Classifying mental states using Brain-Computer Interfaces: A systematic review

Jaime A. Riascos[1,2], Martin Fischer[1], Klaus Gramman[3]

[1]Potsdam University - Potsdam, Germany

[2]Institution University of Envigado (IUE) - Envigado, Colombia

[3]Technical University of Berlin - Berlin, Germany

Synopsis

This study systematically reviews the latest advances in EEG-based brain-computer interfaces (BCIs) for identifying mental states. We examine feature extraction techniques, classification algorithms, and target mental states (e.g., emotion, attention) reported in recent research (2018-2023). Our analysis highlights common practices and key feature extraction and classification trends, achieved classifier accuracies, and provides a valuable resource for researchers seeking to decode mental states using BCI.

Background

Identifying human mental states involves understanding the complex interplay of physiological and cognitive processes. Mental states shape our behavior and well-being. Advances in neuroscience, particularly electroencephalography (EEG), offer a window into the brain's electrical activity, providing potential insights into these internal states (Roy and Frey, 2016), and a real-time observation crucial for Brain-Computer Interface (BCI). BCI translates neural signals into commands for device control (Mridha et al., 2021). While active and reactive BCIs rely on intentional control and stimulus-driven responses, respectively, passive BCIs (pBCIs) aim to infer mental states without explicit control, optimizing the user experience (Alimardani and Hiraki, 2020; Zander & Kothe, 2011). However, BCI development poses challenges in signal processing, feature extraction, and interface design due to EEG's multidimensionality, non-stationarity, and individual variability (Cohen, 2017). To address this, researchers employ diverse methods to extract meaningful insights, yet selecting optimal features and machine learning algorithms remains challenging as careful selection of these methods is crucial for accurate mental state decoding. This paper reviews BCI research over the last five years, focusing on features and machine learning algorithms for decoding mental states involving attention, cognition, emotion, and fatigue. We examined common techniques, accuracies achieved, and primary mental states reported in the literature to provide a valuable resource for researchers seeking to establish robust pipelines for mental state identification using BCI to drive further advancements in BCI research.

Methods

We conducted comprehensive literature searches across prominent repositories (PubMed, Science Direct, IEEE) from 2018 to 2023 (including July 31st, 2023). To focus on EEG-based BCIs, feature extraction, and classification techniques, we employed the following search terms in titles, abstracts, and keywords: ("feature extraction" OR "classification algorithms" OR "machine learning algorithms") AND ("brain-computer interfaces" OR "BCI") AND ("EEG" OR "electroencephalography" OR "EEG-based" OR "electroencephalography-based"). This broad initial search ensured the comprehensive retrieval of relevant studies. We exported results in

RIS format, utilizing Mendeley for data management and duplicate removal. Initial screening excluded 415 out of 953 papers based on title and abstract (e.g., lack of EEG focus, reviews/editorial, non-BCI topics). We then classified the remaining papers by EEG paradigms (SSVEP, Motor Imagery, etc.), with 418 addressing active/reactive BCIs and 120 focused on passive BCIs. Finally, 72 papers were selected to provide comprehensive reporting on EEG procedures, feature extraction methods, and the classification of mental states. We extracted this information from each paper, following data science best practices to construct our final dataset.

Results

We systematically categorized selected papers based on machine learning algorithms and EEG feature categories, then, we analyzed the frequency of use to identify trends in BCI design for mental state classification. Algorithm categories included:

- **Geometric:** Emphasize decision boundaries for class separation.
- **Probabilistic:** Utilize class membership probabilities for predictions.
- **Neural Networks:** Employ complex interconnected units for pattern learning.
- **Ensemble:** Combine predictions from multiple base models.
- **Other:** Encompass algorithms that do not fit neatly into other categories.

Feature categories were:

- **Temporal:** Reflecting changes in signal amplitude over time.
- **Spectral:** Frequency distribution of the signal.
- **Spatial:** Relationships between signals from different brain regions.
- **Information Theoretical:** Quantifies information content.
- **Statistical:** Basic signal descriptions and properties.
- **Other:** Combinations of the above categories.

When combined features are used, we summed this to each corresponding category. Figure 1 illustrates the top classifiers in each category, along with their mean, maximum, and minimum reported accuracies. Most classifiers (excluding Ensemble and AMGLVQ) achieved over 90% accuracy in at least one study. Notably, our findings illustrated the dominance of geometric models, with a total of 66 occurrences and an average accuracy exceeding 70%. Convolutional Neural Networks (CNNs) emerged as the most proficient classifiers, boasting an average accuracy of 84% among the most frequently utilized models. Furthermore, Figure 2 highlights the most popular features and extraction methods. Our investigation revealed the prominence of temporal and spectral features, with 43 and 38 occurrences, respectively. Power Spectral Density (PSD) was the most frequently reported feature (21 occurrences), followed closely by Differential Entropy (18 occurrences). In terms of feature extraction methods, Fourier Transform (16), Wavelet Transform (11), and Empirical Mode Decomposition (5) were the most common techniques. Lastly, we observed a balanced evaluation of various mental states, including Emotion levels (23), Valence (19), Arousal (18), Attention (10), and Dominance (6), highlighting the comprehensive nature of the studies analyzed. From these, we identified the next classes among the most used: Positive/Negative (73), Happiness (8), and Drowsing (7). Sadness (7), Fear (6). Intriguingly, only 2.8% of the works (two papers) performed an online classification. Finally, we also observed a relatively even distribution across frequency bands (Alpha, Beta, Theta, Gamma, Delta).

Discussion

Identifying and classifying mental states using BCIs presents challenges despite advancements in EEG signal processing and artificial intelligence. This systematic review of methods used, and classification accuracies for different mental states achieved, provides important information for researchers in selecting optimal feature extraction and classification configurations from the vast array of options that describe current trends in BCI to guide future research endeavors. Thus, this work represents an opportunity to create stable BCI pipelines and updates about the latest advancements. Deep learning methods, such as convolutional neural networks (CNNs), offer automatic feature extraction and classification, potentially simplifying the process. However, their complexity and real-time implementation constraints must be considered (Craik and Contreras-Vidal, 2019).

On the other hand, spectral features, particularly PSD, are widely used in cognitive assessment, from emotion recognition to fatigue detection (Gao et al, 2020). As mental state identification often involves contrasting classes (e.g., positive/negative valence), information theory measures like differential entropy can quantify differences between EEG signals associated with each state. Thus, trade-offs between complex machine-learning approaches and simple informative features are desirable (Lotte et al., 2018). Our comprehensive approach ensures the inclusion of pertinent literature, providing a robust foundation for further analysis and insights into EEG-based BCI research. Datasets modality should also be extensively studied, as most of the works used images, videos, or sound stimuli. Finally, a discussion about the categories (for features and classifiers) should be carried out to group accordingly each method. Likewise, we advocate for authors to concisely report key findings, facilitating extracting relevant information regarding features, extraction methods, classifiers, labels, and accuracies.

Disclosures

References

(Roy and Frey, 2016) Roy, R.N. and Frey, J. (2016). Neurophysiological Markers for Passive Brain–Computer Interfaces. In Brain–Computer Interfaces 1 (eds M. Clerc, L. Bougrain and F. Lotte). https://doi.org/10.1002/9781119144977.ch5

(Mridha et al., 2021) Mridha MF, Das SC, Kabir MM, Lima AA, Islam MR, Watanobe Y. Brain-Computer Interface: Advancement and Challenges. Sensors (Basel). 2021 Aug 26;21(17):5746. doi: 10.3390/s21175746. PMID: 34502636; PMCID: PMC8433803.

(Alimardani and Hiraki, 2020) Alimardani, M., & Hiraki, K. (2020). Passive Brain-Computer Interfaces for Enhanced Human-Robot Interaction. Frontiers in Robotics and AI, 7.

(Cohen, 2017) Cohen, M. X. (2017). Analyzing neural time series data: theory and practice. MIT press.

(Craik and Contreras-Vidal, 2019) 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.

(Gao et al, 2020) Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S. (2020). EEG-based spatio–temporal convolutional neural network for driver fatigue evaluation. IEEE transactions on neural networks and learning systems, vol. 30, no. 9, pp. 2755-2763, Sept. 2019, doi: 10.1109/TNNLS.2018.2886414.

(Lotte et al., 2018) 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.

(Zander and Kothe, 2011) Zander, T. O., & Kothe, C. (2011). Towards passive brain–computer interfaces: applying brain–computer interface technology to human–machine systems in general. Journal of neural engineering, 8(2), 025005.

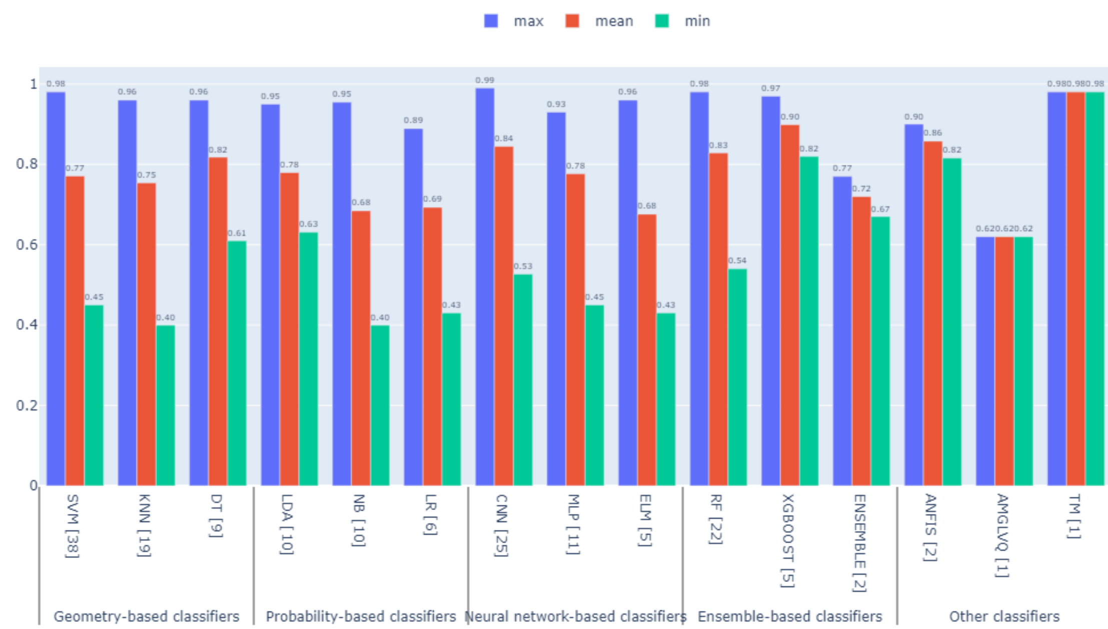

Figure 1. The three most used classifiers by category with the maximal, minimal, and mean accuracy reported. In brackets, the occurrence of the model.

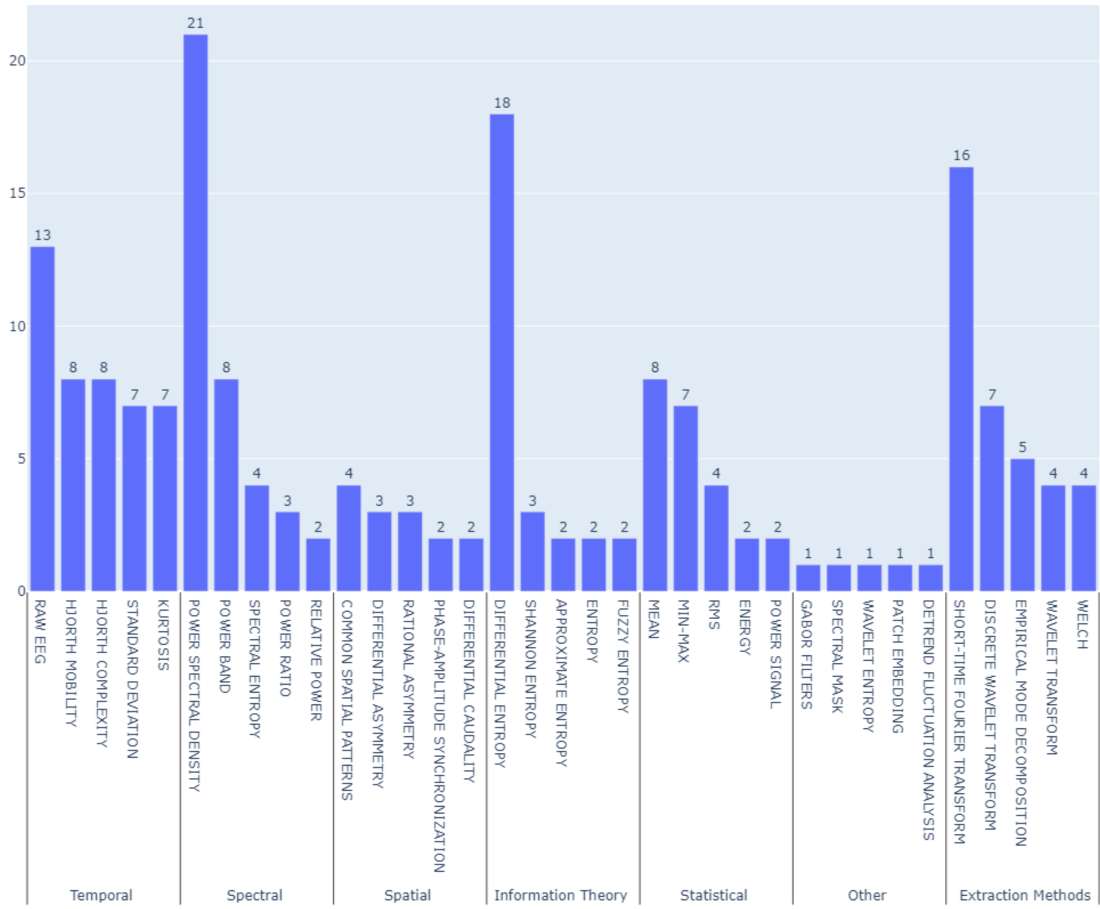

Figure 2. The most used features by category.