Evaluating Contrastive Learning Approaches for Robust P300 Detection in Noisy EEG Environments

MSc Research Project
Data Analytics

Muhammad Awais Quarni
Student ID: 23373326

School of Computing
National College of Ireland

Supervisor: Dr.Hicham Rifai

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Muhammad Awais Quarni
Student ID:	23373326
Programme:	Data Analytics
Year:	2025
Module:	MSc Research Project
Supervisor:	Dr.Hicham Rifai
Submission Due Date:	11/08/2025
Project Title:	Evaluating Contrastive Learning Approaches for Robust P300 Detection in Noisy EEG Environments
Word Count:	XXX
Page Count:	16

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Muhammad Awais Quarni
Date:	10th August 2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Evaluating Contrastive Learning Approaches for Robust P300 Detection in Noisy EEG Environments

Muhammad Awais Quarni
23373326

The extraction of the P300 event-related potential (ERP) electroencephalogram (EEG) data is critical for the development of robust brain-computer interface (BCI) systems, especially in realistic, noisy environments. The study investigates the effectiveness of modern contrastive self-supervised learning methods (SimCLR, TS-TCC, and VICReg) in comparison to a conventional supervised convolutional neural network (CNN) for single-trial P300. The experiments were conducted on the AMBER dataset, a multi-subject, multi-session EEG corpus explicitly collected under both clean and ecologically valid noisy conditions. Due to computational constraints, the analysis was limited to the RSVP tasks X1(clean) and X4 (noisy) from all subjects. EEG epochs were preprocessed, normalized, and augmented, and feature representations were learned via contrastive pretraining. The resulting embeddings were evaluated using downstream classifiers with model performance assessed by accuracy, ROC-AUC, and F1-Score for the rare target (P300) class. Results show that all models struggled to robustly detect P300 targets under real-world noise and data imbalance, with low recall and ROC-AUC across the board. Among the contrastive models, VICReg achieved the highest recall (0.556) and F1-score (0.225) for targets. This finding suggests that variance-invariance-covariance regularisation may enhance representation learning for challenging EEG tasks, but also underscores the persistent difficulty of P300 detection in noisy, resource-limited scenarios. This study provides a transparent benchmark and highlights key challenges for future research with more comprehensive analysis, including all tasks and a larger sample size as well as advanced augmentation and computational strategies are recommended to further advance robust P300 detection in real-world BCI applications.

1 Introduction

Electroencephalography (EEG) is a widely used non-invasive (a procedure that does not require instrumentation through skin or into a body opening) method that is used to measure electrical activity in the brain. It plays a pivotal role in both neuroscience research and practical clinical or commercial applications. One of the most important responses captured by EEG is the P300 component.

The P300 is an event-related potential (ERP) that typically appears as a positive voltage peak around 300 milliseconds after a person perceives an important, rare, or surprising stimulus Alsufyani (2020), Jin et al. (2021). P300 based paradigms are a set of methods that captures brain signal called the P300 wave in the EEG signal when someone is asked to pay attention to a specific target item (like images, a certain letter, sound) among many non-target items that displayed at a certain pace consecutively, this process known as Rapid Serial Visual Presentation (RSVP). The brain triggers a “P300” signal each time the target is detected.

P300 is an ERP marker linked to cognitive processes such as attention and decision making that is usually elicited using “oddball” paradigms, where a rare target is mixed among frequent standard stimuli. The latency of P300 is between 250 and 500 ms after stimulus onset, with a mean of around 300 ms. The amplitude is measured at parietal scalp electrodes, and the extraction of the P300 from the EEG is performed by time locking to stimulus onset and averaging across trials. This process is challenging because single-trial extraction is much harder due to the low signal-to-noise ratio.

The P300 is widely used in brain-computer interface (BCIs) for communication (e.g, spelling devices for paralyzed patients), lie detection, neuromarketing, cognitive assessment, such as in ADHD, Alzheimer’s or ASD research, and even in e-sports for player engagement analysis Du et al. (2023), Das and Mahadevappa (2023). Its reliability in indicating cognitive processing of meaningful events makes it a gold standard for many applied neuroscience tasks. However, the detection of P300 in a single trial and a real-world noisy environment remains a major challenge. EEG signals are often contaminated by various sources of noise, such as muscle activity, eye blinks, or movement artifacts Das and Mahadevappa (2023)

Traditional machine learning classification approaches, such as LDA and SVM, are used to detect the P300 signals, but they rely on manual feature selection and multiple-trial averaging, limiting their use in rapid, real-world BCI systems Cheng et al. (2023), Alsufyani (2020). Deep learning models, especially CNN-based architectures such as EEGNet and PSAEEGNet, have recently improved detection performance by learning features directly from raw EEG data Yuan et al. (2024), Tajmirriahi et al. (2022). Advanced attention-based architectures like PSAEEGNet and Autoencoder-based models have further enhanced detection accuracy in RSVP paradigms Ditthapron et al. (2019). However, these advanced models still struggle with the following:

- Generalizing across users and sessions.
- Robustness to noise and artifacts.
- The requirement for large, labelled datasets (which are expensive to collect).

Self-supervised and contrastive learning approaches have been used as promising tools for addressing these limitations Zhu et al. (2023), Weng et al. (2025). These models learn useful features from unlabeled or weakly labelled data by maximizing the similarity between positive pairs (augmented views of the same trial) and minimizing similarity with negative pairs (different trials). They have the potential to improve P300 extration in realistic, noisy environments where labelled data is scarce or expensive.

Most studies rely on clean, controlled laboratory datasets, limiting their applicability to real-world environments. Despite the development of contrastive learning techniques showing promise in many domains, they have not been applied to detect the P300 in a noisy, real-life EEG. The research aims to address two key gaps:

- First, there is a lack of systematic evaluation of contrastive learning frameworks for P300 detection in comparison with strong classical and deep learning baselines (e.g., attention-based CNNs, hybrid feature models).
- Second, using a self-supervised learning method on large-scale publicly available datasets, such as RSVP based EEG data, is rare.

2 Related Work

The P300 component of the event-related potential (ERP) signals remains one of the most studied signatures in cognitive neuroscience due to its key role in attention, working memory, and decision-making processes Luck (2014). Detection and extraction of the P300 in EEG signals are fundamental to brain-computer interface (BCI) technology, particularly in applications such as assistive communication (spellers), neuromarketing, and cognitive assessment Tajmirriahi et al. (2022). However, the extraction of reliable P300 signals in real-world settings remains challenging, as these environments are characterised by high noise, class imbalance, and variability across different subjects, posing significant technical and methodological challenges.

Recent research has shifted from controlled laboratory environments to real-world, ecologically valid settings, where the robustness of algorithms to noise and their ability to generalise across sessions and individuals become important Awais et al. (2023), Won et al. (2022). This shift has motivated the exploration of both classical machine learning techniques and advanced deep learning strategies, including emerging self-supervised and contrastive learning methods, all aiming to overcome the limitations of traditional supervised models, enhancing the reliability and adaptability of P300 detection systems in more realistic, variable conditions.

2.1 Classical Approaches and Feature Engineering

Early work on P300 extraction relied extensively on manual feature extraction, including time domain, frequency domain, and time frequency features, which were then input into traditional machine learning classifiers such as support vector machines (SVM) and random forests (RF). Peketi et al. (2023), Sarraf et al. (2023) proposed a universal joint feature extraction method using multitask autoencoders that combined classical and deep learning paradigms to improve cross-subject generalization. Although these methods offered incremental improvements, their performance often diminished under conditions of significant inter-subject variability and nonstationary noise problems that are inherent to EEG data in real-world use Saeidi et al. (2021). A persistent limitation of classical feature engineering is its dependence on expert domain knowledge as well as its inability to fully capture the non-linear and hierarchical representation that deep learning models can learn automatically. Nevertheless, these approaches laid the foundation for the critical analysis of deep learning models and motivated the search for more generalizable solutions.

2.2 Supervised Deep Learning for P300: Strengths and Limitations

Deep neural networks, especially convolutional neural networks (CNNs), have become the standard approaches for modern P300 detection due to their ability to automatically extract complex hierarchical features directly from raw or minimally preprocessed EEG data Lawhern et al. (2018). Recent architectures such as PSAEEGNet Yuan et al. (2024) and ERPENet Ditthapron et al. (2019) introduced pyramid squeeze attention mechanisms and computationally efficient blocks to enhance single-trial classification performance.

These innovations have demonstrated significant improvements over traditional methods, particularly in RSVP (Rapid Serial Visual Presentation) paradigms, underscoring the effectiveness of deep learning models in addressing the challenges of real-time EEG analysis. However, even state-of-the-art supervised deep models face significant limitations: Dependence on large, labelled datasets: Training deep models from scratch typically requires large amounts of labelled data, which is costly and often infeasible in real-world settings Awais et al. (2023). This reliance on models for scalability and generalizability, especially important in applications where labelled EEG data is scarce or difficult to obtain. Poor Cross-Subject Generalisation: Despite improvements, models' performance still degrades significantly when applied across different subjections or sessions, largely due to individual variability and nonstationary EEG noise Ditthapron et al. (2019), Tajmirriahi et al. (2022). Addressing this issue remains a critical area for further research to enable more robust and transferable EEG-based P300 detection systems. Sensitivity to Real World Noise: Even with advanced architecture, robustness to movement artifacts, environmental noise, and nonstationary conditions remains a core challenge Yuan et al. (2024). These challenges have directly motivated the investigation of self-supervised and contrastive learning approaches as potential solutions. By leveraging unlabeled data and learning more robust feature representations, these methods aim to improve generalization across subjects and sessions, as well as enhance resilience to real-world noise and artifacts.

2.3 Self-Supervised and Contrastive Learning in EEG-P300

Self-supervised learning aims to learn useful data representations without explicitly relying on labels, typically by solving "pretext" tasks based on the data itself Le-Khac et al. (2020). In the context of EEG, self-supervised learning provides a promising approach for achieving robust, label-efficient and generalization P300 detection, especially under conditions of limited or noisy labelled data Eldele et al. (2021), Bardes et al. (2021). Contrastive learning, a subset of SSL, encourages representations of similar samples (positive pairs) to be close in the latent space, while separating dissimilar samples (negative pairs). For EEG, this is often realised by augmenting the same trial in different ways—through jittering, cropping, or permuting channels—and maximizing agreement between these views Zbontar et al. (2021).

2.4 Recent Methods and Evidence

TS-TCC (Time Series Contrastive and Contextual Coding, Eldele et al. (2021) introduced a dual view of contrastive approach that combines temporal cropping with channel permutations to achieve significant improvements in robustness and generalization on time series data, including EEG datasets. Barlow Twins Bardes et al. (2021), Zbontar et al. (2021) further advanced the self-supervised field by incorporating redundancy reduction and variance-invariance-covariance regularization. These approaches facilitate the learning of compact, invariant representations that open new ways for effective and efficient EEG analysis. However, most of the literature remains focused on controlled benchmarks, with a few studies providing direct comparisons on real-world, multi-subject, and noisy P300 EEG data Yuan et al. (2024), Awais et al. (2023). Recent studies suggest that contrastive methods can approach or even surpass supervised baselines in certain low-label or transfer learning scenarios; their performance is sensitive to data quality, augmentation strategies, and hyperparameter tuning Tajmirriahi et al. (2022).

2.5 Gaps and Future Needs

Despite their promising potential, the most contrastive and self-supervised approaches for EEG are still at the experimental stage. Their effectiveness in real-world, noisy, cross-subject P300 detection, especially in RSVP paradigms, remains an open question. Further, the interpretability of learned features and the integration of domain-specific augmentations (e.g., time-frequency transforms) are active research areas Ditthapron et al. (2019), Tajmirriahi et al. (2022). Addressing these challenges is essential for translating self-supervised methods into robust, real-world EEG applications.

2.6 Summary

In summary, while deep learning has significantly advanced the state of the art in P300 detection, the real-world deployment of BCIs faces challenges such as noise, cross-subject variability, and limited data. Self-supervised and contrastive learning represents a promising new direction, but remains unexplored for realistic P300 Scenarios. There is a critical need for a comprehensive, systematic evaluation of both supervised and self-supervised models under ecologically valid, multi-subject, and noisy conditions, precisely the gap addressed by this research.

3 Methodology

The complete data mining procedure in this study of evaluating contrastive learning approaches for robust P300 Detection in noisy EEG environments follows a set of steps based on CRISP-DM methodology. All these steps are systematically presented in the Figure.

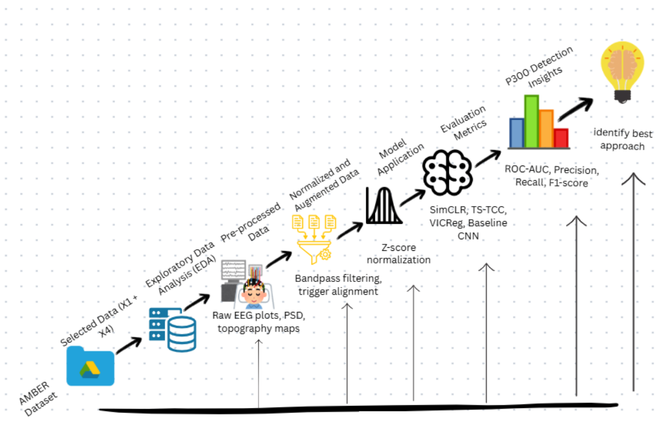


Figure 1: P300 extraction methodology

This study leverages the AMBER dataset, a benchmark specifically designed for evaluation of brain computer interface (BCI) models under authentic, real-world noise environmnt. The dataset consists of high-resolution EEG recordings from participants in four separate sessions incorporating both clean and noisy experimental scenarios.

The AMBER dataset serves as a rigorous testbed for evaluating and comparing contrastive learning methods against traditional supervised approaches for P300 detection. It's a realistic design enables a meaningful assessment of model robustness, signal discrimination, and the ability to lean effectively even with limited labeled data.

3.1 The dataset comprises

- EEG signals recorded using a 32 channel ANT-Neuro eego sports system following the 10-20 internation electrode placement scheme.
- Data samples were taken at 1000 Hz, with CPz set as a reference electrode.
- The configuration ensures high temporal resolution and broad scalp coverage which is crucial for robust detection of the P300 event-related potential (ERP).

For the baseline tasks signal were collected while "Eye Open", "Eyes Movement", "Eyes Closed". For the RSVP experiments, images were shown at a frequency of 4 Hz. The subject must identify infrequent "target" images (e.g., cars) among a stream of "non-target" distractors (e.g., flowers, animals, instruments). As shown in Figure. Awais et al. (2023).

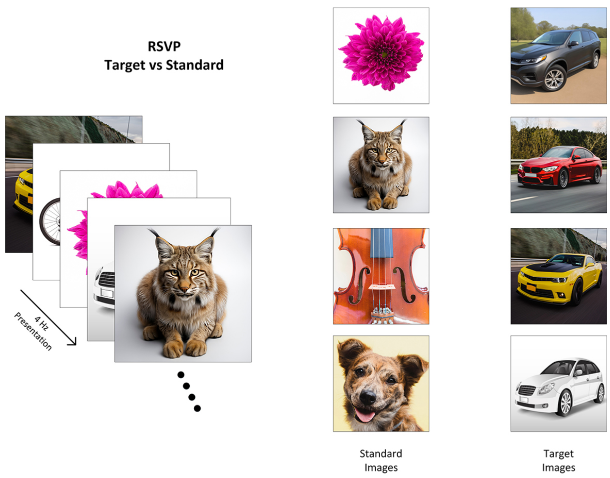


Figure 2: Overview of the AMBER dataset recording setup and protocol. Reproduced from Awais et al. (2023) under the terms of the Creative Commons CC BY license.

A key differentiator of the AMBER dataset Awais et al. (2023) is its inclusion of realistic, naturalistic artifact conditions like body movement, talking, head movement. This simulates a real-world scenario where EEG signals are subject to non-ideal, ecologically valid noise sources. It makes the dataset especially challenging and relevant for robust algorithms development.

Table 1: AMBER Dataset Key Features

Feature	Description
Subjects	10 healthy adults
Sessions	4 per subject (multi-session, multi-day)
EEG Channels	32 channels, 10–20 system, CPz reference
Sampling Rate	1000 Hz
Tasks	Baseline (eyes open/closed/movement), RSVP (with/without artifacts)
Noise Conditions	Body movement, speech, head movement (artifact-included RSVP)
Data Formats	EDF (signals), CSV (annotations)
Stimulus Presentation	RSVP, 4 Hz, target/non-target identification

3.2 Exploratory Data Analysis

Before preprocessing, it is crucial to understand the spatial configuration of the EEG electrodes. The AMBER dataset employs an international 10–20 montage, which is a globally accepted electrode placement system in EEG research. This layout provides a structure and reproducible method for capturing electrical activity from different brain regions, like the frontal, central, parietal, and occipital lobes.

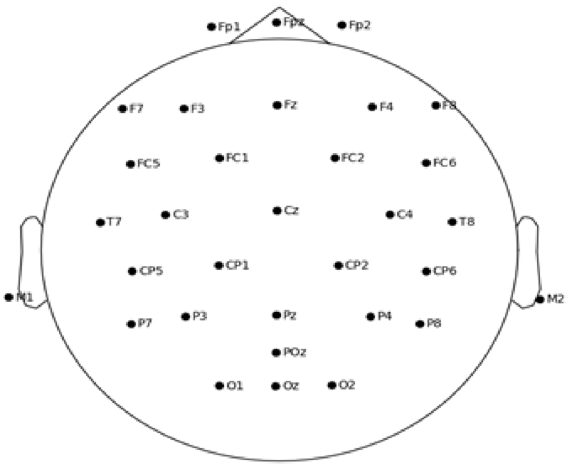


Figure 3: The topographic plot illustrates all 32 electrodes

The topographic plot (see Figure 3) illustrates all 32 electrodes used in the dataset with clear labelling for each site (e.g., Fp1, Fz, Fc, Pz, Oz, etc.). This assures that the data covers the entire scalp and can be used for both spatial and temporal analysis of brain activity.

The raw EEG plot displays unfiltered voltage signals recorded from all 32 channels over the first 10 seconds of the experiment. It's an essential first step to understand the quality of the data, the presence of biological rhythms, and any large artifacts like blinks, muscle activity, and sudden jumps. This helps to identify common noise patterns and plan appropriate preprocessing steps (e.g., filtering, artifact rejection).

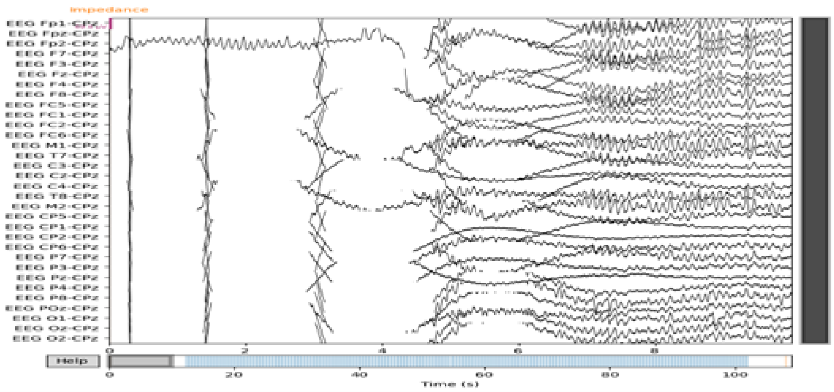


Figure 4: Raw EEG Signal Plot (First 10 Seconds)

In Figure 4 most channels display continuous, oscillatory activity in the microvolt range, with some showing large transients due to movement or eye blinks. The "impedance" trace at the top indicates periods when impedance was measured and is often visible at the start of a recording. Overall, the data exhibits typical EEG properties but also clear evidence of artifacts, underlining the need for careful preprocessing.

The Power Spectral Density (PSD) plot summarizes the distribution of signal power across different frequencies (0-50Hz) for all EEG channels before any filtering is applied. It helps to quantitatively assess which frequency bands are prominent in the raw data and detect slow drifts (high power at low frequencies) and high frequency noise or contamination.

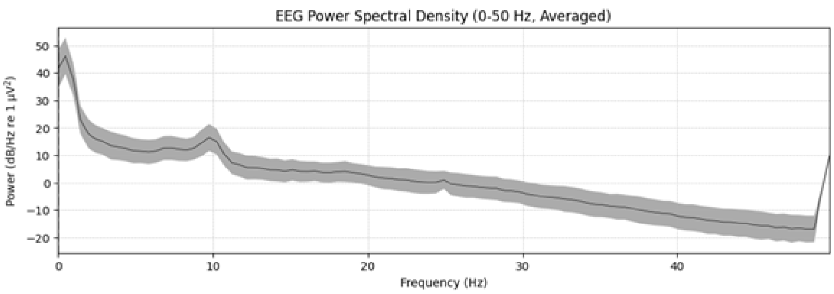


Figure 5: Power Spectral Density (PSD) Before Filtering

In the pre-filtered PSD (see Figure 5) a high amplitude is often observed at lower frequencies (delta and theta), representing slow cortical activity and possible drift. Power typically decreases as frequency increases, but sharp peaks at higher frequencies may indicate muscle artifacts or line noise.

A single plot stimulus locked epoch (trail) from the Cz electrode, corresponding to a "target" event expected to elicit a P300 response. P300 is a well-known positive deflection that occurs approximately 300 ms after a rare, task stimulus.

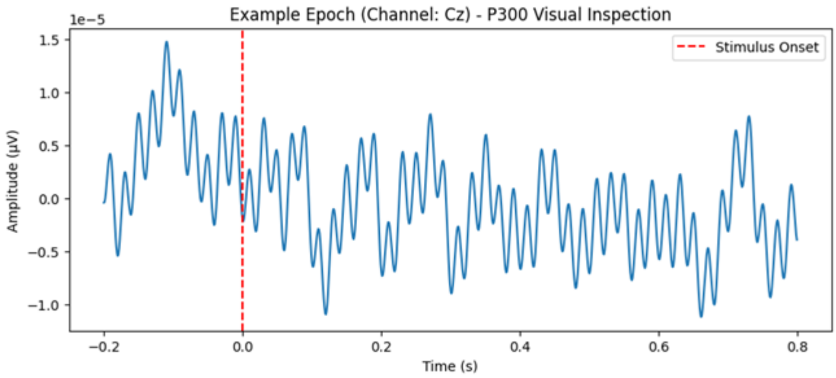
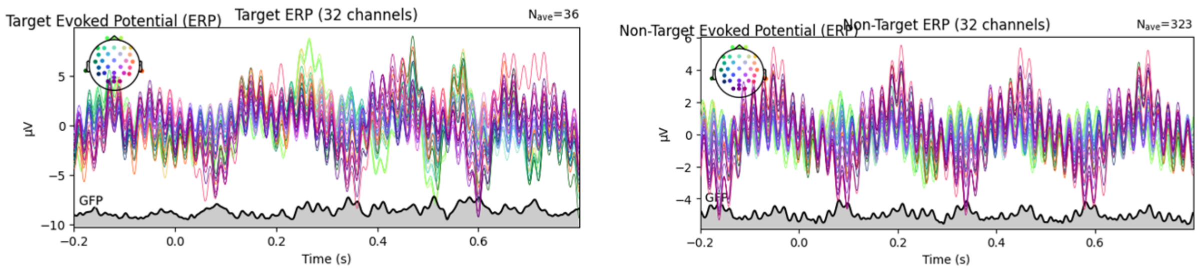


Figure 6: P300 Epoch (Cz channel, Target Trial)

The plot (see Figure 6) shows a clear waveform beginning at – 200 ms (pre-stimulus baseline) with time zero marking stimulus onset (red dashed line). A positive deflection is expected between 250-400 ms which confirms that the epoch extraction process is accurate and that the characteristic of neural response is present in the data.

Grand Average ERPs shows the mean response across all target and non-target trails plotted for all EEG channels. Grand averages increase the signal-to-noise ratio by reducing the impact of random noise revealing consistent event-related brain activity.



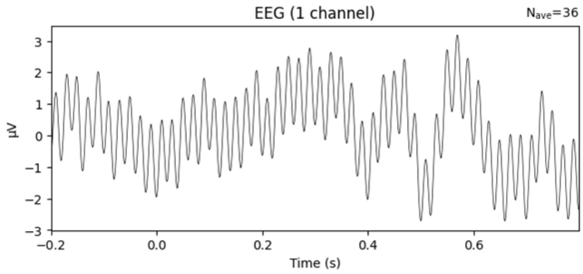
(a) Grand average ERP (target).

(b) Grand average ERP (non-target).

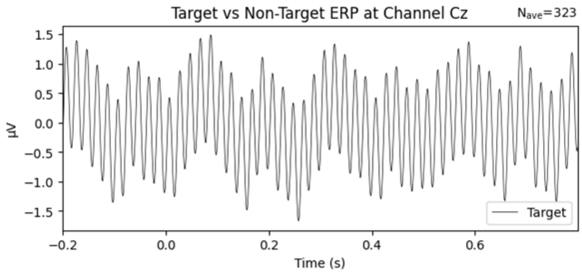
Figure 7: Grand Average ERP (Target and Non-Target, All Channels).

As seen in Figure 7 the target typically displays a pronounced positive peak around 300ms, visible on multiple central and parietal channels. The non-target ERP lacks this strong positive deflection as expected.

The plot overlays the average ERP at the Cz channel for both target and non-target trails. Cz is a canonical site for observing the p300 being centrally located and typically show the strongest amplitude. It does direct comparison of the two conditions and highlights the difference in cognitive processing between rare and frequent stimuli.



(a) Channel-Specific ERP (target).

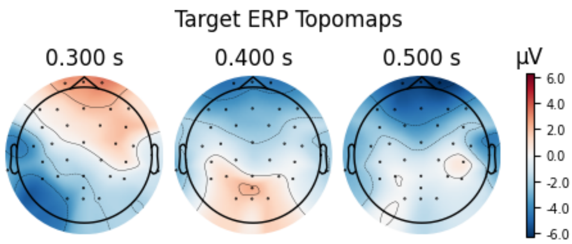


(b) Channel-Specific ERP (non-target).

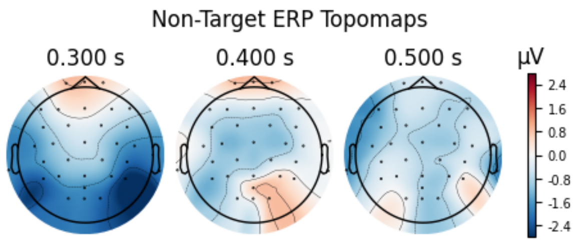
Figure 8: Channel-Specific ERP (e.g., Cz: Target vs non-Target)

Figure 8 illustrates the classic P300 positivity in the target condition with a peak around 300 ms, while the non-target response is comparatively flat. This is the physiological signature the classification models are trained to detect.

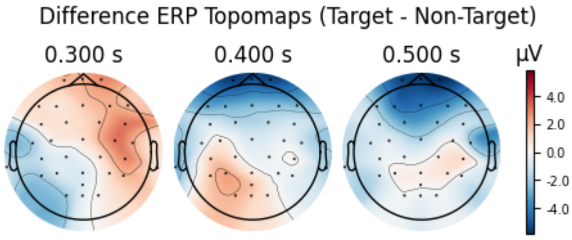
Topographic maps displayed the spatial distribution of EEG voltages across the scalp at selected time point (e.g., 300 ms, 400 ms, 500 ms post-stimulus). Topomaps provide a spatial perspective for P30 response is localized to specific brain region centro-parietal. This is also standard tool in ERP research and visually summarizes the distribution and magnitude of cognitive response.



(a) Target ERP Topomaps



(b) Non-Target ERP Topomaps



(c) Difference ERP Topomaps (Target - Non-Target)

Figure 9: Topographic Maps (Scalp Distribution of ERP at Different Times)

As shown in Figure 9 (9a), (9b), 9c at 300-500 ms the target ERP displays a maximal positive voltage over central and parietal region seen in hot colors, confirming that the response is spatially and temporally consistent with classic P300.

3.3 Data Pre-processing and Transformation

The preprocessing of EEG data is an important step to ensure the quality and interpretability of neural signals for downstream machine learning. In this study, a comprehensive preprocessing pipeline was created to leverage both established EEG analysis tools, such as the MNE library and custom procedures designed for self-supervised and supervised scenarios. Below each stage of the pipeline is described in detail.

3.3.1 Data loading and Structure

EEG datasets were acquired from the AMBER datasetAwais et al. (2023), where each subject's session contained multichannel EEG signals stored in EDF format, accompanied by a marker file (CSV) denoting the occurrence of experimental events. The study focused on two RSVP task conditions for each subject: X1 (clean RSVP) and X4 (noisy RSVP with body movement) due to computational feasibility, as training on all sessions for all subjects would require significant computational resources and extended runtimes. The subset approach ensures meaningful model evaluation under both clean and noisy real-world conditions with available resources.

3.3.2 Bandpass Filtering

EEG signals are particularly susceptible to various noise sources, and to reduce the impact of slow drifts and high-frequency noise, a zero-phase bandpass filter (0.1-30 Hz) was applied to all channels using a FIR windowed filter. This step is critical for extracting reliable event-related potentials (ERPs) such as the P300, which primarily occurs within this frequency range. The effect of filtering can be seen in the PSD plot.

3.3.3 Artifact Identification and Inspection

A visual inspection of several epochs recorded during noisy conditions (e.g., body movement) was conducted to ensure that extreme artifacts could be recognised and understood. This step aids in qualitative assessment and informs later stages if artifact rejection or cleaning is to be considered.

3.3.4 Annotation and Trigger Alignment

Each RSVP task includes triggers that mark the onset of stimulus presentations. The MNE library [1] is used to extract trigger events directly from the EDF file annotations, providing precise sample indices for each trial. These annotations are cross-referenced with corresponding CSV files to assign class labels ("target" or "non-target") to each epoch. One-to-one alignment between the number of epochs and labels is done carefully. Any discrepancies (such as missing or extra triggers) are resolved by either discarding unmatched epochs or labels as justified by the dataset documentation.

3.3.5 Epoch Extraction

For every valid trigger, a time-locked epoch is extracted from continuous EEG data:

[1] `https://mne.tools/stable/index.html`

- Epoch window: The epoch spans from -200 milliseconds (pre-stimulus baseline) to +800 milliseconds (post-stimulus period), totalling 1000 data points per epoch (at 1000 Hz).
- Channels: All 32 channels are retained to preserve the full spatial information of the signals.
- Each epoch is stored as a 2D array (channels x time points), forming a 3D dataset across all epochs.

An example of an epoch from the Cz channel is plotted, which highlights the characteristics of the positive peak of the P300 component around 300ms after stimulus in target trials.

3.3.6 Data Normalization and Augmentation

Each epoch is z-score normalized across time for each channel to ensure that all signals are on a comparable scale and reduce the influence of amplitude variability across trials and subjects.

- Augmentation (for contrastive learning): we implemented realistic augmentations for EEG data, including.
- Jitter: Addition of Gaussian noise to simulate sensor noise or spontaneous fluctuations, which helps to increase robustness to minor signal variations.
- Scaling: Channel-wise amplitude scaling to mimic variability in electrode contact or subject movement, promoting invariance to such differences.

This augmentation creates diverse "views" of the same underlying neural response, which are essential for self-supervised learning frameworks.

3.3.7 Dataset Construction for Modeling

The epoch of (X1) and (X4) RSVP task merged to yield a balanced dataset for contrastive and supervised experiments. Each sample is assigned a domain label (0 for clean, 1 for noisy), enabling domain-robust evaluation and facilitation contrastive learning across conditions. Each epoch is labeled as target (rare, P300) or non-target (common, background stimulus).

4 Implementation

This section describes the modelling framework adopted in this study to investigate the use of contrastive self-supervised learning for P300 detection in EEG under real-world noisy conditions. Multiple methods were used systematically to evaluate performance, which include SimCLR [2], TS-TCC [3], VICReg [4] (all contrastive learning paradigms), and a conventional supervised baseline CNN. The performance of each approach was evaluated

[2] `https://docs.lightly.ai/self-supervised-learning/examples/simclr.html`
[3] `https://github.com/emadeldeen24/TS-TCC`
[4] `https://docs.lightly.ai/self-supervised-learning/examples/vicreg.html`

using both feature-based logistic regression and end-to-end fine-tuning strategies. Contrastive learning leverages the creation of positive and negative pairs from unlabeled data to facilitate the learning of robust, noise-invariant features. This approach reduces reliance on extensive labelled datasets. Three contrastive methods were implemented and compared.

4.0.1 SimCLR (Simple Contrastive Learning of Representations)

SimCLR model learns representations by maximizing agreement between differently augmented views (jittering, scaling) of the same EEG epoch. For each batch, the encoder produces features for both augmented versions, and the NT-Xent loss (Normalized Temperature-Scaled Cross Entropy) encourage the model to bring these "positive pairs" closer together in feature space, while pushing apart negative pairs (augmented views of different epochs).

4.0.2 TS-TCC (Temporal and Contextual Contrasts)

TS-TCC extends SimCLR by introducing temporal cropping and channel permutation augmentations to force the encoder to learn both local (temporal) and global (contextual) invariance in EEG. The same CNN backbone was used with NT-Xent loss computed between temporally and contextually augmented versions. The Feature dimensionality (718, 128) and the result were expected to further enhance noise robustness, but in our experience, the performance was similar or slightly lower than SimCLR.

4.0.3 VICReg (Variance-Invariance-Covariance Regularization)

VICReg is a self-supervised framework where the loss function balances three objectives: invariance (bringing positive pairs close), variance (avoiding collapse), and decorrelation (diverse features). Augmented EEG pairs are encoded, and the model is regularized to avoid trivial solutions.

4.0.4 Baseline CNN

As a reference, a supervised CNN classifier was trained end-to-end using the same architecture as the encoder backbone, with a softmax classification head. The model was trained on normalized epochs with class-balanced weights to address class imbalance.

5 Evaluation

This section provides a comprehensive analysis of the results and main findings of the study as well as the implications of these finding both from academic and practitioner perspective are presented.

Table 2: Results Summary of Models for P300 Detection

Model	Accuracy	ROC-AUC	Precision (Target)	Recall (Target)	F1-score (Target)
SimCLR	0.633	0.497	0.100	0.333	0.154
TS-TCC	0.606	0.403	0.066	0.222	0.101
VICReg	0.617	0.538	0.141	0.556	0.225
Baseline CNN	0.633	0.496	0.100	0.333	0.154

Key observations are that VICReg achieved the best recall and F1-score for the target class, outperforming other contrastive models and the baseline CNN. SimCLR and the baseline performed similarly, with limited ability to detect the target due to low recal/F1 score. TS-TCC achieved the lowest recall and F1 for the target class. So, overall ROC-AUC score suggests all the models found it challenging to robustly separate the target from the non-target class under these conditions, but VICReg shows a modest improvement. The results reinforce the research question: can contrastive learning approaches improve the detection of P300 signals under noisy real-world EEG? Here, VICReg showed potential for improvement; it can be improved if the model is trained on a complete dataset and needs some hyper-tuning, but all other methods remain far from ideal performance, especially in recall and AUC.

6 Conclusion and Future Work

The study critically evaluated contrastive learning models (SimCLR, TS-TCC, VICReg) against a baseline CNN for a limited subset (X1: clean RSVP, X4: noisy RSVP), which was used in all experiments owing to computational resource constraints. VICReg demonstrates a high recall and F1-score for targets, suggesting that variance-invariance-covariance regularization may help in learning more robust representations in limited and noisy data scenarios. This modest improvement answers the research question affirmatively but highlights significant remaining challenges. These findings are consistent with prior research noting the difficulty of P300 detection under real-world noise. The results indicate that even advanced self-supervised approaches face substantial limitations with limited, imbalanced data, which is a common challenge in real EEG-based BCI applications. This is the honest reflection of experimental reality and is important for guiding future research. To achieve more robust results, future work should incorporate all available sessions and subjects from the data set to leverage more data. Utilize more computational resources for larger-scale and more complex models. The dataset should experiment with advanced augmentation and oversampling techniques to address class imbalances and evaluate generalization across multiple noisy tasks.

References

Alsufyani, A. (2020). Analyses of the p300 event-related potentials in eeg signals using shape-based kernel: Fixed and random analyses approach, *International Journal of Emerging Trends in Engineering Research* **8**(2): 476–490.

Awais, M. A., Redmond, P., Ward, T. E. and Healy, G. (2023). Amber: advancing multimodal brain-computer interfaces for enhanced robustness—a dataset for naturalistic settings, *Frontiers in Neuroergonomics* **4**: 1216440.

Bardes, A., Ponce, J. and LeCun, Y. (2021). Vicreg: Variance-invariance-covariance regularization for self-supervised learning, *arXiv preprint arXiv:2105.04906* .

Cheng, C., Liu, Y. and Liao, C. (2023). Affective brain-computer interfaces: The significance of p300 components in emotion detection and classification, *2023 2nd International Conference on Artificial Intelligence and Intelligent Information Processing (AIIIP)*, IEEE, pp. 357–361.

Das, M. and Mahadevappa, M. (2023). Enhancing asd cognitive assessment with p300 classification using emd-based qeeg wavelet features, *2023 7th International Conference on Computer Applications in Electrical Engineering-Recent Advances (CERA)*, IEEE, pp. 1–6.

Ditthapron, A., Banluesombatkul, N., Ketrat, S., Chuangsuwanich, E. and Wilaiprasitporn, T. (2019). Universal joint feature extraction for p300 eeg classification using multi-task autoencoder, *IEEE access* **7**: 68415–68428.

Du, P., Li, P., Cheng, L., Li, X. and Su, J. (2023). Single-trial p300 classification algorithm based on centralized multi-person data fusion cnn, *Frontiers in Neuroscience* **17**: 1132290.

Eldele, E., Ragab, M., Chen, Z., Wu, M., Kwoh, C. K., Li, X. and Guan, C. (2021). Time-series representation learning via temporal and contextual contrasting, *arXiv preprint arXiv:2106.14112* .

Jin, F. n. et al. (2021). P300-based brain–computer interface: Recent developments and applied engineering, *IEEE Transactions on Neural Systems and Rehabilitation Engineering* .

Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P. and Lance, B. J. (2018). Eegnet: a compact convolutional neural network for eeg-based brain–computer interfaces, *Journal of neural engineering* **15**(5): 056013.

Le-Khac, P. H., Healy, G. and Smeaton, A. F. (2020). Contrastive representation learning: A framework and review, *Ieee Access* **8**: 193907–193934.

Luck, S. J. (2014). *An introduction to the event-related potential technique*, MIT press.

Peketi, P. et al. (2023). Machine learning–enabled p300 classifier for autism, *Brain Sciences* **13**(2): 315.
URL: *https://www.mdpi.com/2076-3425/13/2/315*

Saeidi, M. et al. (2021). A review on eeg-based brain-computer interface systems for motor rehabilitation applications, *Biocybernetics and Biomedical Engineering* **41**(1): 1–25.
URL: *https://pmc.ncbi.nlm.nih.gov/articles/PMC8615531/*

Sarraf, J. et al. (2023). A study of classification techniques on p300 speller dataset, *Materials Today: Proceedings* **58**: 1668–1673.
URL: *https://www.sciencedirect.com/science/article/abs/pii/S2214785321044655*

Tajmirriahi, M., Amini, Z., Rabbani, H. and Kafieh, R. (2022). An interpretable convolutional neural network for p300 detection: Analysis of time frequency features for limited data, *IEEE Sensors Journal* **22**(9): 8685–8692.

Weng, W., Gu, Y., Guo, S., Ma, Y., Yang, Z., Liu, Y. and Chen, Y. (2025). Self-supervised learning for electroencephalogram: A systematic survey, *ACM Computing Surveys* **57**(12): 1–38.

Won, D.-W., Hwang, H.-J., Lee, S.-W. and Lee, S.-W. (2022). A large-scale electroencephalography dataset for real-time brain-computer interface speller, *Scientific Data* **9**(1): 415.
URL: *https://www.nature.com/articles/s41597-022-01509-w*

Yuan, Z., Zhou, Q., Wang, B., Zhang, Q., Yang, Y., Zhao, Y., Guo, Y., Zhou, J. and Wang, C. (2024). Psaeegnet: pyramid squeeze attention mechanism-based cnn for single-trial eeg classification in rsvp task, *Frontiers in Human Neuroscience* **18**: 1385360.

Zbontar, J., Jing, L., Misra, I., LeCun, Y. and Deny, S. (2021). Barlow twins: Self-supervised learning via redundancy reduction, *International conference on machine learning*, PMLR, pp. 12310–12320.

Zhu, Q., Zhao, X., Zhang, J., Gu, Y., Weng, C. and Hu, Y. (2023). Eeg2vec: Self-supervised electroencephalographic representation learning, *arXiv preprint arXiv:2305.13957* .