ‘Motor imagery EEG task recognition using a nonlinear Granger causality feature extraction and an improved Salp swarm feature selection’, *Biomedical Signal Processing and Control*, vol. 88, p. 105626, 2024.

[21] Z. Wu, X. Tang, J. Wu, J. Huang, J. Shen, and H. Hong, ‘Portable deep-learning decoder for motor imaginary EEG signals based on a novel compact convolutional neural network incorporating spatial-attention mechanism’, *Medical & Biological Engineering & Computing*, vol. 61, no. 9, pp. 2391–2404, 2023.

[22] T. Grear and D. Jacobs, ‘Classifying EEG motor imagery signals using supervised projection pursuit for artefact removal’, in *2021 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2021, pp. 2952–2958.

[23] B. Huang, H. Xu, M. Yuan, M. Z. Aziz, and X. Yu, ‘Exploiting asymmetric EEG signals with EFD in deep learning domain for robust BCI’, *Symmetry*, vol. 14, no. 12, p. 2677, 2022.

[24] R. S. Chowdhury, S. Bose, S. Ghosh, and A. Konar, ‘Attention Induced Dual Convolutional-Capsule Network (AIDC-CN): A deep learning framework for motor imagery classification’, *Computers in Biology and Medicine*, vol. 183, p. 109260, 2024.

[25] D. Zhang, H. Li, and J. Xie, ‘Unsupervised and semi-supervised domain adaptation networks considering both global knowledge and prototype-based local class information for Motor Imagery Classification’, *Neural Networks*, vol. 179, p. 106497, 2024.

[26] X. Yu, M. Z. Aziz, Y. Hou, H. Li, J. Lv, and M. Jamil, ‘An extended computer aided diagnosis system for robust BCI applications’, in *2021 IEEE 9th International Conference on Information, Communication and Networks (ICICN)*, 2021, pp. 475–480.

[27] S. Kumar, A. Sharma, and T. Tsunoda, ‘Brain wave classification using long short-term memory network based OPTICAL predictor’, *Scientific reports*, vol. 9, no. 1, p. 9153, 2019.

[28] D. Park *et al.*, ‘Spatio-temporal explanation of 3D-EEGNet for motor imagery EEG classification using permutation and saliency’, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2023.

[29] M. Zheng and B. Yang, ‘A deep neural network with subdomain adaptation for motor imagery brain-computer interface’, *Medical Engineering & Physics*, vol. 96, pp. 29–40, 2021.

[30] L. Wang, M. Li, D. Xu, and Y. Yang, ‘Cortical ROI importance improves MI decoding from EEG using fused light neural network’, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2024.

[31] Z. Cai, T.-J. Luo, and X. Cao, ‘Multi-branch spatial-temporal-spectral convolutional neural networks for multi-task motor imagery EEG classification’, *Biomedical Signal Processing and Control*, vol. 93, p. 106156, 2024.

[32] C. Zhang, H. Chu, and M. Ma, ‘Decoding Algorithm of Motor Imagery Electroencephalogram Signal Based on CLRNet Network Model’, *Sensors*, vol. 23, no. 18, p. 7694, 2023.

[33] G. Yang and J. Liu, ‘A novel multi-scale fusion convolutional neural network for EEG-based motor imagery classification’, *Biomedical Signal Processing and Control*, vol. 96, p. 106645, 2024.

[34] S. Lian and Z. Li, ‘An end-to-end multi-task motor imagery EEG classification neural network based on dynamic fusion of spectral-temporal features’, *Computers in Biology and Medicine*, p. 108727, 2024.

[35] Y. Song, X. Ying, and J. Yang, ‘Transformer based on temporal-spatial feature learning for motor imagery electroencephalogram signal decoding’, *Journal of Nanjing University (Natural Science)*, vol. 59, no. 2, pp. 313–321, 2023.

[36] H. Gu, T. Chen, X. Ma, M. Zhang, Y. Sun, and J. Zhao, ‘CLTNet: A Hybrid Deep Learning Model for Motor Imagery Classification’, *Brain Sciences*, vol. 15, no. 2, p. 124, 2025.

[37] W. Zhao, X. Jiang, B. Zhang, S. Xiao, and S. Weng, ‘CTNet: a convolutional transformer network for EEG-based motor imagery classification’, *Scientific Reports*, vol. 14, no. 1, p. 20237, 2024.

[38] J. Zhao and M. Liu, ‘A deep temporal network for motor imagery classification based on multi-branch feature fusion and attention mechanism’, *Biomedical Signal Processing and Control*, vol. 100, p. 107163, 2025.

[39] H. Li, M. Z. Aziz, Y. Hou, and X. Yu, ‘An extended computer-aided diagnosis system for multidomain EEG classification’, in *Fifteenth International Conference on Machine Vision (ICMV 2022)*, 2023, vol. 12701, pp. 400–409.

[40] X. Yu, M. Z. Aziz, M. T. Sadiq, K. Jia, Z. Fan, and G. Xiao, ‘Computerized multidomain EEG classification system: A new paradigm’, *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3626–3637, 2022.

[41] B. Huang, H. Xu, M. Yuan, M. Z. Aziz, and X. Yu, ‘Exploiting asymmetric EEG signals with EFD in deep learning domain for robust BCI’, *Symmetry*, vol. 14, no. 12, p. 2677, 2022.

[42] H. Manoharan *et al.*, ‘A machine learning algorithm for classification of mental tasks’, *Computers and Electrical Engineering*, vol. 99, p. 107785, 2022.

[43] M. T. Sadiq, X. Yu, Z. Yuan, M. Z. Aziz, S. Siuly, and W. Ding, ‘A matrix determinant feature extraction approach for decoding motor and mental imagery EEG in subject-specific tasks’, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 14, no. 2, pp. 375–387, 2020.

[44] M. T. Sadiq, X. Yu, Z. Yuan, and M. Z. Aziz, ‘Identification of motor and mental imagery EEG in two and multiclass subject-dependent tasks using successive decomposition index’, *Sensors*, vol. 20, no. 18, p. 5283, 2020.

[45] X. Yu, M. Z. Aziz, M. T. Sadiq, Z. Fan, and G. Xiao, ‘A new framework for automatic detection of motor and mental imagery EEG signals for robust BCI systems’, *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–12, 2021.

[46] S. Siuly, R. Zarei, H. Wang, and Y. Zhang, 'A new data mining scheme for analysis of big brain signal data', in *Databases Theory and Applications: 28th Australasian Database Conference, ADC 2017, Brisbane, QLD, Australia, September 25--28, 2017, Proceedings 28*, 2017, pp. 151–164.

[47] S. Tiwari, S. Goel, and A. Bhardwaj, 'MIDNN-a classification approach for the EEG based motor imagery tasks using deep neural network', *Applied Intelligence*, pp. 1–20, 2022.

[48] N. A. Hashim and Z. S. M. Al-Husseini, 'Classification of brain signals based on proposed method'.

[49] R. Khatun, A. K. Sarkar, and A. Al Mamun, 'Error Detection in P300 Speller Device Applying Differential Entropy Features and Machine Learning Approaches', in *2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*, 2024, pp. 1157–1162.

[50] S. Bhattacharyya, A. Konar, D. N. Tibarewala, and M. Hayashibe, 'A generic transferable EEG decoder for online detection of error potential in target selection', *Frontiers in neuroscience*, vol. 11, p. 226, 2017.

[51] J. Tong, Q. Lin, R. Xiao, and L. Ding, 'Combining multiple features for error detection and its application in brain--computer interface', *Biomedical engineering online*, vol. 15, pp. 1–15, 2016.

[52] K. Sowndhararajan, M. Kim, P. Deepa, S. J. Park, and S. Kim, 'Application of the P300 event-related potential in the diagnosis of epilepsy disorder: a review', *Scientia pharmaceutica*, vol. 86, no. 2, p. 10, 2018.

[53] N. Panahi, M. C. Amirani, and M. Valizadeh, 'An EEG-Based Brain-Computer Interface Using Spectral Correlation Function', *IEEE Access*, vol. 11, pp. 33236–33247, 2023.

[54] M. H. Annaby, M. H. Said, A. M. Eldeib, and M. A. Rushdi, 'EEG-based motor imagery classification using digraph Fourier transforms and extreme learning machines', *Biomedical Signal Processing and Control*, vol. 69, p. 102831, 2021.

[55] P. N. Paranjape, M. M. Dhabu, P. S. Deshpande, and A. M. Kekre, 'Cross-correlation aided ensemble of classifiers for BCI oriented EEG study', *IEEE Access*, vol. 7, pp. 11985–11996, 2019.

[56] H.-R. Hou, Q.-H. Meng, M. Zeng, and B. Sun, 'Improving classification of slow cortical potential signals for BCI systems with polynomial fitting and voting support vector machine', *IEEE Signal Processing Letters*, vol. 25, no. 2, pp. 283–287, 2017.

[57] L. Duan, M. Bao, J. Miao, Y. Xu, and J. Chen, 'Classification based on multilayer extreme learning machine for motor imagery task from EEG signals', *Procedia Computer Science*, vol. 88, pp. 176–184, 2016.

[58] M. M. Makary, H. M. Bu-Omer, R. S. Soliman, K. Park, and Y. M. Kadah, 'Spectral subtraction denoising preprocessing block to improve slow cortical potential based brain--computer interface', *Journal of Medical and Biological Engineering*, vol. 38, pp. 87–98, 2018.

[59] H. Göksu, 'BCI oriented EEG analysis using log energy entropy of wavelet packets', *Biomedical Signal Processing and Control*, vol. 44, pp. 101–109, 2018.

[60] H. Hou, B. Sun, and Q. Meng, ‘Slow cortical potential signal classification using concave--convex feature’, *Journal of neuroscience methods*, vol. 324, p. 108303, 2019.
[61] L. Duan, M. Bao, S. Cui, Y. Qiao, and J. Miao, ‘Motor imagery EEG classification based on kernel hierarchical extreme learning machine’, *Cognitive Computation*, vol. 9, pp. 758–765, 2017.
[62] M. Yazıcı and M. Ulutaş, ‘Classification of EEG signals using time domain features’, in *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, 2015, pp. 2358–2361.
[63] G. Zhang, M. Yu, Y.-J. Liu, G. Zhao, D. Zhang, and W. Zheng, ‘SparseDGCNN: Recognizing emotion from multichannel EEG signals’, *IEEE Transactions on Affective Computing*, vol. 14, no. 1, pp. 537–548, 2021.
[64] N. Pusarla, A. Singh, and S. Tripathi, ‘Normal inverse Gaussian features for EEG-based automatic emotion recognition’, *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–11, 2022.
[65] F. Hou *et al.*, ‘MECAM: A novel multi-axis EEG channel attention model for emotion recognition’, *IEEE Transactions on Instrumentation and Measurement*, 2024.
[66] Y. Gu, X. Zhong, C. Qu, C. Liu, and B. Chen, ‘A domain generative graph network for EEG-based emotion recognition’, *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 5, pp. 2377–2386, 2023.
[67] Y. Li *et al.*, ‘GMSS: Graph-based multi-task self-supervised learning for EEG emotion recognition’, *IEEE Transactions on Affective Computing*, vol. 14, no. 3, pp. 2512–2525, 2022.
[68] P. Zhong, D. Wang, and C. Miao, ‘EEG-based emotion recognition using regularized graph neural networks’, *IEEE Transactions on Affective Computing*, vol. 13, no. 3, pp. 1290–1301, 2020.
[69] Y. Li *et al.*, ‘A novel bi-hemispheric discrepancy model for EEG emotion recognition’, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 13, no. 2, pp. 354–367, 2020.
[70] Y. Li, W. Zheng, Z. Cui, T. Zhang, and Y. Zong, ‘A novel neural network model based on cerebral hemispheric asymmetry for EEG emotion recognition’, in *IJCAI*, 2018, pp. 1561–1567.
[71] X.-H. Wang, T. Zhang, X.-M. Xu, L. Chen, X.-F. Xing, and C. L. P. Chen, ‘EEG emotion recognition using dynamical graph convolutional neural networks and broad learning system’, in *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2018, pp. 1240–1244.
[72] G. Li *et al.*, ‘An EEG data processing approach for emotion recognition’, *IEEE Sensors Journal*, vol. 22, no. 11, pp. 10751–10763, 2022.
[73] T. Zhang, W. Zheng, Z. Cui, Y. Zong, and Y. Li, ‘Spatial--temporal recurrent neural network for emotion recognition’, *IEEE transactions on cybernetics*, vol. 49, no. 3, pp. 839–847, 2018.

[74] W.-L. Zheng and B.-L. Lu, ‘Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks’, *IEEE Transactions on autonomous mental development*, vol. 7, no. 3, pp. 162–175, 2015.

[75] W. Zheng, ‘Multichannel EEG-based emotion recognition via group sparse canonical correlation analysis’, *IEEE Transactions on Cognitive and Developmental Systems*, vol. 9, no. 3, pp. 281–290, 2016.

[76] Y. Liu and X. Chen, ‘Automatic Sleep Staging Based on Data Augmentation with Diffusion Model and Transformer Network’, in *Proceedings of the 2024 9th International Conference on Cyber Security and Information Engineering*, 2024, pp. 591–595.

[77] X. Chen, X. Dai, X. Liu, and X. Chen, ‘SPTESleepNet: Automatic Sleep Staging Model Based On Strip Patch Embeddings And Transformer Encoder’, in *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2024, pp. 1951–1955.

[78] L. Duan, B. Ma, Y. Yin, Z. Huang, and Y. Qiao, ‘MMS-SleepNet: A knowledge-based multimodal and multiscale network for sleep staging’, *Biomedical Signal Processing and Control*, vol. 103, p. 107370, 2025.

[79] Y. Mao, X. Ma, H. Kuang, and X. Liu, ‘MVFSleepNet: Multi-View Fusion Network for Automatic Sleep Staging’, in *2024 IEEE 7th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, 2024, vol. 7, pp. 39–43.

[80] H. Phan *et al.*, ‘L-SeqSleepNet: Whole-cycle long sequence modelling for automatic sleep staging’, *IEEE Journal of Biomedical and Health Informatics*, 2023.

[81] Z. Cong *et al.*, ‘BiTS-SleepNet: An Attention-Based Two Stage Temporal-Spectral Fusion Model for Sleep Staging With Single-Channel EEG’, *IEEE Journal of Biomedical and Health Informatics*, 2024.

[82] P. Tsoi, Y.-S. Kweon, and S.-W. Lee, ‘Clustering-based Augmentation for Effective Self-supervised Learning in Sleep Staging’, in *2024 International Joint Conference on Neural Networks (IJCNN)*, 2024, pp. 1–8.

[83] Y. She, D. Zhang, J. Sun, X. Yang, X. Zeng, and W. Qin, ‘CBLSNet: A concise feature context fusion network for sleep staging’, *Biomedical Signal Processing and Control*, vol. 91, p. 106010, 2024.

[84] Q. Shen *et al.*, ‘LGSleepNet: an automatic sleep staging model based on local and global representation learning’, *IEEE Transactions on Instrumentation and Measurement*, 2023.

[85] M. Singh, S. Chauhan, A. K. Rajput, I. Verma, and A. K. Tiwari, ‘EASM: An efficient AttnSleep model for sleep Apnea detection from EEG signals’, *Multimedia Tools and Applications*, pp. 1–19, 2024.

[86] X. Yu, M. Z. Aziz, M. T. Sadiq, K. Jia, Z. Fan, and G. Xiao, ‘Computerized multidomain EEG classification system: A new paradigm’, *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 8, pp. 3626–3637, 2022.

[87] S. Siuly, S. K. Khare, V. Bajaj, H. Wang, and Y. Zhang, 'A computerized method for automatic detection of schizophrenia using EEG signals', *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 11, pp. 2390–2400, 2020.
[88] B. Thilakvathi, S. S. Devi, K. Bhanu, and M. Malaippan, 'EEG signal complexity analysis for schizophrenia during rest and mental activity', *Biomedical Research-India*, vol. 28, no. 1, pp. 1–9, 2017.
[89] M. F. Amri, N. Swastika, and M. Sadrawi, 'Diagnosis of Schizophrenia Patients in EEG Signals Using Multiple Optimizer-Convolutional Neural Network (CNN)', in *2023 International Conference on Computer, Control, Informatics and its Applications (IC3INA)*, 2023, pp. 166–170.
[90] S. Parija, P. K. Dash, and R. Bisoi, 'Epileptic EEG signal classification using an improved VMD-based convolutional stacked autoencoder', *Pattern Analysis and Applications*, vol. 27, no. 1, p. 9, 2024.
[91] N. Arunkumar *et al.*, 'Classification of focal and non focal EEG using entropies', *Pattern Recognition Letters*, vol. 94, pp. 112–117, 2017.
[92] R. Sharma, M. Kumar, R. B. Pachori, and U. R. Acharya, 'Decision support system for focal EEG signals using tunable-Q wavelet transform', *Journal of Computational Science*, vol. 20, pp. 52–60, 2017.
[93] V. Gupta, T. Priya, A. K. Yadav, R. B. Pachori, and U. R. Acharya, 'Automated detection of focal EEG signals using features extracted from flexible analytic wavelet transform', *Pattern Recognition Letters*, vol. 94, pp. 180–188, 2017.
[94] M. Sharma, A. Dhere, R. B. Pachori, and U. R. Acharya, 'An automatic detection of focal EEG signals using new class of time--frequency localized orthogonal wavelet filter banks', *Knowledge-Based Systems*, vol. 118, pp. 217–227, 2017.
[95] N. Sriraam and S. Raghu, 'Classification of focal and non focal epileptic seizures using multi-features and SVM classifier', *Journal of medical systems*, vol. 41, no. 10, p. 160, 2017.
[96] A. B. Das and M. I. H. Bhuiyan, 'Discrimination and classification of focal and non-focal EEG signals using entropy-based features in the EMD-DWT domain', *biomedical signal processing and control*, vol. 29, pp. 11–21, 2016.
[97] U. R. Acharya *et al.*, 'Characterization of focal EEG signals: A review', *Future Generation Computer Systems*, vol. 91, pp. 290–299, 2019.
[98] R. Sharma, R. B. Pachori, and U. R. Acharya, 'Application of entropy measures on intrinsic mode functions for the automated identification of focal electroencephalogram signals', *Entropy*, vol. 17, no. 2, pp. 669–691, 2014.
[99] A. Bhattacharyya, R. B. Pachori, and U. R. Acharya, 'Tunable-Q wavelet transform based multivariate sub-band fuzzy entropy with application to focal EEG signal analysis', *Entropy*, vol. 19, no. 3, p. 99, 2017.
[100] S. Phadikar, N. Sinha, and R. Ghosh, 'Unsupervised feature extraction with autoencoders for EEG based multiclass motor imagery BCI', Expert Systems with Applications, vol. 213, p. 118901, 2023.

[101] S. Siuly, Y. Guo, O. F. Alcin, Y. Li, P. Wen, and H. Wang, 'Exploring deep residual network based features for automatic schizophrenia detection from EEG', Physical and Engineering Sciences in Medicine, vol. 46, no. 2, pp. 561–574, 2023.