Kolmogorov-Arnold Networks (KANs) for motor imagery classification: A benchmark with MLP

Jaime A. Riascos-Salas[12] Hernan Villota[1] D Betancur-Vasquez[1] Edwin A. Moncada[1]

1- GITESI - Engineering Faculty - Institution University of Envigado (IUE)
Carrera 27B-39A Sur 57- Envigado, Colombia

2- PECoG - Department of Psychology, University of Potsdam
Karl-Liebknecht-Str. 24-25, 14476 - Potsdam, Germany

Abstract.
This study evaluates the performance of Kolmogorov-Arnold Networks (KANs), a novel neural network architecture with learnable activation functions, in motor imagery classification tasks. Benchmarked against Multi-Layer Perceptrons (MLPs), KANs exhibited comparable performance with Common Spatial Patterns (CSP) features and superior training accuracy with raw EEG data. However, both models faced challenges generalizing to unseen data. While KANs require longer training times, their inherent explainability through explicit inference functions offers potential advantages for understanding brain signal classification. Future research is needed to enhance KANs' robustness and generalization capabilities in real-world BCI applications.

1 Introduction

Motor imagery Brain-Computer Interfaces (MI-BCIs) establish a direct communication pathway between the human brain and external computing devices by leveraging the neural correlates of imagined movements [1]. While significant advancements have been made in recent years, the accurate decoding of electroencephalography (EEG) signals, translating motor-related patterns into actionable commands, remains a formidable challenge in BCI research [2]. Traditional Machine Learning approaches have demonstrated moderate success in achieving acceptable accuracy rates with relatively efficient processing times. However, their performance is heavily contingent on the efficacy of the feature extraction stage [3]. Conversely, Deep Learning models, while boasting superior accuracy, are often encumbered by protracted training and prediction times, hindering their seamless integration into real-time BCI applications [4].

Liu, Ziming and colleagues [5] recently introduced the Kolmogorov-Arnold Networks (KANs) (original paper), a novel neural network architecture rooted in the Kolmogorov-Arnold representation theorem, offering an alternative to traditional Multi-Layer Perceptrons (MLPs). While MLPs learn weights on connections between neurons, KANs focus on learning the activation function at each node. This approach enhances model interpretability, as the role of each node in the overall computation is more transparent. Furthermore, due to the Kolmogorov-Arnold representation theorem, KANs potentially offer computational efficiency advantages over MLPs. This theorem suggests that fewer nodes might be sufficient to represent complex functions, leading to faster training and inference times. Moreover, the inherent structure of KANs facilitates the integration of new information through node addition or modification, making them well-suited for continual learning environments.

This study presents a preliminary assessment of Kolmogorov-Arnold Networks (KANs) as a decoding model for motor imagery-based brain-computer interfaces (MI-BCIs). A benchmark analysis was conducted using Dataset IIa from BCI Competition IV, employing both multiclass and binary classification paradigms. KAN performance was compared against a traditional Multi-Layer Perceptron (MLP) in terms of training, validation, and evaluation accuracy, along with computational time during training and evaluation phases. To assess model efficacy under typical BCI scenarios, both raw electroencephalography (EEG) signals and Common Spatial Patterns (CSP) features were utilized independently for model training. The subsequent sections of this paper detail the experimental methodology, including Dataset IIa description, KAN model overview, and benchmark procedures (Section 2), results (Section 3), and concluding remarks (Section 4).

2 Material and Methods

2.1 Dataset IIa from BCI Competition 4

Electroencephalography (EEG) data was collected from nine subjects during a cue-based brain-computer interface (BCI) paradigm. Participants performed four distinct motor imagery tasks: left hand, right hand, both feet, and tongue movement imagination. Each subject completed two sessions on separate days, with each session consisting of six runs punctuated by brief rest periods. Each run encompassed 48 trials (12 per class), totaling 288 trials per session. We used one session for training and the other for evaluation. Trials began with the presentation of a fixation cross and a short auditory cue. After two seconds, a directional arrow corresponding to one of the four motor imagery tasks was displayed for 1.25 seconds, prompting the subject to initiate the designated mental task. Participants continued the motor imagery task until the fixation cross disappeared at the six-second mark. No performance feedback was provided. The EEG signals were digitized at 250 Hz. The BCI Competition IV-2a (BCI-2a) dataset is one of the most used for benchmarking BCI models [6] because it presents challenges due to its limited sample size, uncontrolled recording conditions, and presence of artifacts. These factors collectively contribute to the complexity of accurately decoding MI tasks using this dataset.

2.2 Kolgomorov-Arnold Networks (KAN)

Vladimir Arnold and Andrey Kolmogorov made a significant contribution to the field of mathematical function analysis with their theorem on the representation of multivariate functions. According to their findings, any multivariate continuous function defined on a bounded domain can essentially be decomposed into a finite combination of continuous univariate functions and the operation of addition. This is exemplified by their formulation for a smooth function:

$$F(X) = f(X_1, \cdots, X_n) = \sum_{q=1}^{2n+1} \phi_q(\sum_{p=1}^{n} \phi_{q,p}(X_p))$$

is a real-valued function. Essentially, this theorem suggests that the only truly multivariate operation is addition, since other complex functions can be represented through combinations of simpler, one-dimensional functions.

The Kolmogorov-Arnold representation theorem, while theoretically promising for machine learning due to its potential to simplify the learning of high-dimensional functions, has faced challenges in practical implementation due to the potential for non-smooth and fractal representations. However, recent advancements in modeling techniques, such as those proposed by Liu et al. (2024), coupled with the prevalence of smooth and sparsely structured functions in real-world scenarios, suggest that the theorem's utility in machine learning is increasingly viable. This is particularly true for tasks that align with common scenarios encountered in physics, where the theorem's focus on practicality over theoretical limitations may be advantageous. In contrast to the well-established Multilayer Perceptron (MLP) approach, which relies on the universal approximation theorem and backpropagation for minimizing error, Kolmogorov-Arnold Networks (KANs) leverage the theorem to conceptualize synaptic weights as activation functions within a multidimensional geometric space. This innovative approach addresses the challenges posed by non-smooth and fractal functions, thereby bridging theoretical mathematical frameworks with practical applications in machine learning, potentially enhancing adaptability and efficacy in classification tasks.

2.3 Experimental setup

Four distinct experimental paradigms were conducted: (i) raw EEG binary classification, (ii) raw EEG multiclass classification, (iii) CSP-filtered EEG binary classification, and (iv) CSP-filtered EEG multiclass classification. Binary classification focused on differentiating left-right motor imagery. Model performance was evaluated using 10-fold cross-validation, assessing accuracy during training, testing, and an independent evaluation set derived from the second session of the dataset. To ensure a fair comparison, the training setup from KAN was transferred to an MLP architecture with 10 hidden layers. Both networks utilized default parameters as proposed by Liu et al. (2024). Training was conducted over 20 epochs with a learning rate of 0.1, a batch size of 16, and cross-entropy loss optimization. All experiments were executed on a Windows 11 machine equipped with an Intel CoreTM i7-13700K CPU (30 MB Cache, 16 Cores, 24 threads, 3.4-5.3 GHz, 125 W), 32 GB DDR5 RAM, and 1TB SSD, utilizing only CPU resources. Training and prediction times were recorded independently. The full codes used are available in the GitHub repository [1]

[1] `https://github.com/JARS29/KAN_MLP_MI-BCI`

3 Experimental results

Initially we report the results of the binary classification scenario. Table 1 indicates the outcomes of the binary classification task. For the CSP-filtered data, both KAN and MLP models exhibited comparable performance. The KAN model achieved a mean training accuracy of 0.82 (SD = 0.08), while the MLP model attained 0.81 (SD = 0.08). In the testing phase, the KAN model maintained consistent performance, averaging 0.71 (SD = 0.10), with a range of 0.61 to 0.95. The MLP model demonstrated slightly higher accuracy, averaging 0.76 (SD = 0.10) and ranging from 0.67 to 0.98. In contrast, for raw EEG classification, KAN outperformed MLP in training accuracy, achieving a mean of 0.79 (SD = 0.09) compared to 0.72 (SD = 0.01) respectively. However, both models failed to achieve competitive accuracy during testing, with all test accuracies at 0.50. This indicates that neither model effectively generalized to unseen data in the raw EEG condition.

	CSP				RAW			
	KAN		MLP		KAN		MLP	
ID	Train	Test	Train	Test	Train	Test	Train	Test
1	0.77	0.66	0.75	0.70	0.88	0.50	0.74	0.50
2	0.77	0.66	0.77	0.73	0.79	0.50	0.76	0.50
3	0.91	0.81	0.93	0.90	0.85	0.51	0.71	0.50
4	0.74	0.63	0.74	0.70	0.87	0.50	0.71	0.50
5	0.80	0.69	0.80	0.70	0.79	0.50	0.73	0.50
6	0.74	0.61	0.76	0.67	0.65	0.50	0.70	0.50
7	0.81	0.69	0.82	0.77	0.63	0.50	0.72	0.50
8	1.00	0.95	0.98	0.98	0.84	0.50	0.73	0.50
9	0.85	0.73	0.81	0.76	0.87	0.50	0.72	0.50
AVG (STD)	0.82 (0.08)	0.71 (0.10)	0.81 (0.08)	0.76 (0.10)	0.79 (0.09)	0.50 (0.00)	0.72 (0.01)	0.50 (0.00)

Table 1: Binary classification

Table 2 presents the accuracies for multiclass classification tasks. In the CSP-filtered condition, both KAN and MLP models achieved similar average training accuracies of 0.60. However, the MLP model outperformed the KAN model in terms of testing accuracy, with an average of 0.51 (SD = 0.10) compared to 0.47 (SD = 0.08). Despite comparable training performance, the MLP model demonstrated superior generalization capabilities on unseen data. In the raw EEG condition, both KAN and MLP models exhibited consistent training accuracies of 0.59 (SD = 0.11) and 0.46 (SD = 0.00), respectively. However, similar to the binary classification scenario, both models struggled to generalize, achieving test accuracies near chance level. This suggests that neither model adequately captured the complex patterns in the raw EEG data necessary for accurate multiclass discrimination. The inference results (no reported in this paper) further corroborate the models' limited generalization capabilities, mirroring their suboptimal performance during training. Both KAN and MLP models failed to generalize to unseen data, even with prior exposure during training. However, CSP filtering demonstrated a marginal improvement in accuracy, with KAN outperforming MLP in the binary classification task, while the reverse occurred in the multiclass scenario.

	CSP				RAW			
	KAN		MLP		KAN		MLP	
ID	Train	Test	Train	Test	Train	Test	Train	Test
1	0.67	0.53	0.67	0.60	0.46	0.25	0.48	0.25
2	0.61	0.51	0.64	0.56	0.62	0.25	0.46	0.25
3	0.69	0.57	0.66	0.61	0.74	0.26	0.46	0.25
4	0.50	0.35	0.50	0.41	0.64	0.25	0.47	0.25
5	0.47	0.36	0.43	0.36	0.38	0.25	0.47	0.25
6	0.50	0.38	0.51	0.39	0.60	0.25	0.47	0.25
7	0.63	0.51	0.64	0.54	0.61	0.25	0.48	0.25
8	0.68	0.56	0.69	0.61	0.66	0.25	0.46	0.25
9	0.65	0.48	0.66	0.56	0.68	0.26	0.48	0.25
AVG (SD)	0.60 (0.08)	0.47 (0.08)	0.60 (0.09)	0.51 (0.10)	0.59 (0.11)	0.25 (0.00)	0.46 (0.00)	0.25 (0.00)

Table 2: Multiclass classification

Finally, while there was no huge differences between models in accuracy, we report that KAN models consistently require significantly more time for training compared to MLPs across all conditions (Table 3). This could be attributed to the more complex nature of KANs and their training process, which involves optimizing activation functions at each node rather than simple weights. The difference in training time is particularly pronounced in the raw EEG condition, suggesting that KANs might require more computational resources to learn from unprocessed, high-dimensional EEG data.

Cond	Feature	Model	Train (pred) time
Multiclass	CSP	KAN	1138 (4.49e-3)
		MLP	70 (4.46e-3)
	RAW	KAN	3183 (8.02)
		MLP	209 (8.03)
Binary	CSP	KAN	711 (7.52e-2)
		MLP	50 (7.63e-2)
	RAW	KAN	1880 (1.41)
		MLP	184 (1.41)

Table 3: Time for training and inference (pred) in seconds.

The ability to derive explicit inference functions for KAN models represents a significant step towards greater transparency and explainability in BCI research. By deriving explicit inference functions for each class (Equations 1-4 multiclass CSP), KANs offer a level of transparency that is often lacking in traditional "black-box" machine learning models like MLPs. This approach holds promise for enhancing our understanding of the relationship between neural activity and motor imagery, ultimately leading to more effective and reliable BCI systems.

$$
\begin{aligned}
f_1(x_n) = &-126.29(-x_1 - 0.28)^2 - 1.4(-x_3 - 0.54)^3 \\
&+ 79.68(-x_4 - 0.09)^3 + 5.63\sin(4.08x_2 + 5.42) + 132.31
\end{aligned} \tag{1}
$$

$$
\begin{aligned}
f_2(x_n) = &\ 36.8(0.05 - x_4)^3 - 75.07\sin(2.31x_1 - 7.42) \\
&- 367.97\tanh(3.3x_2 + 4.69) + 424.81 + 0.1\exp(-4.88x_3)
\end{aligned} \tag{2}
$$

$$
\begin{aligned}
f_3(x_n) = &\ 12.13(0.07 - x_3)^4 - 172.78(-x_1 - 0.15)^2 + 95.51(-x_2 - 0.2)^4 \\
&- 0.8\,|8.36x_4 + 3.66| + 131.67
\end{aligned} \tag{3}
$$

$$
\begin{aligned}
f_4(x_n) = &\ 142.03\sin(1.41x_1 + 8.2) - 1.65\sin(4.71x_3 + 8.19) \\
&- 2.38\sin(3.26x_4 - 1.98) + 520.46\tanh(2.41x_2 + 3.2) - 530.95
\end{aligned} \tag{4}
$$

4 Conclusions

The assessment of Kolmogorov-Arnold Networks (KANs) and Multi-Layer Perceptrons (MLPs) in motor imagery (MI) classification reveals that while KANs demonstrate comparable performance to MLPs when coupled with Common Spatial Patterns (CSP) and even superior training accuracy with raw EEG data, both models struggle with generalization to unseen data, highlighting the challenges inherent in real-world BCI applications due to individual variability and session-to-session fluctuations. This study provides preliminary evidence for the potential of KANs in MI-BCI, emphasizing the need for further research to improve model robustness, generalization, and exploit the inherent explainability of KANs for deeper understanding of brain signal classification.

References

[1] J. R. Wolpaw, N. Birbaumer, D. J. McFarland, G. Pfurtscheller, and T. M. Vaughan. Brain-computer interfaces for communication and control. *Clinical neurophysiology*, 113(6):767–791, 2002.

[2] F Lotte, L Bougrain, A Cichocki, M Clerc, M Congedo, A Rakotomamonjy, and F Yger. A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update. *J Neural Eng*, 15(3):031005, February 2018.

[3] K. K. Ang, Z. Y. Chin, H. Zhang, and C. Guan. Filter bank common spatial pattern (FBCSP) in brain-computer interface. In *2012 IEEE International Joint Conference on Neural Networks (IJCNN)*, pages 1–8. 2012.

[4] A. Craik, Y. He, and J. L. Contreras-Vidal. Deep learning for electroencephalogram (EEG) classification tasks: a review. *Journal of neural engineering*, 16(3):031001., 2019.

[5] Ziming Liu, Yixuan Wang, Sachin Vaidya, Fabian Ruehle, James Halverson, Marin SoljaÄiÄ, Thomas Y. Hou, and Max Tegmark. Kan: Kolmogorov-arnold networks, 2024.

[6] Arpaia P, Esposito A, Natalizio A, and Parvis M. How to successfully classify EEG in motor imagery BCI: a metrological analysis of the state of the art. *J Neural Eng. 2022 Jun 17;19(3). doi:*, 10., 2022.