Undergraduate Conference on Intelligent Computing and Systems (UCICS 2025)
26-27 February, 2025; Varendra University, Rajshahi, Bangladesh

Comparative Study of DAE-Based CNN and CSP-SVM for Motor Imagery Classification

Kumar Shanu Sarkar
Dept. of Computer Science and Engineering
Varendra University
Rajshahi,Bangladesh
212311012@vu.edu.bd

Mst Malia Khatun
Dept. of Computer Science and Engineering
Varendra University
Rajshahi,Bangladesh
212311001@vu.edu.bd

Md. Golam Shahriar
Dept. of Computer Science and Engineering
Varendra University
Rajshahi,Bangladesh
shahriar@vu.edu.bd

Md Khademul Islam Molla, PhD
Dept. of Computer Science and Engineering
University of Rajshahi
Rajshahi,Bangladesh
khademul.cse@ru.ac.bd

***Abstract*— Brain-Computer Interfaces (BCIs) enable people with motor impairments to communicate and control their environment; nevertheless, the performance of electroencephalography (EEG) based motor imagery (MI) classification is still limited due to low signal-to-noise ratios, differences between individuals, and the complexity of datasets. This study investigates two methodologies, the classic Common Spatial Patterns (CSP) used with Support Vector Machines (SVM) and a deep learning approach that is composed of a Denoising Autoencoder (DAE) and a Convolutional Neural Network (CNN). We demonstrate these techniques applied to the BCI Competition IV 2a and 2b datasets. The results show that the DAE-CNN framework achieves an accuracy of 65.4% on the BCIC IV 2a dataset. While the CSP and SVM approach achieved 72.8% accuracy on the BCIC 2b dataset. This study provides meaningful insights for improving MI classification and paves the way for hybrid models that can increase BCI performance.**

Keywords— ***Motor Imagery (MI), Common Spatial Pattern (CSP), Support Vector Machine (SVM), Denoising Autoencoder (DAE), EEGNet, Convolutional Neural Network (CNN).***

I. Introduction

Brain-Computer Interfaces (BCIs) allow for direct communication between the brain and external devices, leading to remarkable opportunities for applications in neurorehabilitation and assistive technologies for people with severe motor disabilities, such as neuroprosthetics. Out of many techniques available, classifying motor imagery (MI) through EEG signals has drawn attention due to being non-invasive, portable and relatively inexpensive. However, owing to low signal-to-noise ratio, prone to artifacts and high dimensionality of data, EEG based MI systems face multiple challenges. The above-mentioned are challenges in making generalization and practical [1, 2, 3]. Conventional approaches such as Common Spatial Patterns (CSP) in combination with Support Vector Machines (SVM) have shown success in this area, directly using spatial features that are extracted to classify data with a robust decision boundary. But their reliance on manual feature engineering limits their adaptability with regards to noisy and complex datasets, particularly in multi-class or few channels' problems [4, 5]. There are now exciting alternatives in deep learning that have emerged more recently. On the other hand, CNNs are capable of discovering spatial-temporal features from raw EEG data independently, and Denoising Autoencoders (DAE) enhance denoising during feature extraction [6,7]. Deep learning models hold the potential to achieve incredible results, but in order to do so they require a significant amount of computational power, preprocessing and large datasets to build optimal performance. This work thoroughly compares classical CSP-SVM and recent DAE-CNN paradigms, utilizing data from BCI Competition IV 2a and 2b, each of them poses distinct challenges in terms of complexity and channels. The 2a type dataset is a multi-class classification with 22 channels and the 2b type dataset is a binary classification with just three bipolar channels to measure computational efficiency. The evaluation of these methodologies provides important insights into the trade-off between computational efficiency and reliable feature learning, suggesting the possibility of hybrid approaches. Bridging the interpretability and efficiency of CSP with the powerful feature extraction abilities of neural networks could significantly enhance the real-world applicability of BCI systems. In conclusion, the results of this work contribute to the state-of-the-art of EEG-based MI classification by highlighting the merits and disadvantages of both classical and modern approaches. These changes may dramatically enhance the performance of BCI systems for real-world applications such as neurorehabilitation, adaptive prosthetics, and assistive devices.

II. Related Works

The performance of motor imagery (MI) classification based on EEG is critical to the success of Brain-Computer Interface (BCI) systems, but facing challenges including low signal-to-noise ratio (SNR), high inter-subject variability, and high data dimensionality. Traditional methods (such as the combination of Common Spatial Patterns (CSP) with Support Vector Machines (SVM) or Linear Discriminant Analysis (LDA)) are efficient in terms of computations and have straightforward interpretations in binary classification problems. However, they usually poor in noisy environments, high dimensions or multi-class datasets, such as BCI Competition IV 2a [12, 13, 14, 15]. Despite all these challenges, CSP-SVM frameworks have the potentiality and are relatively easy to implement, so still work good for resource limitation systems. On the other hand, the surge of deep learning methods also fundamentally reshaped MI classification algorithms by learning spatial-temporal features from raw EEG data using such techniques CNNs and DAEs. EEGNet. However, deep learning approaches need vast amounts of labeled data, significant computational power, and extensive preprocessing before training, while their ability to generalize between data.. CSP-SVMs provide a mathematically tractable framework for subject-general analyses but can struggle to capture complex patterns, while CNNs and DAEs outperform in terms of pattern recognition but may struggle with the trade-off between performance and generalization. It was proposed that combine the interpretability of CSP and the feature extraction advantages of deep learning could be a good strategy [20, 21]. Adaptation of systematic assessment of these approaches across datasets of varied complexities and scalability of real-world applications are major areas that would require further evaluation, reinforcing the need to augment MI classification methodologies for successful BCIs.

III. Materials And Methods

To test the proposed method, BCI Competition IV 2a and 2b datasets have been taken into account. An 2b dataset has 2 MI tasks with 3 bipolar channels and from 2a dataset there are 9 subjects performing 4 MI tasks for 22 channels of EEG recordings. The two datasets were both preprocessed via baseline correction, normalization, and segmenting into overlapping sliding windows to increase the temporal resolution. All datasets were then randomly partitioned into stratified sampling training (70%), validation (15%), and test sets (15%) with the goal of retaining balanced class representation. The 2a dataset consisted of images with high dimensionalities, and 2b low-channel efficiency images.

A. Data Description

The BCI Competition IV 2a[22] dataset is a well-known benchmark dataset for motor imagery (MI) classification, which consists of EEG recordings from 9 subjects performing 4 different MI tasks (left hand, right hand, feet, tongue), using 22 electrodes according to the 10-20 system, at 250 Hz sampling rate filtered between 0.5–100 Hz. Each subject performed 288 trials (72 for each task), where visual cues were followed by 2 seconds of MI in a trial, providing very variable and noisy data to test MI classification methods. The 2b dataset[23] contains binary MI tasks (left and right hand) recorded by 3 bipolar electrodes (C3, Cz, C4) with a sample rate of 250 Hz, along with EOG channels for artifact removal. It consists of six runs of 90 trials (45 for each task) each.

B. Signal Preprocessing

The preprocessing pipeline was mainly designed to enhance the quality of the EEG signal. The pipeline can be described as the following figure:

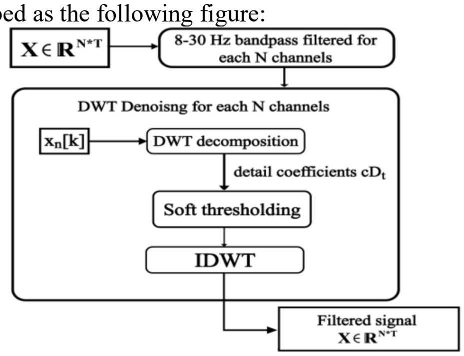

Fig. 1. Preprocessing Pipeline

The raw EEG signal is denoted $X \in R^{N \times T}$, where N is the number of channels and T is the number of time samples. At first the signal X is bandpass filtered with $f_{LOW} = 8.0$ and $f_{HIGH} = 30.0$ Hz using Butterworth filter. This is done for each channel. The filtered signal is then decomposed using discrete wavelet transform using *DB2* wavelet basis and 4 levels of decomposition. A threshold value is computed as following. At first the variance parameter is computed as

$$\sigma = \frac{1}{N}\sum_{k=1}^{N} |d_{lk}| \ for\ all\ l \tag{1}$$

In the equation above, d_{lk} denotes the details coefficients of level l. Then the thresholding value β is computed as:

$$\beta = \alpha\sigma\sqrt{2\ log\ log\ T} \tag{2}$$

Where α is the strength of thresholding.
The preprocessed detail coefficients are computed using a soft thresholding technique as:

$$d_{lk} = \frac{d_{lk}}{|d_{lk}| \times max(|d_{lk}| - \beta, 0)} \tag{3}$$

Finally, each channel k of EEG trial j is whitened as

$$x_{jk}[n] = \frac{x_{jk}[n] - \mu_{jk}}{\sigma_{jk}^2} \tag{4}$$

C. Classification

This study utilized two different classification methods: a DAE-CNN framework focused on harnessing deep learning for effective feature extraction and classification, and a CSP-SVM technique, which represents a conventional machine learning approach frequently applied in motor imagery (MI) classification.

1) DAE-CNN Framework:

An autoencoder is an unsupervised machine learning technique that aims to compute a compressed representation of an input signal. An autoencoder tries to compress the signal in such a way that will help reconstruct the original signal with minimum mean squared error loss. In this process, the compressed representation can be used as a good latent space feature for the original signal. The idea can be visualized as following:

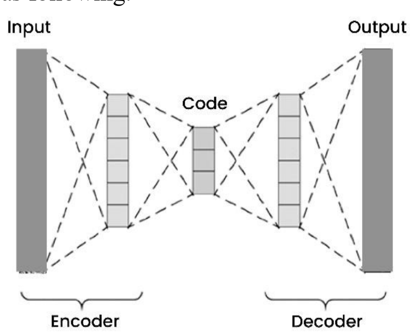

Fig. 2. DAE Architecture

Let us denote a high dimensional input signal as $x \in R^m$, the target is to compute a lower dimensional signal $x \in R^n$ where $n \ll m$. Let us denote the encoder parameterized by ϕ as

$$z = E_\phi(x) \tag{5}$$

And the decoder as

$$x' = D_\theta(z) \tag{6}$$

For example, the encoder can be a multilayer perceptron

$$E_{\phi}(x) = \sigma(Wx + b) \quad (7)$$

Finally, the training objective is

$$\frac{1}{N}\sum_{i=1}^{N} \|x_i - D_{\theta}(E\phi\,(x_i))\|^2 \qquad (8)$$

The denoising is achieved by deliberately injecting noise to the input data during training as:

$$x_i = x_i + \epsilon \qquad (9)$$
$$\epsilon \sim N(0, 1)$$

In this study, the particular architecture of the denoising autoencoder based CNN is as follows:

TABLE I. PROPOSED DAE BASED CNN ARCHITECTURE

Layer	Layer description	Output size	Connected to
Input layer	EEG data with dimensions (22, 256)	(22, 256)	—
1st DAE Layer	Applies denoising to the input EEG signal.	(22, 256)	Input Layer
2nd DAE Layer	Learns latent representations by reducing noise further.	(11, 128)	1st DAE Layer
3rd DAE Layer	Final denoised output with compressed and clean features.	(5, 64)	2nd DAE Layer
1D Convolution Layer	Learns spatial-temporal features from the denoised output.	(8, 64)	3rd DAE Layer
Pooling Layer	Reduces temporal resolution to focus on dominant features.	(8, 32)	1D Convolution Layer
Dense Layer	Fully connected layer for classification.	(4)	Pooling Layer
Softmax Layer	Outputs class probabilities for classification tasks.	(4)	Dense Layer

Overall, the workflow is the following:

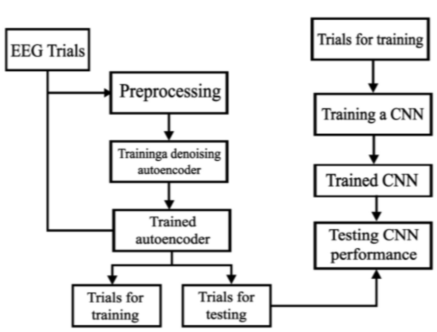

Fig. 3. Workflow Diagram

2) CSP-SVM Method:

The CSP-SVM method employs the well-established Common Spatial Patterns (CSP) algorithm for feature extraction. CSP applies spatial filters W on the raw EEG signals X to output discriminative components: $Z = W^T X$, where Z is the spatially filtered signals. The spatial filter matrix W is computed as

$$w = \arg\min_{w} \frac{w^T \Sigma_1 w}{w^T \Sigma_2 w} \qquad (10)$$

Where

$$\Sigma_c = X_c X_c^T \; for \; c \in \{1, 2\} \qquad (11)$$

A spatial filter matrix $W = [w_1\, w_2\, \dots\, w_{2m}\,] \in R^{N \times 2M}$.For each EEG trial X, the feature vector $z \in R^{2M}$ is computed as

$$z_m = log\,(var(w_m^T X)) \qquad (12)$$

In particular, six CSP features were chosen for the BCI IV 2a dataset because of its high channel number, and only three features worked well for the lower spatial resolution 2b dataset [23].

After all trial were spatially filtered, a kernel SVM with radial basis function kernel was used to perform classification according to the following flow diagram:

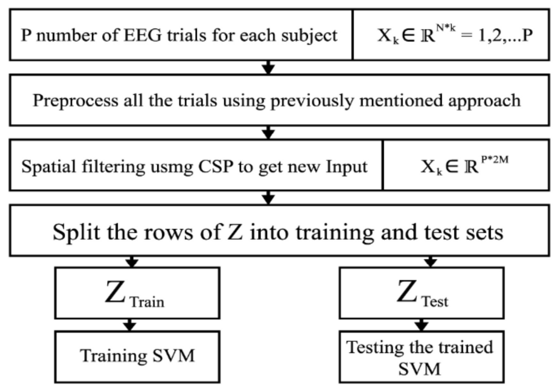

Fig. 4. Classification Diagram

IV. Result And Discussion

DAE-CNN and CSP-SVM were evaluated using five-fold cross-validation, with DAE-CNN leveraging data augmentation for complex feature learning and CSP-SVM prioritizing computational efficiency for binary classification tasks where

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (13)$$

$$Precision = \frac{TP}{TP + FT} \qquad (14)$$

$$Recall = \frac{TP}{TP + FN} \qquad (15)$$

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \qquad (16)$$

The performance of the DAE-CNN and CSP-SVM models were compared with different methodology shown in Table 2. Their DAE-CNN outperformed the CSP-SVM in the multi-class BCI IV 2a dataset with 22 channels and four MI classes, achieving an accuracy of 65.4%, 4.2% better that CSP-SVM's task with 22 channels and 4 MI classes, obtaining an accuracy of 61.2%. However, both models exhibited difficulties with inter-subject variability and complexity of the dataset, suggesting that a performance optimization is necessary to improve multi-class high-dimensional classification performance. In contrast, CSP-SVM out-performed on the binary-class BCI IV 2b dataset, with only three channels of data, achieving accuracy of 72.8% against DAE-CNN's result of 70.3%. It shows CSP-SVM is superior in simple low channel conditions because of its efficient extracting and discrimination capabilities of the features. Additionally, cross-validation results confirmed that CSP-SVM is more consistent in the 2b dataset, and DAE-CNN possesses greater potential to process higher order, 2D data in future applications.

TABLE II. COMPARISON OF RESULTS

Paper	Method	Accuracy
R. Zhang et al. [24]	CSP and SVM	58.2% - 61.2%.
Alimardani et al. [25]	MIN2Net, EEGNet and DeepConvNet	51.7% - 62.5%
S. S. Mohseni Salehi et al. [26]	HCSP	64.5%
This work	**CSP-SVM and DAE-CNN**	**61.2% - 72.8%**

The key results emerged in this latter, two-dimensional setting but highlight the strengths of DAE-CNN over CSP-SVM. In the future, one possible research direction is to study hybrid models that blend the strengths of both models.

V. CONCLUSION

The following study was performed comparing DAE-CNN and CSP-SVM for motor imagery classification on BCI Competition IV datasets. DAE-CNN was limited in the results that it achieved, reaching 65.4% accuracy on the high-dimensional 2a dataset where as CSP-SVM outperformed, achieving 72.8% accuracy on the simpler 2b dataset. However, there are still limitations for both models, the performance needs to be improved. Hybrid models, high-level preprocessing of the data, and domain adaptation are all possible future avenues of exploration to improve performance. Similarly, testing on diverse datasets from various subjects would increase generalizability and find applicability in the real world.

REFERENCES

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan, "Brain–computer interfaces for communication and control," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 10, no. 2, pp. 167–185, 2002. doi: 10.1109/TNSRE.2002.1022070.

[2] N. Birbaumer and L. G. Cohen, "Brain–computer interfaces: Communication and restoration of movement in paralysis," Journal of Physiology, vol. 579, no. 3, pp. 621–636, 2007. doi: 10.1113/jphysiol.2006.125948.

[3] B. Blankertz, G. Curio, and K. R. Müller, "Classifying single trial EEG: Towards brain computer interfacing," in Advances in Neural Information Processing Systems (NIPS), vol. 14, 2002.

[4] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," IEEE Transactions on Biomedical Engineering, vol. 47, no. 4, pp. 847–856, 2000. doi: 10.1109/10.847807.

[5] K. F. LaFleur, K. Cassady, T. Doud, K. Shades, E. Rogin, and B. He, "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," Journal of Neural Engineering, vol. 10, no. 4, p. 046003, 2013. doi: 10.1088/1741-2560/10/4/046003.

[6] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," IEEE Transactions on Biomedical Engineering, vol. 65, no. 9, pp. 2228–2238, 2018. doi: 10.1109/TBME.2017.2771300.

[7] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in Proceedings of the 25th International Conference on Machine Learning (ICML), pp. 1096–1103, 2008.

[8] M. Tangermann et al., "Review of the BCI Competition IV," Frontiers in Neuroscience, vol. 6, p. 55, 2012. doi: 10.3389/fnins.2012.00055

[9] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: A general-purpose brain–computer interface (BCI) system," IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, pp. 1034–1043, 2004. doi: 10.1109/TBME.2004.827072.

[10] S. G. Mason and G. E. Birch, "A general framework for brain–computer interface design," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 1, pp. 70–85, 2003. doi: 10.1109/TNSRE.2003.814481.

[11] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," Proceedings of the IEEE, vol. 89, no. 7, pp. 1123–1134, 2001. doi: 10.1109/5.939829.

[12] H. Ramoser, J. Müller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," IEEE Transactions on Biomedical Engineering, vol. 47, no. 4, pp. 847–856, 2000. doi: 10.1109/TBME.2000.843506.

[13] S. G. Mason and G. E. Birch, "A general framework for brain–computer interface design," IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 11, no. 1, pp. 70–85, 2003. doi: 10.1109/TNSRE.2003.814481.

[14] B. Blankertz, G. Curio, and K. R. Müller, "Classifying single trial EEG: Towards brain computer interfacing," in Advances in Neural Information Processing Systems (NIPS), vol. 14, 2002.

[15] G. Schalk et al., "BCI2000: A general-purpose brain–computer interface (BCI) system," IEEE Transactions on Biomedical Engineering, vol. 51, no. 6, pp. 1034–1043, 2004. doi: 10.1109/TBME.2004.827072.

[16] V. J. Lawhern et al., "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," IEEE Transactions on Biomedical Engineering, vol. 65, no. 9, pp. 2228–2238, 2018. doi: 10.1109/TBME.2017.2771300.

[17] K. F. LaFleur et al., "Quadcopter control in three-dimensional space using a noninvasive motor imagery-based brain–computer interface," Journal of Neural Engineering, vol. 10, no. 4, p. 046003, 2013. doi: 10.1088/1741-2560/10/4/046003.

[18] P. Vincent, H. Larochelle, Y. Bengio, and P. A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in Proc. 25th International Conference on Machine Learning (ICML), pp. 1096–1103, 2008.

[19] M. Tangermann et al., "Review of the BCI Competition IV," Frontiers in Neuroscience, vol. 6, p. 55, 2012. doi: 10.3389/fnins.2012.00055.

[20] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," Proceedings of the IEEE, vol. 89, no. 7, pp. 1123–1134, 2001. doi: 10.1109/5.939829.

[21] N. Birbaumer and L. G. Cohen, "Brain–computer interfaces: Communication and restoration of movement in paralysis," Journal of Physiology, vol. 579, no. 3, pp. 621–636, 2007. doi: 10.1113/jphysiol.2006.125948.

[22] Brunner, R. Leeb, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller, “BCI Competition 2008 – Graz data set A,” Institute for Knowledge Discovery, Graz University of Technology, Austria, 2008.

[23] R. Leeb, C. Brunner, G. R. Müller-Putz, A. Schlögl, and G. Pfurtscheller, “BCI Competition 2008 – Graz data set B,” Institute for Knowledge Discovery, Graz University of Technology, Austria, 2008.

[24] R. Zhang et al., "Using Brain Network Features to Increase the Classification Accuracy of MI-BCI Inefficiency Subject," in IEEE Access, vol. 7, pp. 74490-74499, 2019, doi: 10.1109/ACCESS.2019.2917327.

[25] Alimardani, Maryam & Kocken, Steven & Leeuwis, Nikki. (2023). End-to-End Deep Transfer Learning for Calibration-free Motor Imagery Brain Computer Interfaces. 10.48550/arXiv.2307.12827.

[26] S. S. Mohseni Salehi, M. Moghadamfalahi, F. Quivira, A. Piers, H. Nezamfar and D. Erdogmus, "Decoding complex imagery hand gestures," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jeju, Korea (South), 2017, pp. 2968-2971, doi: 10.1109/EMBC.2017.8037480.