DOI: doi.org/10.58797/cser.040104

A Subject-Independent Comparison of EEGNet and CSP+LDA on EEGMMIDB: Implications for Assistive Educational Technology

Nur Aziezah[1,*], Fajar Nur Hamzah[2], Walidatush Sholihah[3]

[1]*Software Engineering Technology Study Program, Vocational School, IPB University, Jl. Kumbang, No. 14 Bogor 16128, Indonesia*

[2]*Faculty of Science, University of Sydney, Eastern Ave, Camperdown NSW 2050, Australia*

[3]*Faculty of Science and Engineering, University of Groningen, Broerstraat 5, 9712 CP Groningen, The Netherlands*

*Corresponding Author Email: nuraziezah@apps.ipb.ac.id

Received: 25 January 2026
Revised: 16 March 2026
Accepted: 25 March 2026
Published: 9 April 2026
Issued: 30 April 2026

Current Steam and Education Research
e-ISSN: 3025-8529

Abstract

Motor imagery (MI) electroencephalography (EEG) classification remains difficult in cross-subject settings because signals vary substantially between individuals. This study compares a compact deep learning model (EEGNet) with a classical Common Spatial Pattern and Linear Discriminant Analysis pipeline (CSP+LDA) for hands-versus-feet MI classification. Data came from the EEG Motor Movement/Imagery Dataset (EEGMMIDB) on PhysioNet. A leave-one-subject-out (LOSO) evaluation was conducted across 109 subjects using runs 6, 10, and 14. Preprocessing applied a 7-30 Hz band-pass filter, a 1-2 s post-cue window, and subject-wise z-score normalization. Because normalization used unlabeled statistics from each held-out subject, the setting is LOSO with unsupervised test-time normalization rather than a fully inductive calibration-free protocol. EEGNet reached a mean balanced accuracy of 0.660 ± 0.136 (macro-F1 0.640 ± 0.147; Cohen's kappa 0.312 ± 0.271), whereas CSP+LDA reached 0.635 ± 0.141 (macro-F1 0.605 ± 0.164; kappa 0.270 ± 0.281). A descriptive paired comparison showed that EEGNet scored higher for 60 of 109 subjects, CSP+LDA scored higher for 47 subjects, and 2 subjects were tied. The aggregated EEGNet confusion matrix showed higher recall for hands than for feet (0.735 vs 0.574), and balanced accuracy varied widely across subjects. These results provide a reproducible subject-independent comparison on EEGMMIDB and show that inter-subject variability remains a major barrier to robust MI decoding. The reproducible pipeline also offers a practical neuroinformatics case for teaching biosignal preprocessing, leakage-aware validation, and model comparison, while the subject-level variability shows what must be addressed before assistive educational deployment.

Keywords: CSP+LDA, EEGMMIDB, EEGNet, leaveone-subject-out (LOSO), motor imagery EEG

INTRODUCTION

Motor imagery (MI) electroencephalography (EEG) decoding is central to brain–computer interface (BCI) systems that translate brain activity into control signals for communication, rehabilitation, and assistive interaction (Wolpaw et al., 2002). A major barrier to practical use is generalization to a new user whose labeled EEG data were not available during training. Cross-subject variability therefore remains a core challenge for subject-independent MI decoding (Kwon et al., 2020; Wu et al., 2022).

From a neurophysiological perspective, motor imagery modulates sensorimotor rhythms in the mu/alpha and beta bands. These changes are commonly described as event-related desynchronization or synchronization (ERD/ERS), which are frequency-specific decreases or increases in oscillatory activity associated with cortical processing (Pfurtscheller & Lopes da Silva, 1999). This neurophysiological basis motivates the use of mu–beta information in MI pipelines, but low signal-to-noise ratios and large differences between individuals still limit subject-independent decoding.

CSP+LDA is a classic, well-trusted method for classifying motor-imagery EEG signals. Common Spatial Patterns (CSP) finds spatial filters that highlight the brainwave patterns linked to each class, and Linear Discriminant Analysis (LDA) is a simple, lightweight classifier used on top (Blankertz et al., 2008; Lotte et al., 2018). Filter-bank extensions further show that CSP performance depends on the selected frequency ranges (Ang et al., 2012). Performance also depends on epoch definition, normalization, feature estimation, and validation design. Consequently, a fair comparison requires both methods to share the same data, preprocessing, and held-out-subject protocol (Borra et al., 2025; Wang et al., 2024).

Deep learning can reduce dependence on handcrafted EEG features. Convolutional approaches can learn temporal and spatial representations directly from EEG, although their performance is sensitive to architecture, preprocessing, and validation choices (Craik et al., 2019; Roy et al., 2019; Schirrmeister et al., 2017). EEGNet, introduced by Lawhern et al. (2018), uses temporal, depthwise, and separable convolutions to provide a compact model with relatively few parameters. Other convolutional designs have also shown the value of learning temporal and spectral information from MI signals (Sakhavi et al., 2018). However, strong performance in subject-dependent or mixed-subject settings does not establish generalization to an unseen person. Cross-subject and cross-dataset variability can substantially change model performance (Kwon et al., 2020; Xu et al., 2020), and EEG pipelines are sensitive to hyperparameter selection and random initialization (Borra et al., 2025).

Existing MI-EEG studies use heterogeneous datasets, preprocessing pipelines, model families, and validation schemes, which makes reported gains difficult to compare directly (Craik et al., 2019; Roy et al., 2019; Wang et al., 2024). Recent models include multi-task subject-independent learning, filter-bank multi-scale convolution, and convolutional transformers (Autthasan et al., 2022; Liu et al., 2023; Song et al., 2023). These new methods give us more models to choose from, but a fair comparison is only possible if the people used for testing and the way the data is processed stay separate from what the model trains on.

To test how well models work on people they were never trained on, researchers have evaluated them on individuals left out of training, and have also tried compact networks, transfer learning, and domain-generalization strategies (Keutayeva et al., 2024; Kwon et al., 2020; Liang et al., 2023; Zhang et al., 2021). Even so, it is still useful to compare EEGNet and CSP+LDA side by side on the same large public dataset, using the same pipeline and one held-out subject per fold, because this helps separate which model has a higher average score from which one works more consistently across users.

To fill this gap, we used the EEG Motor Movement/Imagery Dataset (EEGMMIDB), a public PhysioNet resource with recordings from 109 volunteers. The data were collected with the BCI2000 system using 64 EEG channels at 160 Hz, and each participant completed 14 runs covering baseline, motor execution, and motor imagery tasks (Goldberger et al., 2000; Schalk et al., 2004). This study focused on the bilateral motor imagery runs, where T1 means imagining moving both fists and T2 means imagining moving both feet.

Beside brain–computer interface research, classifying EEG in a subject-independent way could one day support adaptive and assistive educational technologies. Researchers have already used EEG in education, for example to measure cognitive load and to explore portable neurotechnology applications (Antonenko et al., 2010; Xu & Zhong, 2018). Educational neurotechnology also includes wider EEG-, BCI-, and AI-supported systems, but using them responsibly requires solid science, proven educational value, and care for how users differ (Nouri, 2025). A model that works well for only some users could give people unequal access or uneven interaction quality. The educational value of this study is therefore about the future: it checks a technical requirement for possible future interfaces, it does not measure learning outcomes or test an educational intervention.

Accordingly, this study addresses the following research question: ‘to what extent do EEGNet and CSP+LDA differ in subject-independent motor imagery classification performance and consistency across subjects under a leave-one-subject-out evaluation on EEGMMIDB.’ The comparison focuses on average performance, class-wise behaviour, and variability across held-out users.

METHOD

Dataset and Task Definition

This study used the public EEG Motor Movement/Imagery Dataset (EEGMMIDB) hosted on PhysioNet (Goldberger et al., 2000). Data were loaded through MNE-Python's eegbci loader; MNE provides an open-source framework for reproducible EEG and MEG processing (Gramfort et al., 2014). The loader downloads the original EDF recordings and annotations. The selected motor imagery condition asked participants to imagine both fists versus both feet and corresponds to EEGBCI runs 6, 10, and 14 (Schalk et al., 2004). T1 was mapped to the hands class and T2 to the feet class. Rest segments were excluded from supervised learning.

EEG Preprocessing and Epoching

For each subject, EDF files from the selected runs were concatenated, standardized with MNE-Python's EEG routines, and mapped to the standard_1005 montage. Signals were re-referenced to the average reference, band-pass filtered from 7 to 30 Hz, and resampled to 128 Hz. Epochs were extracted from 1.0 to 2.0 s after each T1 or T2 cue. The resulting trial tensors had the form trials × channels × time, and the same channel set, filtering, resampling, epoch window, trial inclusion criteria, and class definitions were used for EEGNet and CSP+LDA.

Subject-Wise Z-Score Standardization

To reduce differences in amplitude scale between subjects and to stabilize training, we applied subject-wise z-score normalization per EEG channel. This step used only each subject's own unlabeled data; no label information was used. For a subject's trial tensor $X \in R^{(N \times C \times T)}$, where N is the number of trials, C is the number of channels, and T is the number of time points, the normalization was computed channel-wise over trials and time as shown in Equations (1)-(3).

$$\mu_c = (1 / NT) \cdot \Sigma_{n=1..N} \Sigma_{t=1..T} X_{n,c,t} \quad (1)$$

$$\sigma_c = \sqrt{[\, (1 / NT) \cdot \Sigma_{n=1..N} \Sigma_{t=1..T} (X_{n,c,t} - \mu_c)^2 + \varepsilon \,]}, \quad \varepsilon = 10^{-6} \quad (2)$$

$$\tilde{X}_{n,c,t} = (X_{n,c,t} - \mu_c) / \sigma_c \quad (3)$$

This standardization used the held-out subject's unlabeled distribution. No test labels were used; however, the procedure constitutes unsupervised test-time normalization. The results should therefore be interpreted as LOSO evaluation with unsupervised test-time normalization, not as a fully inductive calibration-free setting. Model training and supervised feature estimation still excluded the held-out subject's labels and trials.

Deep Learning Model: EEGNet

We implemented EEGNet, the compact convolutional neural network introduced by Lawhern et al. (2018), for end-to-end MI-EEG decoding. The network uses temporal convolution, depthwise spatial filtering, and separable convolutions. Inputs were reshaped to batch × 1 × channels × time. The reported implementation used F1 = 8, depth multiplier D = 2, F2 = 16, dropout = 0.5, and a temporal kernel length of 63 samples at 128 Hz. Training used a batch size of 64, Adam with a learning rate of 10^{-3}, weight decay of 10^{-4}, and early stopping with a patience of 12 epochs.

Classical Baseline: CSP + LDA

As a classical baseline, we implemented Common Spatial Patterns (CSP) followed by Linear Discriminant Analysis (LDA). CSP finds spatial filters that make the variance as large as possible for one motor-imagery class and as small as possible for the other; the log of these variances then serves as the features for classification (Lotte et al., 2018). The covariance matrices for each class were regularized before CSP was applied, and LDA, as a simple linear classifier, made the final decision. CSP and LDA were given exactly the same inputs and the same one-subject-out folds as EEGNet, so any difference in results came from the models themselves, not from different preprocessing or different test subjects.

Evaluation Protocol: Leave-One-Subject-Out (LOSO)

We used leave-one-subject-out (LOSO) evaluation to measure cross-subject generalization. Across 109 folds, one participant was held out for testing and the remaining 108 participants formed the training pool. This split prevents trials from the same person from appearing in both the training and test sets, a requirement for evaluating a classifier on an unseen user (Keutayeva et al., 2024; Kwon et al., 2020). Holding out every subject in turn approximates the practical requirement that a deployed classifier operate on a user whose labeled EEG data were not included during model training. Performance was reported for every held-out subject and summarized across all 109 subjects as mean ± standard deviation. Because test-subject normalization statistics were used, this is a transductive LOSO-z setting rather than a fully inductive LOSO setting.

Metrics and Paired Comparison

We evaluated performance per LOSO fold using accuracy, balanced accuracy, macro-averaged F1, and Cohen's kappa. We also report an aggregated confusion matrix obtained by summing the per-fold confusion matrices across all 109 folds. Let TP, TN, FP, and FN denote the entries of the confusion matrix for the binary hands-versus-feet task. Accuracy, balanced accuracy, class-wise precision (P_k) and recall (R_k), per-class F1 ($F1_k$), macro-F1, and Cohen's kappa (κ) were computed as shown in Equations (4)-(9).

$$Acc = (TP + TN) / (TP + TN + FP + FN) \tag{4}$$

$$bAcc = ½ \cdot [\ TP / (TP + FN)\ +\ TN / (TN + FP)\ ] \tag{5}$$

$$Pk = TPk / (TPk + FPk) \qquad Rk = TPk / (TPk + FNk) \tag{6}$$

$$F1k = 2 \cdot Pk \cdot Rk / (Pk + Rk) \tag{7}$$

$$Macro\text{-}F1 = (1 / K) \cdot \Sigma k=1..K\ F1k, \quad K = 2 \tag{8}$$

$$\kappa = (po - pe) / (1 - pe) \tag{9}$$

Balanced accuracy was computed as the mean of class-wise recalls, macro-F1 as the unweighted mean of the per-class F1 scores, and Cohen's kappa as agreement beyond chance. The paired subject-level comparison was summarized descriptively using the mean and median paired differences and the numbers of EEGNet wins, CSP+LDA wins, and ties.

Implementation Environment

All experiments ran in Google Colab. We used Python with MNE-Python for data loading and preprocessing, PyTorch for training EEGNet, and standard scientific libraries for evaluation and visualization. This environment matches the pipeline used to generate the reported results and figures.

RESULTS AND DISCUSSION

Overall LOSO Performance on EEGMMIDB

Using EEGMMIDB, we evaluated hands-versus-feet motor imagery under a leave-one-subject-out protocol across 109 subjects, using runs 6, 10, and 14 with subject-wise z-score normalization. Table 1 summarises the overall LOSO performance of EEGNet and the CSP+LDA baseline across all subjects.

Table 1. LOSO subject-wise z-score results (109 subjects)

Model	Accuracy (mean ± SD)	Balanced Accuracy (mean ± SD)	Macro-F1 (mean ± SD)	Cohen’s Kappa (mean ± SD)
EEGNet	0.656 ± 0.136	0.660 ± 0.136	0.640 ± 0.147	0.312 ± 0.271
CSP+LDA	0.634 ± 0.141	0.635 ± 0.141	0.605 ± 0.164	0.270 ± 0.281

EEGNet reached a mean accuracy of 0.656 ± 0.136 and a balanced accuracy of 0.660 ± 0.136, with a macro-F1 of 0.640 ± 0.147 and Cohen's kappa of 0.312 ± 0.271 (Table 1). The mean was above chance in the evaluated LOSO-z setting, but the large standard deviations show substantial variation between held-out subjects.

Under the same LOSO-z protocol and normalized inputs, CSP+LDA reached a mean accuracy of 0.634 ± 0.141, a balanced accuracy of 0.635 ± 0.141, a macro-F1 of 0.605 ± 0.164, and Cohen's kappa of 0.270 ± 0.281 (Table 1). Using the unrounded subject-level values reported in Figure 4, EEGNet's mean balanced-accuracy advantage was 0.0249, which appears as approximately 0.025 when the Table 1 means are rounded to three decimals. This is a modest average difference and must be interpreted together with the subject-level overlap.

Aggregated Confusion Matrix and Class-Wise Behaviour

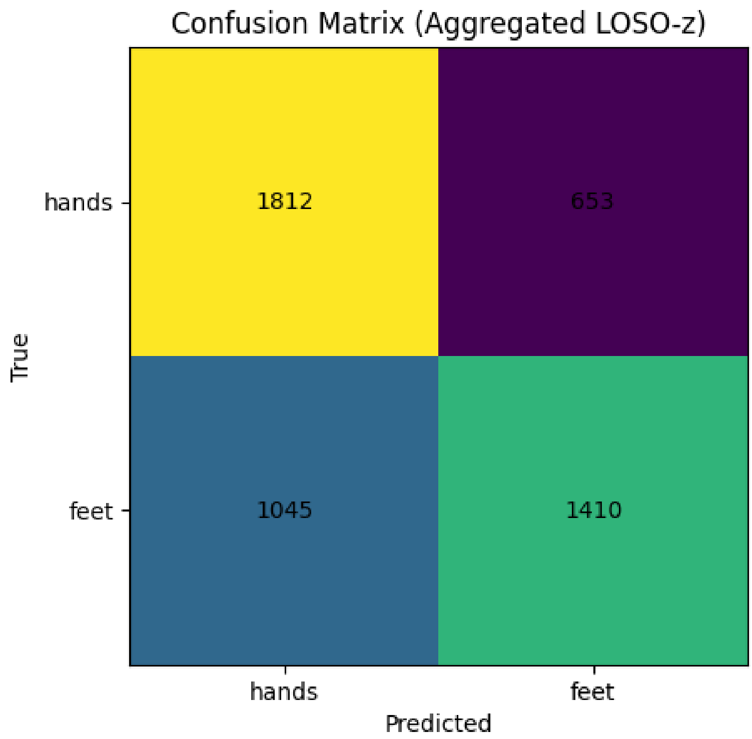

Figure 1. Aggregated confusion matrix of EEGNet under LOSO-z evaluation (109 subjects)

Figure 1 shows the aggregated confusion matrix of EEGNet across all LOSO folds under subject-wise z-score normalization. It reveals a clear class asymmetry. Hands imagery is recognised more reliably than feet imagery: most hands trials are classified correctly, while a large number of feet trials are misclassified as hands. From this matrix, recall is 0.735 for hands and 0.574 for feet. This confirms that feet imagery is harder to recognise under LOSO-z. This imbalance also explains why overall accuracy can look reasonable while balanced accuracy stays limited, since balanced accuracy is held down by the lower recall on the feet class.

For CSP+LDA, the aggregated confusion matrix shows a more even class-wise pattern, with recall of 0.641 for hands and 0.626 for feet. Compared with EEGNet, CSP+LDA reduced the feet-to-hands bias but had lower hands recall. Thus, EEGNet achieved the higher average score in this LOSO-z setting, while CSP+LDA distributed its errors more evenly across the two classes.

Inter-Subject Variability and Subject-Level Dispersion

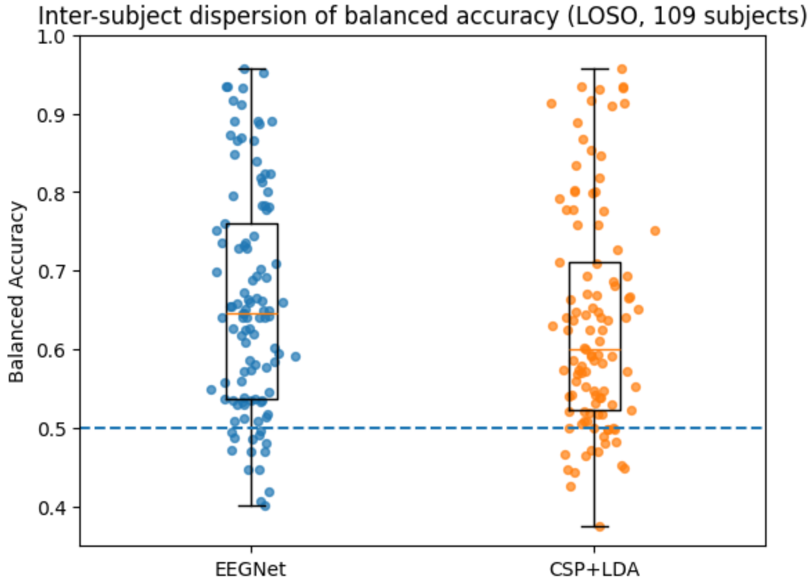

Figure 2. Inter-subject dispersion of balanced accuracy under LOSO-z for EEGNet and CSP+LDA

Subject-level performance shows large dispersion in balanced accuracy across the 109 subjects (Figure 2). The overall mean performance is above chance (Table 1), but the boxplots and the per-subject points show that the same pipeline can perform very well for some subjects and stay close to chance for others. Both EEGNet and CSP+LDA have a sizeable group of subjects near the chance line (0.5). This shows that subject-independent motor imagery decoding stays highly sensitive to differences between subjects under the LOSO protocol.

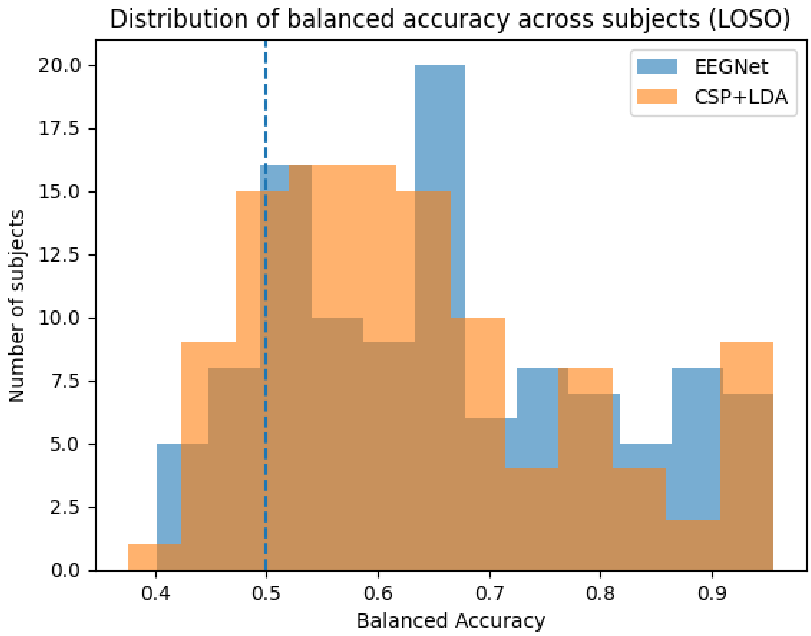

Figure 3. Distribution of subject-level balanced accuracy across 109 subjects under LOSO-z

Figure 3 shows how subject-level balanced accuracy is spread across the cohort for both methods. Most subjects cluster in a moderate range, roughly 0.55-0.70. Fewer subjects reach very high performance, and a smaller group stays near chance. Taken together, Figures 2 and 3 point to inter-subject variability as the main limit on cross-subject MI classification. This spread likely comes from a mix of recording factors, such as signal quality and electrode contact, and subject or session factors, such as strategy and fatigue. This motivates methods that directly target the shift between subjects for subject-independent MI decoding.

Paired Subject-Level Comparison

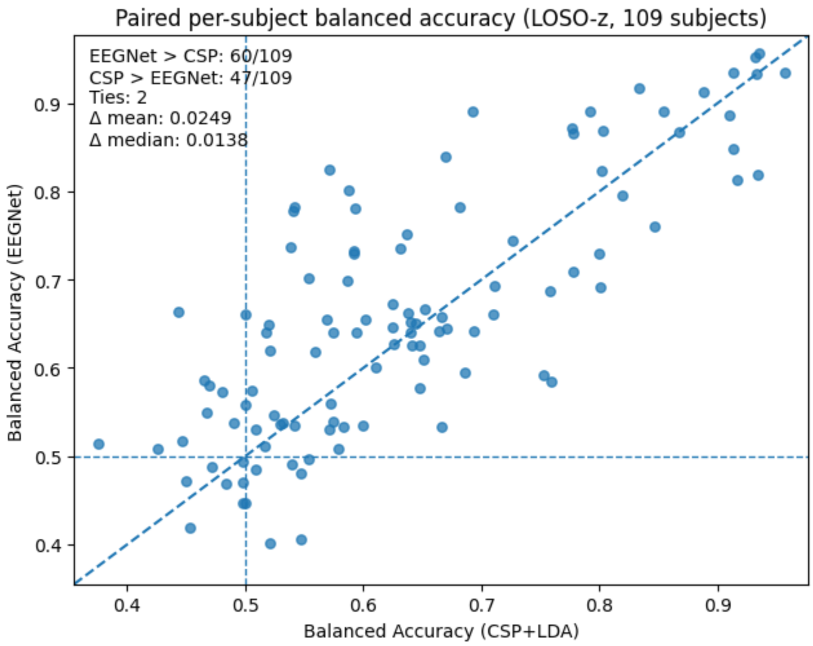

Figure 4. Paired per-subject balanced accuracy comparison between EEGNet and CSP+LDA under LOSO-z

Looking at each subject individually, EEGNet did better than CSP+LDA for 60 of the 109 subjects, CSP+LDA did better for 47, and 2 subjects were tied (Figure 4). The median difference between paired subjects was 0.0138. These results show that EEGNet scored higher on average, but they do not mean it won for everyone. This overlap in how subjects perform matters in practice. EEGNet had the higher average under LOSO-z, yet CSP+LDA was still the better model for a large share of individuals. The result therefore supports a scope-limited comparison rather than a claim of universal architectural superiority. Transfer learning and domain-generalization studies show that cross-subject shift can be addressed through adaptation or invariant feature learning, but these procedures require validation that keeps the target subject separate from model selection (Kwon et al., 2020; Liang et al., 2023; Wu et al., 2022; Zhang et al., 2021).

Discussion and Implications for Neuroinformatics and Informatics Education

In general, EEGNet produced slightly higher average scores than CSP+LDA for hands-versus-feet MI classification on EEGMMIDB under the evaluated LOSO-z setting. However, the dominant finding is inter-subject variability. EEGNet showed higher recall for hands than for feet (0.735 vs 0.574), whereas CSP+LDA produced a more even recall profile (0.641 vs 0.626). Balanced accuracy also ranged widely across subjects. These patterns agree with evidence that subject and dataset shifts limit EEG generalization (Kwon et al., 2020; Wang et al., 2024; Xu et al., 2020). They also show why architectural comparisons should not be separated from reporting decisions about filtering, validation, and model selection (Ang et al., 2012; Blankertz et al., 2008; Borra et al., 2025; Craik et al., 2019; Roy et al., 2019).

From a neuroinformatics perspective, the comparison shows that model choice cannot be separated from preprocessing, validation design, and subject-level reporting. The public EEGMMIDB dataset and the transparent LOSO-z workflow can therefore serve as an authentic case for informatics and software-engineering education. Students can reproduce the signal-processing pipeline, compare a feature-engineered method with a compact deep network, audit possible subject leakage, inspect aggregate and individual-level metrics, and document trade-offs among interpretability, computation, and generalization. Open data, analysis code, and structured laboratory resources can widen access to electrophysiology training, while good-practice guidance stresses transparent preprocessing, statistical decisions, machine-learning validation, and reporting (Bukach et al., 2021; Niso et al., 2022). This is a curriculum implication of the reproducible workflow; the present study did not measure student learning.

The same subject-level variability has a prospective implication for assistive educational systems. A classifier that works well for some users but remains near chance for others could create unreliable or unequal interactions. Practical systems would therefore need uncertainty reporting, user-specific calibration or adaptation, fallback interaction methods, and validation across diverse users before deployment. Physiological outputs should also be connected to clearly defined learning or interaction outcomes rather than treated as direct measures of learning (Antonenko et al., 2010; Nouri, 2025; Xu & Zhong, 2018). These are future design requirements only; the study did not evaluate educational users, usability, or learning outcomes.

CONCLUSION

This study compared EEGNet and CSP+LDA for hands-versus-feet motor imagery classification on EEGMMIDB across 109 held-out subjects. Under the evaluated LOSO-z setting, EEGNet achieved higher average accuracy, balanced accuracy, macro-F1, and Cohen's kappa than CSP+LDA. The advantage was not consistent across all subjects: EEGNet scored higher for 60 subjects, CSP+LDA for 47 subjects, and 2 subjects were tied. Inter-subject dispersion and EEGNet's lower recall for feet imagery remained the main practical limitations. Because normalization used the held-out subject's unlabeled statistics, these findings apply to LOSO with unsupervised test-time normalization and should not be described as fully inductive calibration-free performance. Future work should evaluate alignment, domain-generalization, and lightweight adaptation methods under clearly separated training, validation, and test procedures.

For neuroinformatics and informatics education, this benchmark provides a concrete case for teaching reproducible biosignal processing, classical-versus-deep model comparison, leakage-aware validation, and the interpretation of subject-level variability. For future assistive educational technology, the findings show why uncertainty handling, calibration, and equitable performance across users must be treated as design requirements rather than optional refinements. These implications remain prospective: usability, educational effectiveness, fairness, and deployment with diverse learners require separate empirical evaluation (Bukach et al., 2021; Nouri, 2025).

STATEMENTS AND DECLARATIONS

Competing Interests. The authors declare no competing interests.

Author Contributions. NA: Conceptualization, methodology, project administration, software development, data curation, formal analysis, visualization, and writing the original draft. FNH and WS: Methodology, validation, interpretation of the findings, and review and editing of the manuscript. All authors reviewed and approved the final version of the manuscript.

Ethics Approval. Not applicable. This study involved a secondary analysis of a publicly available EEG dataset and did not involve direct participant recruitment, interaction with participants, or the collection of new human-subject data.

Informed Consent to Participate. No participants were recruited and no new human-subject data were collected.

Consent for Publication. Not applicable.

Data Availability. The EEG Motor Movement/Imagery Dataset analyzed in this study is publicly available through PhysioNet.

Use of Artificial Intelligence Tools. No generative artificial intelligence tools were used in the preparation of this manuscript. Grammarly was used for grammar and language correction.

REFERENCES

Ang, K. K., Chin, Z. Y., Wang, C., Guan, C., & Zhang, H. (2012). Filter bank common spatial pattern algorithm on BCI Competition IV datasets 2a and 2b. *Frontiers in Neuroscience, 6*, 39. https://doi.org/10.3389/fnins.2012.00039

Antonenko, P., Paas, F., Grabner, R., & van Gog, T. (2010). Using electroencephalography to measure cognitive load. *Educational Psychology Review, 22*(4), 425–438. https://doi.org/10.1007/s10648-010-9130-y

Autthasan, P., Chaisaen, R., Sudhawiyangkul, T., Rangpong, P., Kiatthaveephong, S., Dilokthanakul, N., Bhakdisongkhram, G., Phan, H., Guan, C., & Wilaiprasitporn, T. (2022). MIN2Net: End-to-end multi-task learning for subject-independent motor imagery EEG classification. *IEEE Transactions on Biomedical Engineering, 69*(6), 2105–2118. https://doi.org/10.1109/TBME.2021.3137184

Blankertz, B., Tomioka, R., Lemm, S., Kawanabe, M., & Müller, K.-R. (2008). Optimizing spatial filters for robust EEG single-trial analysis. *IEEE Signal Processing Magazine, 25*(1), 41–56. https://doi.org/10.1109/MSP.2008.4408441

Borra, D., Magosso, E., & Ravanelli, M. (2025). A protocol for trustworthy EEG decoding with neural networks. *Neural Networks, 182*, 106847. https://doi.org/10.1016/j.neunet.2024.106847

Bukach, C. M., Bukach, N., Reed, C. L., & Couperus, J. W. (2021). Open science as a path to education of new psychophysiologists. *International Journal of Psychophysiology, 165*, 76–83. https://doi.org/10.1016/j.ijpsycho.2021.04.001

Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. *Journal of Neural Engineering, 16*(3), 031001. https://doi.org/10.1088/1741-2552/ab0ab5

Goldberger, A. L., Amaral, L. A. N., Glass, L., Hausdorff, J. M., Ivanov, P. C., Mark, R. G., Mietus, J. E., Moody, G. B., Peng, C.-K., & Stanley, H. E. (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation, 101*(23), e215–e220. https://doi.org/10.1161/01.CIR.101.23.e215

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Parkkonen, L., & Hämäläinen, M. S. (2014). MNE software for processing MEG and EEG data. *NeuroImage, 86*, 446–460. https://doi.org/10.1016/j.neuroimage.2013.10.027

Keutayeva, A., Fakhrutdinov, N., & Abibullaev, B. (2024). Compact convolutional transformer for subject-independent motor imagery EEG-based BCIs. *Scientific Reports, 14*, 25775. https://doi.org/10.1038/s41598-024-73755-4

Kwon, O.-Y., Lee, M.-H., Guan, C., & Lee, S.-W. (2020). Subject-independent brain–computer interfaces based on deep convolutional neural networks. *IEEE Transactions on Neural Networks and Learning Systems, 31*(10), 3839–3852. https://doi.org/10.1109/TNNLS.2019.2946869

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. *Journal of Neural Engineering, 15*(5), 056013. https://doi.org/10.1088/1741-2552/aace8c

Liang, S., Xuan, C., Hang, W., Lei, B., Wang, J., Qin, J., Choi, K.-S., & Zhang, Y. (2023). Domain-generalized EEG classification with category-oriented feature decorrelation and cross-view consistency learning. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31*, 3285–3296. https://doi.org/10.1109/TNSRE.2023.3300961

Liu, K., Yang, M., Yu, Z., Wang, G., & Wu, W. (2023). FBMSNet: A filter-bank multi-scale convolutional neural network for EEG-based motor imagery decoding. *IEEE Transactions on Biomedical Engineering, 70*(2), 436–445. https://doi.org/10.1109/TBME.2022.3193277

Lotte, F., Bougrain, L., Cichocki, A., Clerc, M., Congedo, M., Rakotomamonjy, A., & Yger, F. (2018). A review of classification algorithms for EEG-based brain–computer interfaces: A 10 year update. *Journal of Neural Engineering, 15*(3), 031005. https://doi.org/10.1088/1741-2552/aab2f2

Niso, G., Krol, L. R., Combrisson, E., Dubarry, A. S., Elliott, M. A., François, C., Héjja-Brichard, Y., Herbst, S. K., Jerbi, K., Kovic, V., Lehongre, K., Luck, S. J., Mercier, M., Mosher, J. C., Pavlov, Y. G., Puce, A., Schettino, A., Schön, D., Sinnott-Armstrong, W., … Chaumon, M. (2022). Good scientific practice in EEG and MEG research: Progress and perspectives. *NeuroImage, 257*, 119056. https://doi.org/10.1016/j.neuroimage.2022.119056

Nouri, A. (2025). A scoping review of educational neurotechnology: Methods, applications, opportunities, and challenges. *Review of Education, 13*, e70070. https://doi.org/10.1002/rev3.70070

Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: Basic principles. *Clinical Neurophysiology, 110*(11), 1842–1857. https://doi.org/10.1016/S1388-2457(99)00141-8

Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: A systematic review. *Journal of Neural Engineering, 16*(5), 051001. https://doi.org/10.1088/1741-2552/ab260c

Sakhavi, S., Guan, C., & Yan, S. (2018). Learning temporal information for brain–computer interface using convolutional neural networks. *IEEE Transactions on Neural Networks and Learning Systems, 29*(11), 5619–5629. https://doi.org/10.1109/TNNLS.2018.2789927

Schalk, G., McFarland, D. J., Hinterberger, T., Birbaumer, N., & Wolpaw, J. R. (2004). BCI2000: A general-purpose brain–computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering, 51*(6), 1034–1043. https://doi.org/10.1109/TBME.2004.827072

Schirrmeister, R. T., Springenberg, J. T., Fiederer, L. D. J., Glasstetter, M., Eggensperger, K., Tangermann, M., Hutter, F., Burgard, W., & Ball, T. (2017). Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping, 38*(11), 5391–5420. https://doi.org/10.1002/hbm.23730

Song, Y., Zheng, Q., Liu, B., & Gao, X. (2023). EEG Conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 31*, 710–719. https://doi.org/10.1109/TNSRE.2022.3230250

Wang, X., Liesaputra, V., Liu, Z., Wang, Y., & Huang, Z. (2024). An in-depth survey on deep learning-based motor imagery electroencephalogram (EEG) classification. *Artificial Intelligence in Medicine, 147*, 102738. https://doi.org/10.1016/j.artmed.2023.102738

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology, 113*(6), 767–791. https://doi.org/10.1016/S1388-2457(02)00057-3

Wu, D., Xu, Y., & Lu, B.-L. (2022). Transfer learning for EEG-based brain–computer interfaces: A review of progress made since 2016. *IEEE Transactions on Cognitive and Developmental Systems, 14*(1), 4–19. https://doi.org/10.1109/TCDS.2020.3007453

Xu, J., & Zhong, B. (2018). Review on portable EEG technology in educational research. *Computers in Human Behavior, 81*, 340–349. https://doi.org/10.1016/j.chb.2017.12.037

Xu, L., Xu, M., Ke, Y., An, X., Liu, S., & Ming, D. (2020). Cross-dataset variability problem in EEG decoding with deep learning. *Frontiers in Human Neuroscience, 14*, 103. https://doi.org/10.3389/fnhum.2020.00103

Zhang, K., Robinson, N., Lee, S.-W., & Guan, C. (2021). Adaptive transfer learning for EEG motor imagery classification with deep convolutional neural network. *Neural Networks, 136*, 1–10. https://doi.org/10.1016/j.neunet.2020.12.013