Using Temporal Features to Improve Accuracy in Multivariate Pattern Analysis (MVPA) of M/EEG Data

Fatemeh Barazesh Morgani(✉), Erin Goddard, and Gustavo Batista

University of New South Wales (UNSW), Sydney, Australia
{f.barazesh_morgani,erin.goddard,g.batista}@unsw.edu.au

Abstract. Multivariate Pattern Analysis (MVPA), or 'brain decoding' methods, have become a popular approach for analyzing time series brain recordings such as magnetoencephalography (MEG) and electroencephalography (EEG) to address experimental questions in cognitive neuroscience. Currently, the most common approach is to epoch data relative to an event of interest (e.g., stimulus onset) and apply the classification analyses to relative time points individually. This approach reveals how neural information unfolds over time but neglects any information that may be present in the temporal dynamics of brain activity. In this work, we employ several time series methods (classifiers that incorporate temporal information) with varying time window durations to analyze MEG and EEG data and compare their utility for revealing the temporal dynamics of neural information. As the window duration increases, classifier accuracy improves, but the results become less precise regarding the timing of neural information. We found that a window of ∼20 ms resulted in accurate classifiers while maintaining the temporal precision of the results. We also tested time series classifiers based on convolutional kernels, which often outperform simpler methods in benchmark datasets. However, these did not yield better performance in our datasets. Our results suggest that using a brief window, rather than a single time point, could lead to improved sensitivity for MVPA methods to detect neural representations from MEG and/or EEG data.

Keywords: Multivariate Pattern Analysis · Time series Classification · Magnetoencephalography · Electroencephalography

1 Introduction

Multivariate pattern analysis (MVPA) has become a standard approach in cognitive neuroscience to extract information from high-dimensional neuroimaging signals. Applied to time series modalities like MEG and EEG, it enables tracking representational dynamics on millisecond timescales. MVPA [1] encompasses a range of methods for analyzing neuroimaging data. In MVPA, machine learning classifiers are used to identify brain activity patterns associated with participants' experiences (e.g., stimulus) or their internal state (e.g., decision or task)

P. Bruno et al. (Eds.): HC_AIxIA_HYDRA 2025, CCIS 2830, pp. 182–195, 2026.
https://doi.org/10.1007/978-3-032-16708-8_16

(Fig. 1-A). Brain responses are recorded using M/EEG, and activation levels across experimental conditions are represented as patterns in a high-dimensional space. A classifier is trained on a subset of data to associate these patterns with specific conditions, then tested on unseen data. Successful generalization to new data indicates that brain responses contain condition-specific information (Fig. 1-B). Training and testing classifiers at each time point independently reveals when information about a stimulus or task becomes detectable, enabling the temporal mapping of sensory and cognitive processes [1]. However, most studies have relied on conventional classifiers that ignore temporal structure, despite M/EEG being inherently a multidimensional time series.

In this paper, we extend MVPA with time series classifiers to improve decoding accuracy, i.e., predicting experimental conditions from neural data. Unlike conventional methods that treat each time point in isolation, time series classifiers leverage the temporal context provided by neural activations that occur before and after a given moment. We systematically evaluate three approaches for time series classification and benchmark their performance on three M/EEG datasets, collected by our research group and external collaborators.

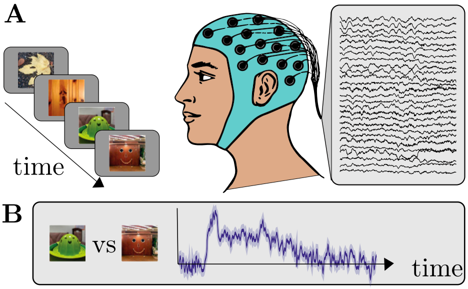

Fig. 1. MVPA in cognitive neuroscience.

Previous attempts to incorporate temporal information into MVPA have primarily relied on averaging strategies to improve the signal-to-noise ratio (e.g., [2,3]. However, they found that performance gains were marginal, with slight variations across window sizes. These findings demonstrate that temporal smoothing can offer incremental benefits, but may not fully utilize the temporal structure inherent in M/EEG signals. Our experiments demonstrate that incorporating time series classifiers into MVPA improves its sensitivity compared to traditional methods.

This paper is organized as follows. Section 1 introduces the background and related work. Section 2 describes the datasets and preprocessing. Section 3 details the methodology. Section 4 presents and discusses results, and Sect. 5 concludes and provides directions for future work.

1.1 Background and Problem Formulation

We begin with a formal notation for time series. A time series $T = (t_1, t_2, \ldots, t_m)$ is an ordered sequence of m data points. If each point $t_i \in \mathbb{R}$, the series is called *univariate*, whereas if $t_i \in \mathbb{R}^d$ is a d-dimensional vector capturing multiple variables at each time step, it is termed *multivariate*. In this work, we focus on multivariate time series.

To capture local temporal patterns, a sliding window extracts subsequences of length l from a time series T of length m:

$$S_{i,l} = (t_i, t_{i+1}, \ldots, t_{i+l-1}), \quad 1 \leq i \leq m - l + 1. \tag{1}$$

A time series classifier can be formally defined as a function $f : \mathcal{T} \rightarrow \mathcal{Y}$ that maps an entire time series to a class label from a set of classes $\mathcal{Y} = \{1, \ldots, c\}$. While this formulation is standard in time series classification, it is not directly applicable to MVPA, where the objective is to track how discriminative information unfolds over time. In this context, classification is not performed on whole sequences but on temporally localized portions of the data (Fig. 1). To address this, we define *time series MVPA* as the application of a subsequence classifier $g : \mathcal{S} \rightarrow \mathcal{Y}$, which maps a subsequence to a class label. A sliding window is used to extract subsequences across the signal, each of which is classified independently. The decoding accuracy of these subsequences is then evaluated as a function of time. The choice of window length l plays a crucial role in this analysis: larger windows capture richer contextual information and tend to improve classification accuracy, albeit at the expense of the temporal precision of the results. Conversely, smaller windows preserve temporal resolution but may fail to capture the full context necessary for robust decoding [3]. Thus, the window length introduces a trade-off between sensitivity and temporal specificity that must be carefully balanced in practice.

This work focuses on M/EEG data, typically organized into *channels*, *trials*, and *time points*. Each channel corresponds to a sensor or electrode (typically ∼64–128 electrodes for EEG, ∼150–300 sensors for MEG). Trials capture repetitions of experimental conditions (e.g., viewing a stimulus or making a decision), often several hundred per condition across sessions. Within each trial, signals are sampled at high temporal resolution (e.g., 1000 Hz), yielding hundreds to thousands of time points defined relative to task events, such as stimulus onset or a participant's behavioral response (e.g., decision time).

1.2 Related Work

This section reviews the time series classification approaches considered in this work. Middlehurst et al. [4] extended the taxonomy of Bagnall et al. [5], originally based on the type of discriminatory features exploited, to incorporate recent developments, refining the categories into distance–, feature–, interval–, shapelet–, dictionary–, convolution–, deep learning–based, and hybrid approaches.

In this paper, we focus on distance-based, feature-based, and convolution-based methods, which demonstrated the strongest performance in preliminary experiments with M/EEG data. While recent advances in multivariate time series classification have centered on deep learning models, particularly those leveraging attention mechanisms, we found that such models did not yield competitive accuracy in our setting. We attribute this to the limited size of M/EEG datasets, which are inherently constrained by the practical challenges of collecting human data over long experimental sessions. These limitations underscore the need for approaches that strike a balance between expressive power and data efficiency, motivating our focus on classical yet robust methods in this study.

Distance-Based. These methods utilize a distance function to calculate the similarity between time series. This work employs a different approach, using distance to compute a feature vector representation. This approach begins by computing a distance matrix that contains pairwise distances between all time series training samples, as illustrated in Equation (2). Each row of this matrix then serves as a feature vector representing a time series, which is subsequently used as input to a standard classifier [6].

$$\begin{bmatrix} D(TS_1, TS_1) & D(TS_1, TS_2) & \cdots & D(TS_1, TS_n) \\ \vdots & \vdots & \ddots & \vdots \\ D(TS_n, TS_1) & D(TS_n, TS_2) & \cdots & D(TS_n, TS_n) \end{bmatrix} = \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} \tag{2}$$

Paparrizos et al. [7] present a taxonomy of time series distance measures across seven categories. Lock-step measures compare the i-th points of two series and, with $O(m)$ complexity (m is the series length), are widely applied. In this work, we use Euclidean distance, the most common lock-step measure, to transform time series into feature vectors.

Distance measures for multivariate time series extend those for univariate data to quantify dissimilarity between two sequences [7]. For example, the Euclidean distance, shown in Equation (3), is computed by summing the squared differences across all time points and dimensions and then taking the square root.

$$d(X, Y) = \sqrt{\sum_{j=1}^{d} \sum_{i=1}^{n} (x_{i,j} - y_{i,j})^2} \tag{3}$$

Distance-based time series classification represents each series as a feature vector of its distances from other series, allowing standard classifiers to be applied. The k-Nearest Neighbor method leverages these distances, with 1-NN being popular for its simplicity and strong performance [8].

Gudmundsson et al. [9] first explored integrating time series distance measures with classifiers beyond k-NN, combining Support Vector Machines (SVMs) with Dynamic Time Warping. Building on this, Kate [10] studied extensively using various distance measures as SVM input features for time series classification.

Feature-Based. These techniques extract summary statistical features from time series, converting each series into a single vector to enable standard classifiers to handle temporal data [4]. The Highly Comparative Time Series Analysis (HCTSA) toolbox [11] generates over 7,700 time series features. Lubba et al. [12]

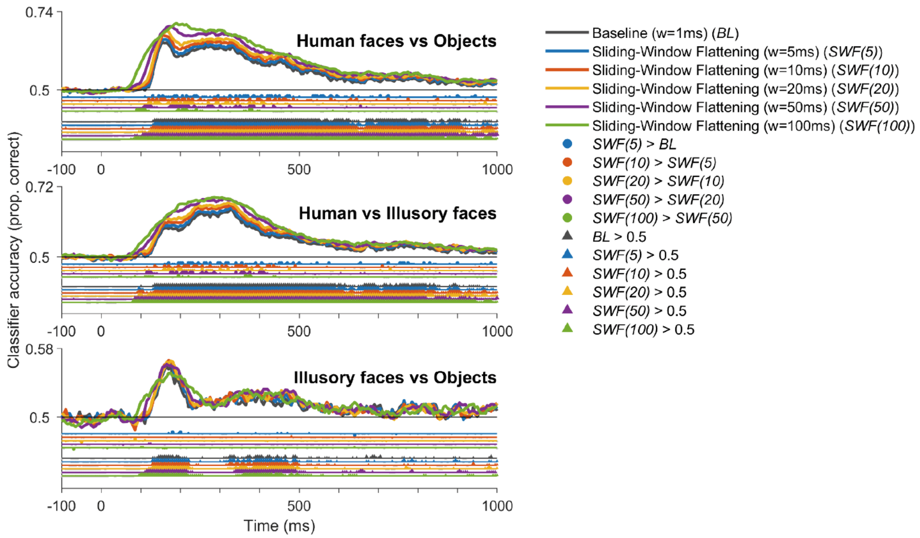

Fig. 2. Classification accuracy and temporal precision for Sliding-window flattening across three category pairs of the Pareidolia-MEG dataset. Dots below the x-axis show Bayes Factor results: colored dots mark above-chance or above-baseline performance (small dots: moderate evidence, $BF > 3$; large dots: strong evidence, $BF > 10$); Small gray dots mark moderate evidence for chance or no baseline difference ($BF < 0.3$). (Color figure online)

selected the 22 most discriminative features, called canonical time series characteristics (Catch22)[1], representing a compact and informative subset. Here, we use Catch22 as our feature-based approach for time series classification.

Convolution-Based. This approach applies convolutions between time series and specialized kernels, short subseries patterns that capture discriminative features. Each kernel produces an activation map via a sliding dot product. These maps are then summarized into feature vectors for classification [4].

Convolution (cross-correlation) applies a kernel ω of length l to a time series T via a sliding dot product, producing an activation map M, where the value at the i-th position is:

$$M_i = (T_{i,l} * \omega) = \sum_{j=0}^{l-1} T_{i+j} \cdot \omega_j \tag{4}$$

The Random Convolutional Kernel Transform (ROCKET) was proposed by Dempster et al. [13]. It generates thousands of randomly parameterized convolutional kernels. Each kernel applies two pooling operations—maximum value and proportion of positive values—and the resulting features are concatenated to form a $2k$-dimensional feature vector for k kernels.

[1] https://github.com/DynamicsAndNeuralSystems/catch22.

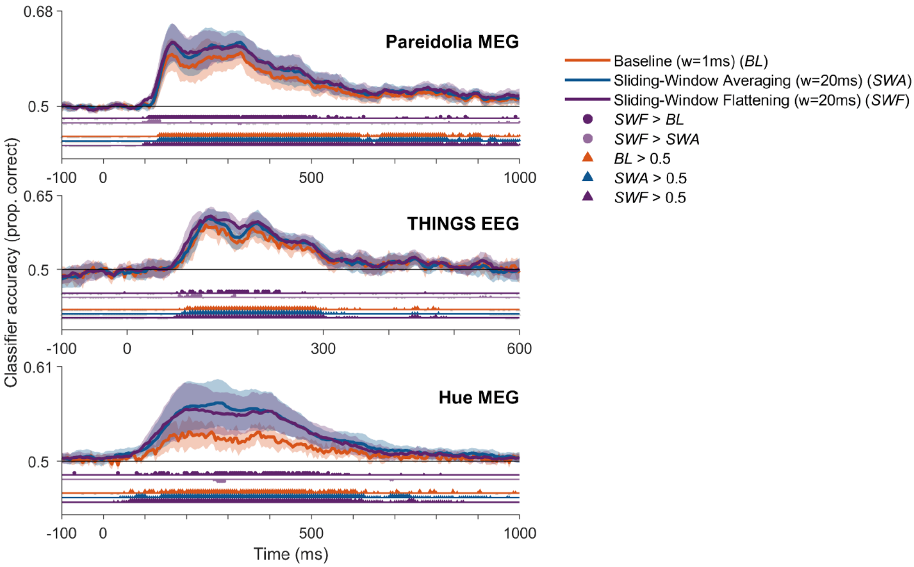

Fig. 3. Comparison of classifier performance for baseline methods across datasets (conventions as in Fig. 2). Sliding-window flattening outperforms Baseline 1 ms throughout post-stimulus. On Pareidolia-MEG and THINGS-EEG, it also exceeds Sliding-window averaging early on, while on Hue-MEG, averaging briefly outperforms flattening.

MiniROCKET[2] [14] is a faster variant of ROCKET, improving speed by over an order of magnitude while maintaining accuracy. It simplifies the method by using mostly fixed kernels and only the proportion of positive values for pooling, discarding maximum value extraction. Using 10,000 kernels, MiniROCKET efficiently generates 10,000 features.

2 Datasets

We evaluate various time series classification methods on three M/EEG datasets: *Pareidolia-MEG*, *THINGS-EEG*, and *Hue-MEG*. In each case, we performed pairwise classification of stimulus categories, since this matches the approaches used in the original works, where the patterns of relative classifier accuracy across different pairs are used to infer the brain's representation of the stimuli.

2.1 Pareidolia MEG

The first dataset contains MEG recordings (from a 160 sensor system) of brain responses to illusory faces, from [15]. Humans possess remarkable face detection and recognition abilities, supported by specialized cortical regions. This sensitive system can produce false positives, triggering the perception of faces in inanimate objects, such as trees or fruit—a phenomenon known as face pareidolia.

[2] https://github.com/angus924/minirocket.

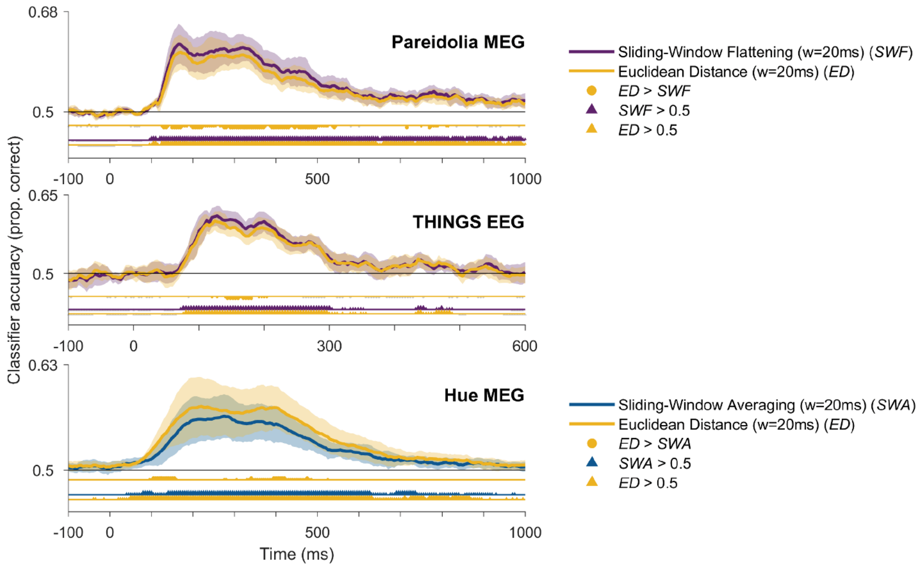

Fig. 4. Comparison of classifier performance for Euclidean Distance across datasets (conventions as in Fig. 2). On Pareidolia-MEG, Sliding-window flattening outperforms Euclidean Distance throughout post-stimulus, whereas on THINGS-EEG this advantage is brief. On Hue-MEG, Euclidean Distance outperforms Sliding-window averaging.

21 participants viewed 96 visual stimuli divided into three categories: 32 illusory faces, 32 matched non-face objects, and 32 human faces. Stimuli were presented in random order across six runs per participant, each containing four repetitions of all 96 stimuli, yielding a total of 2,304 trials (24 repetitions per stimulus).

The original study implemented cross-decoding using leave-one-exemplar-out cross-validation, where a linear discriminant analysis classifier was trained to distinguish between two categories (e.g., human faces vs. objects) using all exemplars except one from each category, then tested on the held-out exemplars. For our analysis, we used MEG recordings from the 9 participants for which head shape information was available. We calculated pairwise classification performance across the three stimulus categories using the same cross-validation scheme, meaning that all trials corresponding to the held-out stimulus exemplar were excluded from training and used for testing, but employed a ridge regression classifier instead.

2.2 THINGS EEG

The second dataset contains EEG recordings (from a 64-channel system) of brain responses to a systematic collection of objects and concepts designed to understand how the human brain categorizes and represents objects, from [16]. The full dataset comprises EEG recordings from 50 participants who viewed 22,248 images depicting 1,854 object concepts, each presented twice. In separate validation data, a further 200 images were repeated 12 times across 12 sequences. The original study employs a leave-one-sequence-out cross-validation procedure and trains a regularized linear discriminant classifier to discriminate between images.

We used the validation data with 12 repeats per image, as it is most similar to a traditional M/EEG design. To improve computational efficiency, we reduced the dataset by selecting EEG recordings from 10 participants for 20 object concepts, using 200 validation images (10 per concept), making it comparable in size to the other datasets. Pairwise classification performance was then computed using a ridge regression classifier with the same cross-validation scheme, ensuring that all trials from the held-out sequence were excluded to prevent image overlap between training and testing.

2.3 Hue MEG

The third dataset contains MEG responses (275-sensor system) to 14 different hues from 8 participants [17], designed to investigate how neural responses to color evolve over time. Stimuli were large circular blobs presented 84 times in counterbalanced order across two tasks: color categorization and color discriminability. Linear SVM classifiers were trained to discriminate stimulus hue within each task using leave-one-trial-out cross-validation.

In our study, we used MEG responses from 8 participants to 14 hues and computed pairwise classification, combining data from both tasks. We used the same cross-validation to split training and test data and a ridge regression classifier to compute accuracy.

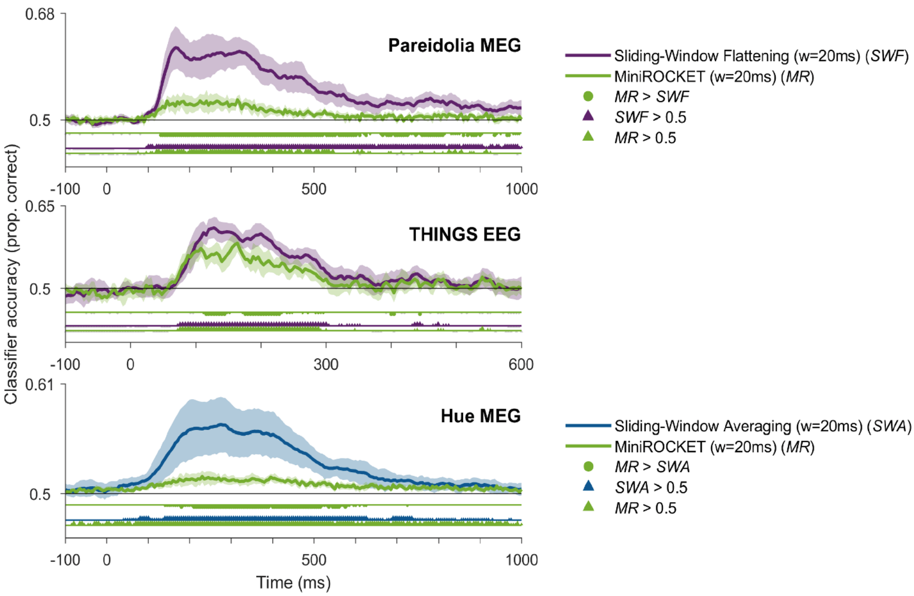

Fig. 5. Comparison of classifier performance for MiniROCKET across datasets (conventions as in Fig. 2). Sliding-window methods consistently outperform MiniROCKET.

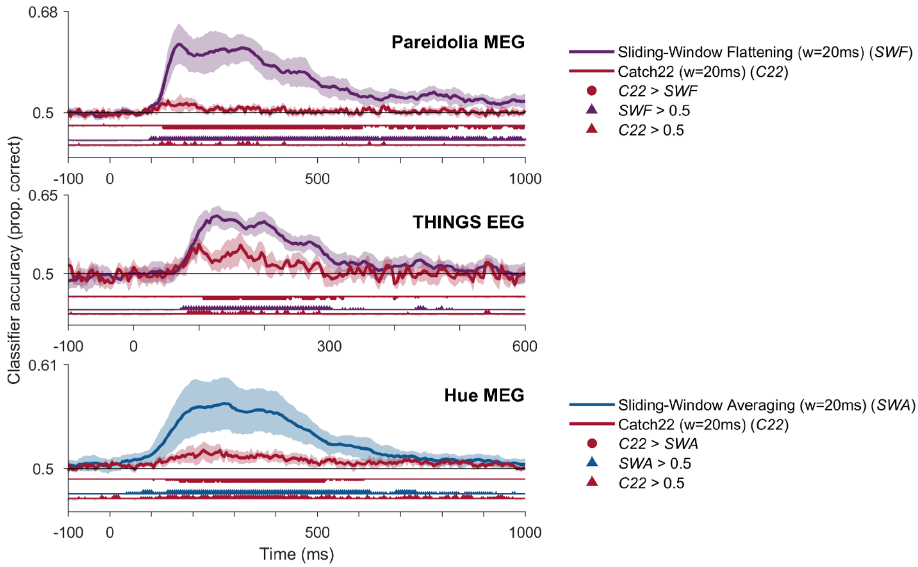

Fig. 6. Comparison of classifier performance for Catch22 across datasets (conventions as in Fig. 2). Sliding-window methods consistently outperform Catch22.

2.4 Data Preprocessing

Preprocessing of M/EEG data was performed using Brainstorm ([18], http://neuroimage.usc.edu/brainstorm), as described previously for the Hue-MEG data [17]. All M/EEG functional data were notch filtered to remove line frequency interference (50, 100, 150, and 200 Hz for THINGS and Pareidolia data, acquired in Australia, and 60, 120 and 180 Hz for Hue data, acquired in Canada), then bandpass filtered (0.2–200 Hz, using an even-order linear phase FIR filter). For MEG data (Pareidolia and Hue), cardiac and eye blink artifacts were removed using signal space projection (SSP) [19]. For THINGS-EEG data, eye blink artifacts were removed using independent component analysis (ICA) [20]. In each case, 1000 Hz functional data were extracted for each trial: from -100 to 1000 ms (MEG) or 600 ms (EEG) relative to stimulus onset. After epoching, whole-head EEG data were used in subsequent analyses.

For MEG data, we applied a further step, using minimum norm source reconstruction of the trial data to correct for participant movement between recordings during the same session. This is described elsewhere for the Hue-MEG dataset [17]. For the Pareidolia-MEG dataset, we employed a similar procedure, but since anatomical MRI images were unavailable, we created individual head models from head shape data, as in [21]. In source reconstruction, the recordings of 160 or 275 sensors are used to estimate responses from 15,000 cortical sources. To reduce this redundancy, we transformed the source-reconstructed data using principal component analysis (PCA), retaining data from the first n components that accounted for 99.99% of the variance across sources.

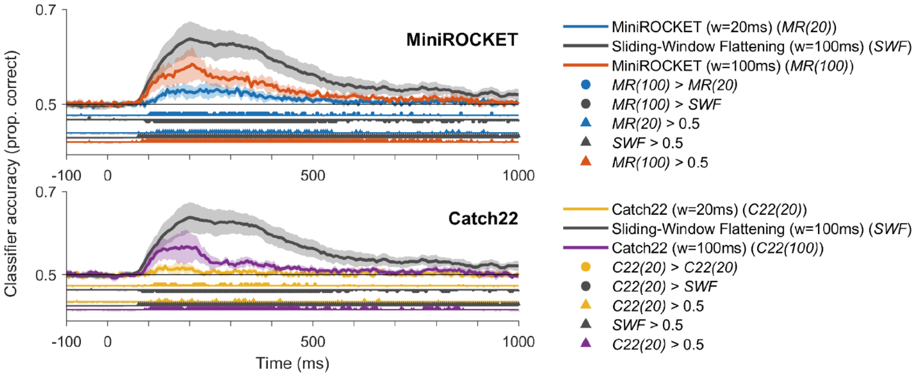

Fig. 7. Classifier performance of MiniROCKET and Catch22 on Pareidolia-MEG with longer time windows (conventions as in Fig. 2). Both methods perform better at 100 ms compared to 20 ms, but Sliding-window flattening (100 ms) still outperforms them.

3 Methodology

We evaluate six decoding approaches: three baselines that do not explicitly model temporal dependencies, and three time series classifiers that capture sequential structure. All methods operate on M/EEG trials that are aligned to stimulus onset and are evaluated using sliding windows of varying durations.

Baseline 1 ms. The standard MVPA approach applies classification independently at each time point, ignoring temporal context.

Sliding-window averaging. Averages samples within a brief temporal segment to compute summary features (here, the mean). This preserves dimensionality but collapses temporal dynamics into static values [5,22].

Sliding-window flattening. Concatenates all samples within a window into the classifier input (e.g., a 20 ms window with 150 channels yields 150×20 features). This increases dimensionality and allows richer local dynamics to be captured, though without encoding temporal order [1].

These baselines illustrate complementary trade-offs: pointwise decoding neglects temporal information, averaging incorporates context but discards dynamics, and flattening captures more temporal features without sequential order.

To overcome these limitations, we additionally assess three established families of time series classifiers, distance-based, feature-based, and convolution-based, which have been shown to perform strongly in prior benchmarks [4]. These methods explicitly model the order and dependencies among time points, designed to capture discriminative temporal patterns in MEG/EEG signals.

For all approaches, we employ `RidgeClassifierCV`, a linear classifier with L_2 regularization, from `sklearn.linear_model`. Regularization strength is selected automatically via cross-validation across a logarithmic grid of parameters ($\alpha \in$

$(-3, 3, 10)$). This yields a favorable bias–variance trade-off, typically outperforming 1-NN while training more efficiently than SVMs in our setting.

4 Experiments and Results

To support reproducible research, we created a public repository[3]. Our classifiers were trained to separate pairwise categories of images in the Pareidolia-MEG and THINGS-EEG datasets, as well as hue pairs in the Hue-MEG dataset. We assess three baseline methods and three time series classification methods—Euclidean Distance, MiniROCKET, and Catch22—with different window lengths.

We treated the time window length as a hyperparameter. We evaluated the classification accuracy and temporal precision across different time window sizes ($t = 5, 10, 20, 50, 100$ ms) for Sliding-window flattening on three category pairs of the Pareidolia-MEG dataset (Human faces vs. Objects, Human faces vs. Illusory faces, and Illusory faces vs. Objects). Figure 2 shows that classification accuracy increases with larger windows for Human faces vs. Objects and Human faces vs. Illusory faces. For Illusory faces vs. Objects, accuracy rises for windows of 5, 10, and 20 ms, but decreases for 50 and 100 ms. Additionally, as the window of time increases, there is a decrease in temporal precision with which the results can be used to draw inferences regarding the underlying neural processes. This is seen in Fig. 2 where the different dynamics or shapes of decoding accuracy curves across classifications become less distinct for the longest windows. Overall, a window of approximately 20 ms provides an effective balance between maximizing accuracy and preserving temporal precision.

Here, we used a Bayes Factor [23], which quantifies how much more likely the data are under one hypothesis (e.g., H_1) than another (e.g., H_0). It provides a continuous measure of evidence, unlike the traditional frequentist approach.

Next, we compare the average pairwise classification accuracy of the three baseline methods using a 20 ms sliding window across all datasets. As shown in Fig. 3, Sliding-window flattening achieves the highest accuracy for the Pareidolia-MEG and THINGS-EEG datasets, while Sliding-window averaging performs best for the Hue-MEG dataset. These baselines are therefore selected for subsequent comparisons with time series classification methods.

We then evaluate three time series methods—Euclidean Distance (Fig. 4), MiniROCKET (Fig. 5), and Catch22 (Fig. 6)—with a 20 ms sliding window, and compare them against their best baselines. Based on Fig. 4, Sliding-window flattening outperforms Euclidean Distance on the Pareidolia-MEG and THINGS-EEG datasets, whereas Euclidean Distance surpasses Sliding-window averaging on the Hue-MEG dataset. Furthermore, as shown in Figs. 5 and 6, baseline methods significantly outperform MiniROCKET and Catch22, suggesting that these time series approaches may require longer time windows to be more effective.

To test whether the MiniROCKET and Catch22 approaches can outperform the baseline methods at longer time windows, we compared classifier accuracy using 20-ms and 100-ms windows for both methods on the Pareidolia-MEG

[3] https://github.com/F-Barazesh-Morgani/MVPA_TemporalFeatures.

dataset. As shown in Fig. 7, although MiniROCKET and Catch22 with 100 ms windows outperform their 20 ms counterparts, Sliding-window flattening with a 100 ms window still outperforms both, even at 100 ms.

5 Conclusion and Future Work

The results of computing the average pairwise classification accuracy across stimuli in the Pareidolia-MEG, THINGS-EEG, and Hue-MEG datasets demonstrate that using a brief time window—as in Sliding-window flattening, which incorporates multiple adjacent time points as input features—enables the classifier to capture richer temporal dynamics compared to traditional single-time-point approaches (Baseline 1 ms). This strategy enhances the sensitivity of MVPA methods for detecting neural representations in both MEG and EEG data across all three datasets. However, Sliding-window flattening does not account for the temporal order of observations within each window.

To address this limitation and model the sequential structure of the data, we employed three time series classification methods—representing distance-based, feature-based, and convolution-based approaches—that explicitly capture dependencies among time points in dynamic brain signals.

Among these, we found that applying Euclidean distance with a 20 ms window to transform the time series into feature vectors improved the sensitivity of MVPA compared to traditional methods (Baseline 1 ms), which compute classifier accuracy at individual time points. This improvement was observed across all three datasets, and in the Hue-MEG dataset, the Euclidean distance method even outperformed Sliding-window flattening and averaging.

While methods like MiniROCKET and Catch22 often outperform simpler approaches in other domains (e.g., gesture recognition), they did not yield better performance in our analysis when using a 20 ms window. These methods typically benefit from longer time windows, but extending the window length would compromise temporal precision, which is critical for our research goals.

As future work, we will explore deep learning methods such as LSTMs, RNNs, TCNs, and Transformers, combined with data augmentation, to better capture temporal dependencies and dynamic patterns in M/EEG data.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Grootswagers, T., Wardle, S.G., Carlson, T.A.: Decoding dynamic brain patterns from evoked responses: a tutorial on multivariate pattern analysis applied to time series neuroimaging data. J. Cogn. Neurosci. **29**(4), 677–697 (2017)
2. Ashton, K., et al.: Time-resolved multivariate pattern analysis of infant EEG data: a practical tutorial. Dev. Cogn. Neurosci. **52**, 101044 (2022)

3. Wu, H., et al.: Decoding n400m evoked component: a tutorial on multivariate pattern analysis for op-meg data. Brain Sci. **12**(5), 1–19 (2022)
4. Middlehurst, M., Schäfer, P., Bagnall, A.: Bake off redux: a review and experimental evaluation of recent time series classification algorithms. Data Min. Knowl. Discov. (2023)
5. Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E.: The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min. Knowl. Discov. **31**(3), 606–660 (2017)
6. Abanda, A., Mori, U., Lozano, J.A.: A review on distance based time series classification. Data Min. Knowl. Discov. **33**(2), 378–412 (2019)
7. Paparrizos, J., et al.: A survey on time series distance measures. ACM Comput. Surv. **55**(4) 2023
8. Ding, H., Trajcevski, G., Scheuermann, P., Wang, X., Keogh, E.: Querying and mining of time series data: experimental comparison of representations and distance measures. Proc. VLDB Endow. **1**(2), 1542–1552 (2008)
9. Gudmundsson, S., Runarsson, T.P., Sigurdsson, S.: Support vector machines and dynamic time warping for time series. In: Proceedings of the Joint Conference Neural Networks (IEEE World Congress on Computational Intellignce), pp. 2772–2776. IEEE (2008)
10. Kate, R.J.: Using dynamic time warping distances as features for improved time series classification. Data Min. Knowl. Discov. **30**(2), 283–312 (2016)
11. Fulcher, B.D., Jones, N.S.: hctsa: a computational framework for automated time series phenotyping using massive feature extraction. Cell Syst. **5**(5), 527-531.e3 (2017)
12. Lubba, C.H., et al.: catch22: canonical time series characteristics selected through highly comparative time series analysis. Data Min. Knowl. Discov. **33**(6), 1821–1852 (2019)
13. Dempster, A., Petitjean, F., Webb, G.: Rocket: exceptionally fast and accurate time series classification using random convolutional kernels. Data Min. Knowl. Discov. **34**, 1454–1495 (2020)
14. Dempster, A., Schmidt, D.F., Webb, G.I.: Minirocket: a very fast (almost) deterministic transform for time series classification. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 248–257. ACM (2021)
15. Wardle, S.G., Taubert, J., Teichmann, L., Baker, C.I.: Rapid and dynamic processing of face pareidolia in the human brain. Nat. Commun. **11**, 4518 (2020)
16. Grootswagers, T., Zhou, I., Robinson, A.K., Hebart, M.N., Carlson, T.A.: Human EEG recordings for 1,854 concepts presented in rapid serial visual presentation streams. Sci. Data **9**(1), 3 (2022)
17. Goddard, E., Mullen, K.T.: Temporal evolution of color representations measured with magnetoencephalography reveals a "coarse to fine" dynamic. J. Cogn. Neurosci. 1–25 (2025)
18. Tadel, F., Baillet, S., Mosher, J.C., Pantazis, D., Leahy, R.M.: Brainstorm: a user-friendly application for MEG/EEG analysis. Comput. Intell. Neurosci. **2011**, 879716 (2011)
19. Uusitalo, M.A., Ilmoniemi, R.J.: Signal-space projection method for separating MEG or EEG into components. Med. Biol. Eng. Comput. **35**(2), 135–140 (1997)
20. Ablin, P., Cardoso, J., Gramfort, A.: Faster independent component analysis by preconditioning with Hessian approximations. IEEE Trans. Sign. Process. **66**(15), 4040–4049 (2018)

21. Goddard, E., Carlson, T.A., Woolgar, A.: Spatial and feature-selective attention have distinct, interacting, effects on population-level tuning. **34**(2), 290–312
22. Fulcher, B.D., Jones, N.S.: Highly comparative feature-based time series classification. IEEE Trans. Knowl. Data Eng. **26**(12), 3026–3037 (2014)
23. Kass, R.E., Raftery, A.E.: Bayes factors. J. Am. Stat. Assoc. **90**(430), 773–795 (1995)