Improving Subject-Invariant Electroencephalographic Signal Classification via Autoregressive Modeling

Sara Stalevska
University of Primorska
Faculty of Mathematics, Natural Sciences and Information Technologies
Koper, Slovenia

Peter Rogelj
University of Primorska
Faculty of Mathematics, Natural Sciences and Information Technologies
Koper, Slovenia

Abstract

Electroencephalography (EEG) recordings carry subject-specific information. In reality, we observe that inter-subject tasks are more challenging than intra-subject ones, as classification accuracy in this case is always lower due to subject variability. This gap in accuracy is a major challenge for developing scalable Brain-Computer Interfaces (BCIs). We investigated how autoregressive (AR) models' training data affects the performance of classification using innovative data and compared this with the classification of raw EEG signals. For the classification model, we selected EEGNet. We analyzed seven different types of data used to extract AR models. The tests were performed for three different tasks: eyes open versus eyes closed (EO/EC), resting versus movement, and left versus right hand movement. We found out that in some cases the gain of accuracy is considerable (EO/EC), while in some cases the changes are not statistically significant (movement tasks). We conclude that innovative signals may considerably improve classification if AR models are computed from the same kind of data that is later being classified.

Keywords

electroencephalography, autoregressive modelling, inter-subject classification, EEGNet, innovative signal, brain-computer interface.

1 Introduction

Electroencephalography (EEG) is a noninvasive biosignal acquisition technique widely used in neuroscience research and clinical applications to help identify neurological disorders and analyze cognitive processes [1]. While machine learning models can classify mental states and motor imagery from EEG signals [5], a core challenge in EEG classification is inter-subject variability. Brain signals vary significantly between individuals performing the same task. The reason may be in different brain functioning dependent on anatomical differences, brain connectivity patterns, and other physiological factors. As a result, inter-subject models (models trained to classify data from previously unseen subjects) generally yield considerably lower accuracy than intra-subject models (models trained and tested on data from the same individual).

This performance gap is particularly relevant in Brain-Computer Interfaces (BCIs), neuroscience research, and especially applications in medical diagnosis, where testing must often be conducted on subjects whose data cannot be collected in advance to train the models [4]. Traditionally, researchers use concepts like domain adaptation to alleviate this problem by artificially adapting classification models to fit new individuals. We propose an alternative approach to this problem. Instead of adapting complex models to fit individuals, we aim to reduce the use of the subject-specific information that causes the problems in the first place.

We tend to decompose information obtained from brain signals into static and dynamic components. Static ones include subject specifics, intrasubject variation of neurophysiological mechanisms, and various other, often unknown factors. The dynamic ones are expected to contain all the information of current brain activity. In our approach, the static information is contained in AR models, while the dynamic one is present in the innovative signals, which are obtained as AR residual errors.

The AR model weights capture the predictable, static information (reflecting *who* the person is), while the errors carry the dynamic, task-related information (reflecting *what* they are doing). Because knowing the static parameters of the AR model allows us to completely reconstruct the raw EEG signal just from these prediction errors, it means the errors contain all required information about EEG dynamics. Consequently, they are referred to as innovative signals.

Because subject-specific information is primarily static, using these innovative signals instead of raw EEG recordings should reduce subject-specific noise and improve inter-subject classification. We have analyzed the impact of replacing raw EEG data with innovative signals on three classification tasks: eyes open and eyes closed (EO/EC, a lower difficulty problem that often serves as a sanity check for detecting methodological problems of classification algorithms), rest and movement (T0/T1-T2, distinguishing resting from active motor tasks), and left and right hand movement (T1/T2, distinguishing different motor movement tasks performed with the left or right hand).

2 Methods

2.1 Autoregressive Modeling Strategy

To extract the dynamic information from brain activity, our method relies on fitting a univariate Autoregressive (AR) model to individual EEG signals. An AR model predicts the current value of a raw EEG signal, $y(t)$, as a linear combination of its p most recent past values [2]:

$$y(t) = \sum_{i=1}^{p} a_i y(t-i) + e(t) \tag{1}$$

This process produces two distinct components. First, it yields a set of weights, a_i, which represent the AR model itself. These parameters define a transfer function that captures the static, stable properties of the individual's brain signal. Second, it yields a residual error, $e(t)$, representing the difference between the predicted value and the actual observed value.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).
Information Society 2026, Ljubljana, Slovenia
© 2026 Copyright held by the owner/author(s).
https://doi.org/10.70314/is.2026.cogni.99

This error sequence, $e(t)$, is the innovation signal. It represents the new, unpredictable information entering the system at any given moment. Importantly, by applying the AR model weights a_i as an infinite impulse response (IIR) filter to the innovation signal $e(t)$, we can perfectly reconstruct the raw signal $y(t)$. Therefore, no dynamic task-related information is lost by transitioning from the raw signal to the innovation signal.

Conventionally, for applications like computing Granger causality, AR models are estimated from individual signal recordings or even very short window sections. In our proposed approach, we wish to obtain a static model that best describes the subject's overall signal characteristics across various states. As this part of information is expected to be static, it can be extracted from other recordings, not necessarily the analyzed signals alone. This splits the process into two steps: first, computing the AR model weights a_i from a chosen subset of training data; second, applying this fixed model as an inverse filter across the subject's remaining data to compute the isolated innovation signals $e(t)$ used for classification.

2.2 Evaluation Approach

We evaluate the improvements of the classification capabilities by comparing the classification results obtained by our proposed approach (using innovation signals) against a baseline of classification results on raw EEG data.

For classification of EEG data, we selected EEGNet [3], a widely used convolutional neural network specifically designed for EEG data. Its capabilities have been demonstrated in various EEG classification tasks, making it an ideal, reliable evaluator for our preprocessed signals. It is important that EEGNet was used with parameters proposed by its authors, not to introduce any bias by adapting to specific type of inputs.

In this phase, we used the following EEG recording types to extract the AR models:

- **Raw data:** Raw EEG signal without AR modeling.
- **Per-epoch AR:** Model fitted locally on the short epoch being classified.
- **EO model:** Model fitted on the subject's eyes-open resting data.
- **EC model:** Model fitted on the subject's eyes-closed resting data.
- **EO+EC model:** Model fitted on both resting states combined.
- **Task T1/T2:** Model fitted exclusively on active fist movement epochs.
- **Tasks and resting:** Model fitted on all task recordings (rest and movement).
- **All data:** Model fitted using all available recordings for the subject.

Finally, we apply a one-way analysis of variance (ANOVA) to the cross-validation accuracies to show whether the choice of AR training data significantly affects the classification capabilities compared to the raw baseline.

3 Results

3.1 Experimental Setup and Data

The experiments utilize the PhysioNet EEG Motor Movement/Imagery Dataset [6], containing recordings from 109 healthy subjects from 64 channels at 160 Hz. Each subject completed resting runs (EO, EC) and task runs where different sections correspond to rest (T0), left fist movement (T1), and right fist movement (T2). Note that T0, T1, and T2 segments come from the same continuous recording session.

Preprocessing consisted of high-pass filtering at 1 Hz and ICA-based artifact removal. Afterward, 19 standard 10-20 system electrodes were retained.

For the AR modeling, we selected an order of $p = 20$, following the recommendation to use a window size of 125 ms, above the lowest recommended value of 100 ms. Signals were segmented into 4-second epochs. EEGNet was trained for 30 training epochs using the Adam optimizer with a mini-batch size of 50. Model evaluation utilized a 5-fold cross-validation, repeated 6 times to improve statistical analysis.

Table 1 summarizes the mean accuracies and standard deviations across the 30 cross-validation results (5 folds×6 repetitions). Pairwise p-values comparing each strategy against the raw data baseline are included to denote statistical significance.

4 Discussion

The primary objective of this study was to analyze the capabilities of AR modeling for reducing the impact of subject variability on classification of EEG data. The empirical observations across all three tasks demonstrate that substituting raw EEG data with AR innovation signals can improve inter-subject classification when the discriminative information is concentrated in the dynamic component of the brain signal.

For the EO/EC task, sustained alpha oscillations are static properties of each of the two states, but when considered together, the differences between them are dynamic. By using the same model for extracting innovative signal from both, the classification results improve, if only the models are related to the analyzed brain states. Both, EO and EC models considerably improve the results, while training the models on other types of data, e.g., tasks, provides lower improvement. Interestingly, the EC model performs slightly better than the EO+EC model (86.58% vs. 86.48%); however, this marginal difference is likely noise, as there is no significant difference between them (comparing them provides a p-value of 0.835, which is greater than the 0.05 threshold), though both provide highly significant improvements over the raw baseline ($p < 0.001$). Whenever we use the same type of data for training AR models as for the classification, the results are close to the maximum. By removing the static subject traits, the classifier learned a general representation of the EO/EC distinction across subjects.

For the T0/T1-T2 classification task, we can observe higher mean accuracy when AR models are trained with task recordings (reaching 80.44% for Task T1/T2 and 80.09% for Tasks and resting). However, the statistical significance of these results is low, as p-values are much above 0.05, indicating that no significant improvement can be claimed over the raw baseline.

Conversely, for the T1/T2 task, the innovative signals do not show any decent improvement over raw signals. It is obvious that differences between left and right-hand tasks are subtle. It may be that these differences overlap with intra-subject variability, causing the subtraction of static data to be less influential. However, we can observe that using different types of data for AR extraction maintains stable performance across all models without severely degrading raw capabilities.

Furthermore, throughout all tests, the per-epoch AR strategy consistently worsened classification. When the AR model is fitted locally on a short 4-second segment, the model's transfer

Table 1: Mean classification accuracy (%), standard deviation, and pairwise p-values (vs. Raw data) across 30 cross-validation results.

	EO/EC		T0/T1-T2		T1/T2	
Training Data Strategy	Mean (Std)	p-val	Mean (Std)	p-val	Mean (Std)	p-val
Raw data	80.52 (3.28)	—	79.56 (2.70)	—	78.92 (2.01)	—
Per-epoch AR	77.51 (3.83)	**0.028**	75.47 (2.24)	**<0.001**	78.55 (1.96)	0.998
EO model	84.06 (3.91)	**0.004**	79.07 (2.05)	0.988	78.03 (2.46)	0.732
EC model	**86.58 (2.89)**	**<0.001**	78.31 (2.39)	0.323	78.92 (2.12)	1.000
EO+EC model	86.48 (3.17)	**<0.001**	79.24 (1.95)	0.999	78.41 (1.78)	0.983
Task T1/T2	82.96 (4.53)	0.148	**80.44 (1.92)**	0.759	**78.93 (2.20)**	1.000
Tasks and resting	82.82 (3.86)	0.209	80.09 (1.69)	0.980	**78.93 (2.05)**	1.000
All data	85.93 (3.13)	**<0.001**	79.98 (2.11)	0.995	78.90 (2.27)	1.000

Values in bold represent the best results for accuracy mean and std, while p-values are bold if there is a significant difference noted with respect to raw data ($p < 0.05$).

function absorbs the sustained task patterns within that specific window. Class-related brain processes dynamics is constant through the whole epoch, so all the AR model training data is as such included in the static AR model instead of left out in the dynamic component. The resulting innovation signal is left containing mostly noise. This finding heavily justifies our proposed two-step approach: AR models must be calculated from global or state-specific data arrays, not localized windows, to properly separate static traits from task dynamics.

In all three tasks, the best results are obtained when AR models are in the first step trained using the same kind of data as it is later analyzed in the second step. This confirms that the division between static properties and dynamic information is task-specific, allowing the analysis to focus on specific types of brain activity, i.e., the one with high enough variability in the AR training data. This observation sets the basis for future improvements that could better target specific activity and potentially improve classification accuracy even in more challenging tasks, such as T1/T2.

5 Conclusion

Replacing raw EEG recordings with innovation signals, obtained as residual autoregressive errors, can improve inter-subject classification accuracy. Training AR models across all target classes isolates static, subject-specific traits, allowing dynamic, class-specific information to be effectively distinguished. Consequently, compared to raw EEG data, innovative signals yield accuracy gains when target brain processes produce distinct signal changes that outweigh general variability in the subject's AR training data. This effect is most pronounced in the EO/EC task, whereas improvements in more complex tasks like T1/T2 are subtle. Nonetheless, these observations highlight how targeted AR training strategies can help reduce the challenges of inter-subject variability in EEG classification.

References

[1] A. Chaddad, Y. Wu, R. Kateb, and A. Bouridane, "Electroencephalography Signal Processing: A Comprehensive Review and Analysis of Methods and Techniques," *Sensors*, vol. 23, no. 14, p. 6434, 2023.

[2] C. W. J. Granger, "Investigating Causal Relations by Econometric Models and Cross-spectral Methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, 1969.

[3] C. M. Köllőd, A. Adolf, K. Iván, G. Márton, and I. Ulbert, "Deep Comparisons of Neural Networks from the EEGNet Family," *Electronics*, vol. 12, no. 12, p. 2743, 2023.

[4] 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," *Journal of Neural Engineering*, vol. 15, no. 5, p. 056013, 2018.

[5] A. Nogales and Á. J. García-Tejedor, "A Systematic Review of Electroencephalography Open Datasets and Their Usage With Deep Learning Models," *IEEE Access*, vol. 11, pp. 72384–72399, 2023.

[6] P. Rogelj, "Estimation of Task-Related Dynamic Brain Connectivity via Data Inflation and Classification Model Explainability," *Neuroinformatics*, vol. 23, no. 2, p. 33, 2025.

[7] Y. Roy, H. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: a systematic review," *Journal of Neural Engineering*, vol. 16, no. 5, p. 051001, 2019.

[8] G. Schalk, "EEG Motor Movement/Imagery Dataset," *PhysioNet*, 2009.

[9] A. Delorme and S. Makeig, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of Neuroscience Methods*, vol. 134, no. 1, pp. 9–21, 2004.