Proceedings Track

Geometric Machine Learning on EEG Signals

Anonymous Authors Email and *Address*

Abstract

Brain-computer interfaces (BCIs) offer transformative potential, but decoding neural signals presents significant challenges. The core premise of this paper is built around demonstrating methods to elucidate the underlying low-dimensional geometric structure present in high-dimensional brainwave data in order to assist in downstream BCI-related neural classification tasks. We demonstrate two pipelines related to electroencephalography (EEG) signal processing: (1) a preliminary pipeline removing noise from individual EEG channels, and (2) a downstream manifold learning pipeline uncovering geometric structure across networks of EEG channels. We conduct preliminary validation using two EEG datasets and situate our demonstration in the context of the BCI-relevant imagined digit decoding problem. Our preliminary pipeline uses an attention-based EEG filtration network to extract clean signal from individual EEG channels. Our primary pipeline uses a fast Fourier transform, a Laplacian eigenmap, a discrete analog of Ricci flow via Ollivier's notion of Ricci curvature, and a graph convolutional network to perform dimensionality reduction on high-dimensional multi-channel EEG data in order to enable regularizable downstream classification. Our system achieves competitive performance with existing signal processing and classification benchmarks; we demonstrate a mean test correlation coefficient of >0.95 at 2 dB on semi-synthetic neural denoising and a downstream EEG-based classification accuracy of 0.97 on distinguishing digit- versus non-digit thoughts. Results are preliminary and our geometric machine learning pipeline should be validated by more extensive follow-up studies; generalizing these results to larger inter-subject sample sizes, different hardware systems, and broader use cases will be crucial.

Keywords: Geometric machine learning, brain-computer interfaces, manifold learning, Ricci flow, Laplacian eigenmaps, graph convolutional networks, transformers.

1. Introduction

Developing systems capable of interpreting the underlying thoughts corresponding to raw brainwave patterns is an important research objective in the neural interfacing field (Lopez-Bernal et al., 2022), offering numerous applications from human rehabilitation (Huang et al., 2022) to augmentation (Cinel et al., 2019). The use cases resulting from advanced brain-computer interfaces (BCIs) include neuroprosthetics (Linderman et al., 2008), nonverbal communication systems (Moreira et al., 2022), thought-based interfaces (v. d. Berg et al., 2021), paralysis aides (Yoo, 2023), and beyond.

The primary obstacle to achieving advanced neural interpretation, however, remains poor signal quality. While invasive systems (i.e., systems requiring intracranial electrode implantation) can achieve higher fidelity brainwave data, the intensive requisite open brain surgery poses a major barrier to adoption (Sun et al., 2018). Non-invasive systems, such as electroencephalography (EEG), suffer from low signal-to-noise ratios and decreased resolution as neural signals diffuse across layers of bone and muscle tissue; electromyographic

© 2024 A. Authors.

(EMG) interference remains the most prominent source of EEG contamination (Chen et al., 2019; Muthukumaraswamy, 2013). Crucially, the high dimensionality associated with multi-channel EEG systems (Liang et al., 2021) at 256 Hz and above makes robust downstream classification challenging; locating low-dimensional signal within a high-dimensional space introduces a key sparsity challenge. While some of the signal-to-noise ratio (SNR) challenges inherent in non-invasive EEG data can be remediated via hardware (Narasimhan et al., 2011), software-based algorithmic (Makeig et al., 2012) solutions are crucial.

An underlying hypothesis guiding this paper is that we can leverage strategies from geometric machine learning (ML) to elucidate a lower-dimensional structure associated with underlying brainwave thought patterns located within high-dimensional EEG data streams—and that this geometric structure can be exploited to enable robust neural interpretation at scale. We also, however, first must handle the issue of high noise in individual EEG channels in order to unlock downstream structural exploitation. Inspired by recent studies demonstrating preliminary success with transformer-based neural denoising networks (Pu et al., 2022; Yin et al., 2023), we build and validate our own transformer-based network for EEG denoising. While CNN-based signal filtration algorithms take advantage of the spatial locality inherent in EEG processing (Schirrmeister et al., 2017), transformers can also harness both the spatial locality and non-locality relevant to neural signal reconstruction (Cho et al., 2023; Vaswani et al., 2017). By effectively filtering EMG noise from EEG data, we hope to show preliminary evidence that transformer-based denoising can help induce geometric differentiation of EEG classes in the denoised representational space. Transformer-based neural denoising networks, however, do struggle with issues relating to more prolonged convergence across the parameter space (resulting in long training runs and rendering live re-training in the face of new information impractical), latency issues, and preventing unnecessary signal degradation (Pu et al., 2022; Yin et al., 2023). Moreover, transformer-based signal processing methods struggle from a lack of methods for effective application site determination (Pfeffer et al., 2024). Recently, Pu et al. (2022) achieved relative success in comparison with conventional neural network processing models but did not record a PCC above 0.9 on synthetic blind source separation (BSS) and reconstruction (Pu et al., 2022). To iterate upon existing denoising architectures, we introduce an autoencoder-targeted adversarial transformer (AT-AT) architecture leveraging an autoencoder-based method for transformer application site determination and adversarial training for output regulation. In this study, we provide preliminary demonstration that this AT-AT architecture can effectively support geometric differentiation of classes in the representational space on a BCI-relevant EEG classification problem.

We pair this upstream signal denoising architecture with a downstream *manifold learning* algorithm to perform structure-preserving nonlinear embeddings of AT-AT-denoised EEG data. As previously alluded to, even after addressing the EEG-EMG denoising problem, one must address a curse-of-dimensionality challenge in terms of interpreting high-dimensional, multi-channel EEG data. The sparsity of high-dimensional EEG has posed major problems in terms of developing robust, regularizable EEG classification (Mody and Rangarajan, 2022; Hemmelmann et al., 2004; León et al., 2021); most BCI classification problems revolve around a low-dimensional, structured output—for example, predicting the number that a user is thinking of from EEG data. Hence, to extract low-dimensional information buried within a high-dimensional data source, one must create effective embeddings of high-

Proceedings Track

dimensional EEG data (Wang et al., 2019). However, conventional linear dimensionality reduction methods like principal component analysis (PCA) often fail to facilitate embedding space separation of class representations, as EEG signal includes nonlinearities that fall outside the scope of linear embedding strategies (Liang et al., 2021). In this paper, we build on recent promising results in nonlinear dimensionality reduction of neural data (Armonaite et al., 2024; Anuragi et al., 2024; Pan et al., 2022; Liang et al., 2021) by creating a geometric machine learning pipeline leveraging Laplacian eigenmaps, an Ollivier-Ricci curvature evolution algorithm, and graph convolutional networks (GCNs) to create EEG embeddings that exploit low-dimensional structure and enable regularizable downstream classification. For preliminary validation, we apply this pipeline on a real-world BCI use case selected because of its low-dimensional classification output structure: classifying imagined digits and non-digits from EEG data. The rationale behind the algorithmic components of our geometric ML pipeline are described in detail in Section 2.2.

2. Methods

2.1. Building a Denoising Model

Our study uses two datasets for validation purposes: the benchmark EEGdenoiseNet (Zhang et al., 2021) dataset for validating the initial transformer-based denoising model, and the MindBigData imagined digit decoding EEG dataset (Vivancos and Cuesta, 2022) as an initial platform through which to demonstrate our geometric machine learning pipeline. In accordance with established standardized methodologies (Zhang et al., 2021) across a benchmark SNR range of -7 dB to 2 dB, we built and trained the AT-AT system (Figure 1) end-to-end on the EEGdenoiseNet data with an EEG-EMG BSS model objective. As alluded to in the Section 1, a key engineering goal of the AT-AT system was to enable faster (re-)training capabilities—essential for real-world live BCI applications. To assess training time in a live BSS environment, it is crucial to account for both the model training duration and the total time of EEG and EMG segments used. In this study, we limited our training run to just four minutes of training data in order to simulate a real-world live retraining situation; the model was given an objective of denoising high-variance EMG artifacts from standard semi-synthetic (Zhang et al., 2021) raw EEG data.

To infer signal SNR before applying a tailored denoising model, we first employ an upstream model based on a hybrid LSTM-CNN-MLP architecture to identify the appropriate SNR target level (Choi et al., 2024) (details provided in Appendix C). The AT-AT system's initial denoising pass uses a convolutional autoencoder, which has shown architectural promise in prior EEG denoising work (Leite et al., 2018). While this autoencoder performs well in lower-noise conditions, it struggles with EEG reconstruction in higher-noise environments. To address this, we selectively invoke a time-series transformer model to handle more challenging reconstructions. However, this introduces additional complexity in the parameter space and extends training time (Pfeffer et al., 2024), so the transformer is only activated when needed. The transformer's invocation is determined by a heuristic based on the correlation between the autoencoder's output and the original signal (Choi et al., 2024). Since the autoencoder is trained to produce a noise-free EEG output, if the autoencoder output highly correlates with the raw input signal, we infer that the original signal contains less noise. Conversely, low-correlation sections (<0.8) are flagged for masking and time-

series transformer reconstruction. While this application site determination proved fruitful in terms of correlation coefficient performance (see Section 3), a notable challenge with this hybrid approach was preserving the spectral characteristics of the EEG signal when integrating transformer outputs. Inserting transformer-generated data can cause irregularities, compromising the spectral fidelity of the model's output. To address this, we incorporated adversarial training into the autoencoder-transformer hybrid architecture. The transformer-augmented autoencoder acts as the "generator," while a 1D-CNN-based discriminator model distinguishes between real and model-generated EEG data. This generator-discriminator pair is trained within a five-cycle GAN loop, ensuring the reconstructed EEG signal adheres to authentic EEG spectral properties.

2.2. Geometric ML Pipeline

Following AT-AT completion, we applied our trained denoising model to a real-world validation case: the MindBigData EEG classification problem. Previous studies have achieved promising results on inter-digit classification with this dataset (Mahapatra and Bhuyan, 2023; Mishra et al., 2021), but we attempt the ambitious task of simultaneously generalizing the imagined digit classification problem to an open-set environment. Definitions of open-set classification vary (Wu et al., 2024), so we add clarification: in open-set learning (Wu et al., 2024), we attempt to both classify digits while simultaneously rejecting data from non-digit, unknown classes based on two-second raw EEG samples. The open-set version of imagined digit classification has high representational space class ambiguity (see Figure 2, left panel), and is not covered by existing MindBigData studies (Mahapatra and Bhuyan, 2023; Mishra et al., 2021); in this paper, we attempt to leverage our AT-AT and geometric machine learning pipelines to fill a gap in current literature by demonstrating successful open-set imagined digit classification (see Section 3 and Appendix B).

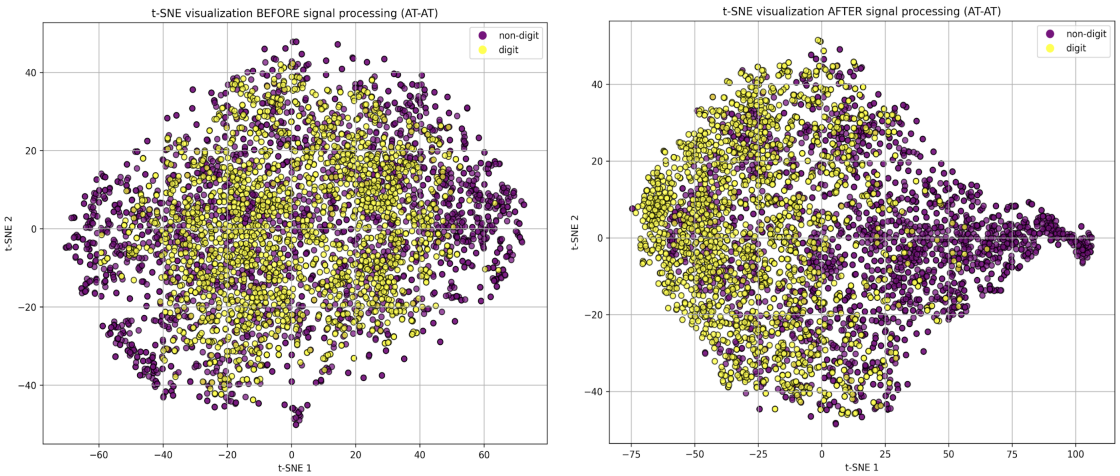


Figure 1: AT-AT processing facilitates latent space separability of neural signal classes. Pre-AT-AT t-SNE depicted left; post-AT-AT depicted right.

Distinguishing between digit- and non-digit-related thoughts from raw EEG data, as in the open-set learning imagined digit classification problem, is made difficult due to low representational space separability between these raw EEG classes (Figure 2, left panel). Crucially, however, after denoising the raw EEG data with AT-AT (see Section 3 for more detailed metrics on AT-AT performance), AT-AT's noise reduction capabilities enabled genuine geometric representational space separability to be elucidated (Figure 2, right panel). The difference between the two t-distributed stochastic neighbor embedding (t-SNE) plots in Figure 2 offer preliminary validation that beyond mere EEG-EMG source separation, our upstream AT-AT model can pull signal from noise and elucidate geometric structure in a neural interpretation context.

In order to bolster our imagined digit classification efforts while solving the aforementioned curse-of-dimensionality problem of multi-channel EEG, we implemented our primary geometric machine learning pipeline. After invoking AT-AT filtration on the individual EEG channels, we ran each channel through a one-sided FFT to create a parallel frequency vector corresponding to each time-series channel. The underlying hypothesis motivating FFT implementation was that the filtered frequency domain can elucidate key information relevant to downstream classification. After performing this FFT upscaling in conjunction with our AT-AT-denoised dataset, we next turned our focus toward tackling the second problem inhibiting successful EEG classification—that of high (multi-channel) dimensionality.

Our first stage of dimensionality reduction involved reducing the corresponding EEG time series vectors (each *512x1*) to match the dimension of the (*256x1*) one-sided FFT (Figure 2, panel 2). Given the inherent non-linearities present in EEG signal—and, crucially, the desire to preserve the local neighborhood structure (Liang et al., 2021; Wang et al., 2019) of EEG channel nodes (as more global structures are handled via the downstream GCN)—we leveraged Laplacian eigenmaps (Belkin and Niyogi, 2003; Liang et al., 2021), implementing a 256-component spectral embedding on an initial multi-channel graph. This graphical representation for the Laplacian eigenmap was performed via bidirectional linking between EEG channels FP1 and FP2 along with TP9 and TP10 (corresponding to known EEG spatial patterns (Liu et al., 2020)); the subsequent reduced output at each node was paired with the corresponding one-sided FFT for further processing. Further details on our Laplacian eigenmap process are included in Appendix A.

While the combined FFT and Laplacian eigenmap representation was folded into an initial graph, we realized that manually determining edge weights and connections might not be sufficient to capture the subtle underlying community detection necessary for optimal downstream classification. Therefore, following the work of Weber et al. (2017) and Tian et al. (2023), we applied Ricci flow via Ollivier-Ricci curvature to discover more nuanced spatial relationships between EEG channels across different brain regions in both time and frequency domains. Ricci flow, a mathematical concept originating from Riemannian geometry, iteratively deforms the metric of a manifold by adjusting its curvature over time. In our context, a discrete Ricci flow analog was applied to the graph edges by leveraging Ollivier's notion of Ricci curvature (Ollivier, 2007), which defines how two nearby distributions of mass (neighbors of EEG nodes) compare in terms of transport cost (using Wasserstein distance). This comparison allowed us to dynamically evolve the relationships between EEG channels as we refined the graph structure.

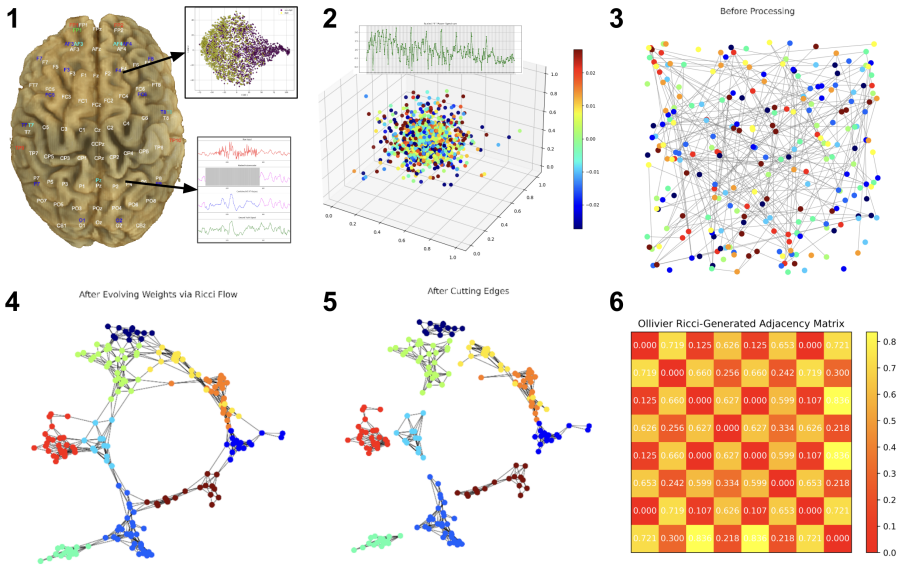


Figure 2: The six manifold learning pipeline stages. (1) AT-AT filtration; (2) FFT upscaling and Laplacian eigenmap reduction; (3) graph initialization; (4) Ricci flow evolution; (5) edge cutting; (6) final representation for GCN reduction.

To effectively compare the one-sided FFT channels and Laplacian eigenmap-reduced time-series channels, we created a custom distance metric (see Appendix A). This metric was essential for determining relationships across the graph, given that standard Euclidean distances could not sufficiently capture the complex interdependencies between both spatial and frequency representations. The metric incorporated both local neighborhood structures and broader spectral features by combining spatial proximity (from the Laplacian eigenmap) with frequency characteristics (from the FFT), enabling a more holistic comparison between EEG channels. Starting from an initial configuration of 64,000 adjacency relationships and 1,000 hypothesized subgraphs, the Ricci flow was iteratively applied over 10 unsupervised iterations per subgraph. The evolution of edge weights under the Ricci flow was driven by the computed Ollivier-Ricci curvature for each edge, which indicated whether the relationship between two nodes (channels) was converging or diverging over time. Edges with low Ricci curvature-based weighting indicated strong, cohesive communities (or clusters) between nodes, whereas highly-weighted edges pointed to more distant relationships. This distinction allowed us to dynamically adapt the graph structure to better reflect underlying community structures in the data.

During each iteration, edge weights were updated according to the formula:

$$w_{uv}(t+1) = w_{uv}(t) \cdot e^{-\alpha\kappa(u,v)}$$

where $\kappa(u,v)$ is the Ollivier-Ricci curvature of edge (u,v) and α is the learning rate controlling how aggressively the graph's structure evolves (see Appendix A). As the iterations progressed, we observed a bimodal edge weight distribution, which naturally occurred due

Proceedings Track

to the distinct nature of inter-frequency, inter-time-series, and intra-channel relationships within the EEG data. The bimodality in this distribution is reflective of the temporal and spectral communities present in the graph structure, with clusters of nodes elucidated during evolution (Weber et al., 2017). After evolving the graph's edge weights over the 10 iterations, we performed edge cutting based on the edge weight distribution. The optimal 0.6 cut ratio was selected to retain the most informative relationships, cutting past core connections while preserving the more subtle trends essential for capturing meaningful community structures. This cut probability allowed us to filter out noisy edges while retaining the strongest interdependencies, resulting in a more refined representation of the underlying EEG channel relationships (see Appendix A for further details).

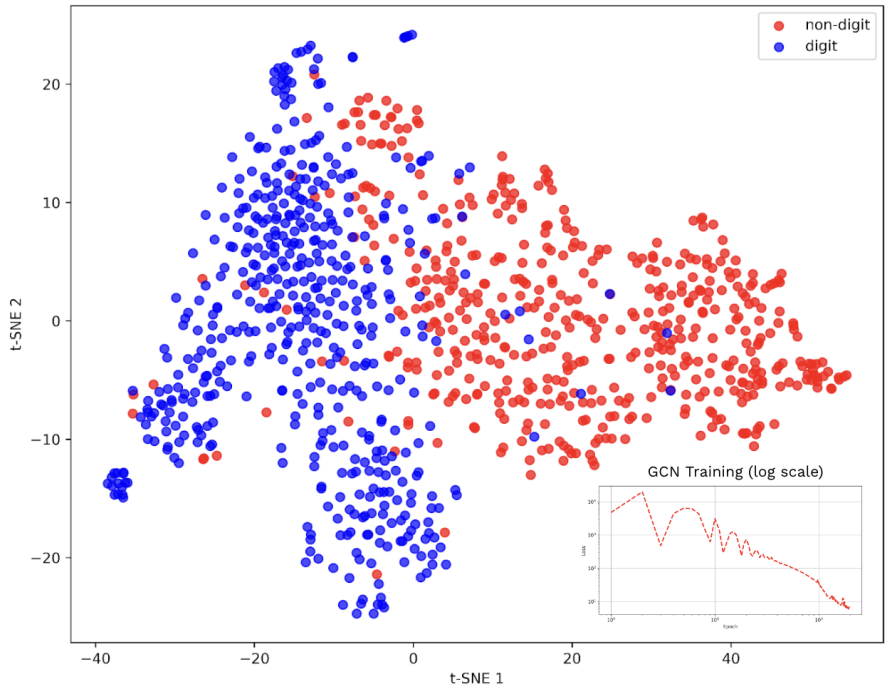


Figure 3: Class embeddings produced by the GCN after the manifold learning pipeline.

The final composite adjacency matrix, produced by aggregating the surviving edges across all subgraphs, served as the foundation for downstream training of a graph convolutional network (GCN). This matrix, derived via the Ricci flow process, provides a nuanced representation of the EEG signal's spatial and spectral structures. We then trained a graph convolutional network (GCN) to reduce the Ricci flow-generated graph into a final embedding for downstream classification. The GCN was trained over 100 epochs to preserve distances while performing eight-to-one dimensionality reduction. Skip connections were implemented to prevent over-smoothing—as shown in the GCN training sub-panel (Figure 3, lower right) and the clear digit versus non-digit separability, our GCN (i.e., the culmination of our manifold learning pipeline) was able to foster clear geometric differentiation of digit versus non-digit classes (Figure 3). Indeed, post-manifold learning, downstream classification was relatively trivial—a lightweight 1D CNN (Appendix B) proved sufficient in terms

of handling final classification (see Section 3). Further details on the adjacency matrix and GCN pipeline are provided in Appendix A.

3. Results

As mentioned, upstream EEG-EMG denoising helped foster some class separability in the representational space. This was a second-order effect; AT-AT was trained solely in the context of the aforementioned EEGdenoiseNet semi-synthetic contamination problem. On EEGdenoiseNet-based benchmarks, AT-AT posted a mean reconstructive correlation coefficient (CC) with ground truth of 0.951 at 2 dB (95% CI: 0.947, 0.954), a temporal relative root mean square error (tRRMSE) of 0.317 (95% CI: 0.305, 0.329), and a spectral relative root mean square error (sRRMSE) of 0.270 (95% CI: 0.238, 0.303). On the low end of the SNR spectrum (-7 dB), AT-AT posted a CC with ground truth of 0.703 (95% CI: 0.679-0.726), tRRMSE of 0.759 (95% CI: 0.731, 0.786), and sRRMSE of 0.800 (95% CI: 0.753, 0.848). Total mean training time on a T4 High-RAM GPU was measured at 249.1 seconds.

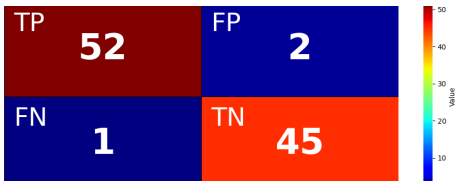
Table 1: AT-AT processing performance compared with existing models.

Model	**C-T-S (-7 dB)**	**C-T-S (2 dB)**	**Est. Size**	**TT**
Novel CNN [A]	0.69-0.72-0.65	0.92-0.33-0.30	41.9M+	unk.
EEGIFNet [B]	*did not test*	0.95-0.32-*unk.*	43.3M+	*unk.*
GCTNet [C]	*did not test*	0.94-0.28-*unk.*	10.3M+	*unk.*
RNN [D]	0.55-1.05-1.00	0.93-0.32-0.40	11.5M+	8 hrs
AT-AT (Ours)	0.70-0.76-0.80	0.95-0.32-0.27	438K+	249s

Our geometric machine learning pipeline achieved a 1D-CNN-based 97.0% test accuracy (95% CI: 93.66, 100.0) on the open-set digit- versus non-digit thought classification from two-second EEG samples. Combined with a downstream bidirectional RNN architecture (Mahapatra and Bhuyan, 2023) to determine final inter-digit granular classification, our geometric ML-driven open-set learning pipeline achieves implied >90% accuracy (see Appendix B) across all individual digits.

Table 2: Downstream classification performance overview after manifold learning. More granular details included in Appendix B.

Metric	**Score**	**Confusion Matrix**
Accuracy	0.970	
F1 Score	0.968	
AUCROC	0.971	
Sensitivity	0.978	
Specificity	0.963	



Proceedings Track

4. Discussion

4.1. Implications

We can observe evidence for the efficacy of our manifold learning pipeline in the induced latent space class separability in Figure 3, especially when contrasted with the initial lack of embedding class differentiation in Figure 1 (left panel). We can also see preliminary evidence for the Ollivier-Ricci curvature-based algorithm in the GCN embedding performance and rapid observed GCN convergence (Figure 3, lower right panel), potentially serving as an indication that evolving edge weights using Ricci flow transforms the EEG graph to elucidate intrinsic geometric relationships between node collections. By uncovering subtle underlying graph structure, this revelation of nuanced communities may facilitate the downstream GCN's ability to learn and preserve distances in the embedding space. Our preliminary validation of the upstream AT-AT denoising model also proved promising; by abstracting out high-noise segment handling and implementing selective transformer invocation, we were able to train a reasonably competitive (Table 1) model performing filtration of high-variance EMG artifacts with a relatively low parameter count.

Beyond direct neural interfacing applications in the realm of imagined digit classification (see Appendix B for details), the manifold learning pipeline presented in this study could open doors to further research on the underlying semantic geometric structure present in high-dimensional brainwave data. The t-SNE visualizations in Figure 1 and Figure 3 offer a preliminary indication that raw high-dimensional EEG signals could contain an underlying geometric structure corresponding to semantic thoughts. More extensive validation of this early idea could open the door to future studies on EEG-based thought modeling with implications for uncovering fundamental properties of intelligence.

4.2. Limitations

It is crucial to note that more extensive studies will be needed to generalize these results to larger sample sizes, broader user studies, different hardware systems, and wider-ranging use cases. In particular, the geometric manifold learning pipeline will require validation on a wider slate of EEG-based classification tasks; the limited deployment of our methods to imagined digit classification (see Appendix B) curtails the pipeline's broader applicability without further adaptation. The stability and robustness of the Ricci flow-based graph optimization also require further validation, particularly across different signal modalities or downstream use cases. Our AT-AT denoising system must also be validated across a wider slate of both semi-synthetic and real-world environments beyond the high-variance EEGdenoiseNet EMG filtration context and subsequent imagined digit classification deployment. Importantly, while intensive and beyond the scope of this initial study (which relied on available open-source EEG data), validation of these methods will also require larger-scale recruitment of human participants for exhaustive, use-case-specific EEG data collection.

It is also important to note that the capabilities unveiled in this study—particularly the ability to potentially discern internal thoughts from non-invasive external brainwave activity—could pose a privacy risk to individuals if misused. Careful consideration, alongside the development of robust ethical and regulatory frameworks, will be vital to ensure that BCI advancements are managed responsibly and in alignment with human values.

References

A. Anuragi, D. S. Sisodia, and R. B. Pachori. Mitigating the curse of dimensionality using feature projection techniques on electroencephalography datasets: An empirical review. *Artificial Intelligence Review*, 57(3):75, February 2024.

K. Armonaite, L. Conti, L. Laura, M. Primavera, and F. Tecchio. Neural signal analysis with non-linear methods. *Neural Signal Analysis*, 2024.

M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15(6):1373–1396, June 2003.

X. Chen, X. Xu, A. Liu, S. Lee, X. Chen, X. Zhang, M. J. McKeown, and Z. J. Wang. Removal of muscle artifacts from the EEG: A review and recommendations. *IEEE Sensors Journal*, 19(14):5353–5368, March 2019.

S. Cho, S. Cho, S. Park, H. Lee, H. Lee, and M. Lee. Curve your attention: Mixed-curvature transformers for graph representation learning. Technical report, arXiv preprint arXiv:2309.04082, 2023.

B. J. Choi, G. Milsap, C. A. Scholl, and F. Tenore. Targeted adversarial denoising autoencoders (TADA) for streamlined neural signal filtration. Technical report, Pending U.S. DoD Clearance to the Journal of Neural Engineering, 2024.

C. Cinel, D. Valeriani, and R. Poli. Neurotechnologies for human cognitive augmentation: Current state of the art and future prospects. *Frontiers in Human Neuroscience*, 13, January 2019. doi: 10.3389/fnhum.2019.00013.

C. Hemmelmann, M. Horn, S. Reiterer, B. Schack, T. Süsse, and S. Weiss. Multivariate tests for the evaluation of high-dimensional EEG data. *Journal of Neuroscience Methods*, 139(1):111–120, October 2004.

J. Huang, G. Li, H. Su, and Z. Li. Development and continuous control of an intelligent upper-limb neuroprosthesis for reach and grasp motions using biological signals. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 52(6):3431–3441, June 2022. doi: 10.1109/TSMC.2021.3069084.

W. Huang, Y. Rong, T. Xu, F. Sun, and J. Huang. Tackling over-smoothing for general graph convolutional networks. Technical report, arXiv preprint arXiv:2008.09864, August 2020.

N. M. Leite, E. T. Pereira, E. C. Gurjao, and L. R. Veloso. Deep convolutional autoencoder for EEG noise filtering. In *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 2605–2612. IEEE, December 2018.

J. León, J. J. Escobar, J. González, J. Ortega, F. M. Arrabal-Campos, A. Ortiz, and M. Damas. Recurrent neural networks and efficiency in high-dimensional EEG classification. In *International Conference on Bioengineering and Biomedical Signal and Image Processing*, pages 297–310, 2021.

Proceedings Track

Z. Liang, R. Zhou, L. Zhang, L. Li, G. Huang, Z. Zhang, and S. Ishii. EEGFuseNet: Hybrid unsupervised deep feature characterization and fusion for high-dimensional EEG with an application to emotion recognition. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 29:1913–1925, 2021.

M. D. Linderman, G. Santhanam, C. T. Kemere, V. Gilja, S. O'Driscoll, M. Y. Byron, A. Afshar, S. I. Ryu, K. V. Shenoy, and T. H. Meng. Signal processing challenges for neural prostheses. *IEEE Signal Process. Mag.*, 25:18–28, 2008. doi: 10.1109/MSP.2008.4408439.

R. Liu, M. Xu, Y. Zhang, E. Peli, and A. D. Hwang. A pilot study on EEG-based evaluation of visually induced motion sickness. *The Journal of Imaging Science and Technology*, 64 (2):20501–1, 2020.

D. Lopez-Bernal, D. Balderas, P. Ponce, and A. Molina. A state-of-the-art review of EEG-based imagined speech decoding. *Front. Hum. Neurosci.*, 16, April 2022. doi: 10.3389/fnhum.2022.867281.

N. Mahapatra and P. Bhuyan. EEG-based classification of imagined digits using a recurrent neural network. *Journal of Neural Engineering*, 20(2), 2023.

S. Makeig, C. Kothe, T. Mullen, N. Bigdely-Shamlo, Z. Zhang, and K. Kreutz-Delgado. Evolving signal processing for brain–computer interfaces. *Proc. IEEE*, 100:1567–1584, May 2012. doi: 10.1109/JPROC.2012.2185009.

R. Mishra, K. Sharma, and A. Bhavsar. Visual brain decoding for short duration EEG signals. In *2021 29th European Signal Processing Conference (EUSIPCO)*, pages 1226–1230. IEEE, August 2021.

S. K. Mody and G. Rangarajan. Sparse representations of high dimensional neural data. *Scientific Reports*, 12(1):7295, 2022.

J. P. C. Moreira, V. R. Carvalho, E. M. A. M. Mendes, A. Fallah, T. J. Sejnowski, C. Lainscsek, and L. Comstock. An open-access EEG dataset for speech decoding: Exploring the role of articulation and coarticulation. Technical report, Postgraduate Program in Electrical Engineering, Federal University of Minas Gerais, Belo Horizonte, MG, Brazil, and Departments of Neurosurgery, Computational Neurobiology Laboratory, Institute for Neural Computation, Division of Biological Sciences, University of California, and Department of Linguistics, HSE University, Moscow, 2022.

S. D. Muthukumaraswamy. High-frequency brain activity and muscle artifacts in MEG/EEG: A review and recommendations. *Frontiers in Human Neuroscience*, 7, 2013. doi: 10.3389/fnhum.2013.00138.

S. Narasimhan, H. J. Chiel, and S. Bhunia. Ultra-low-power and robust digital-signal-processing hardware for implantable neural interface microsystems. *IEEE Trans. Biomed. Circuits Syst.*, 5:169–178, April 2011. doi: 10.1109/TBCAS.2010.2076281.

Y. Ollivier. Ricci curvature of metric spaces. *Comptes Rendus Mathematique*, 345(11): 643–646, 2007.

Y. T. Pan, J. L. Chou, and C. S. Wei. Matt: A manifold attention network for EEG decoding. In *Advances in Neural Information Processing Systems*, volume 35, pages 31116–31129, December 2022.

J. Peng, R. Lei, and Z. Wei. Beyond over-smoothing: Uncovering the trainability challenges in deep graph neural networks. Technical report, arXiv preprint arXiv:2408.03669, August 2024.

M. A. Pfeffer, S. S. Ling, and J. K. Wong. Exploring the frontier: Transformer-based models in EEG signal analysis for brain-computer interfaces. *Computers in Biology and Medicine*, page 108705, June 2024.

X. Pu, P. Yi, K. Chen, Z. Ma, D. Zhao, and Y. Ren. EEGDnet: Fusing non-local and local self-similarity for EEG signal denoising with transformer. *Computers in Biology and Medicine*, 151:106248, 2022.

R. T. Schirrmeister, J. T. Springenberg, L. D. Fiederer, M. Glasstetter, K. Eggensperger, M. Tangermann, F. Hutter, W. Burgard, and T. Ball. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.*, 38(11):5391–5420, November 2017. doi: 10.1002/hbm.23730.

M. Z. Soroush, P. Tahvilian, M. H. Nasirpour, K. Maghooli, K. Sadeghniiat-Haghighi, S. V. Harandi, Z. Abdollahi, A. Ghazizadeh, and N. Jafarnia Dabanloo. EEG artifact removal using sub-space decomposition, nonlinear dynamics, stationary wavelet transform, and machine learning algorithms. *Frontiers in Physiology*, 13, August 2022.

F. T. Sun, S. A. Desai, T. K. Tcheng, and M. J. Morrell. Changes in the electrocorticogram after implantation of intracranial electrodes in humans: The implant effect. *Clinical Neurophysiology*, 129(3):676–686, March 2018.

Y. Tian, Z. Lubberts, and M. Weber. Curvature-based clustering on graphs. Technical report, arXiv preprint arXiv:2307.10155, 2023.

B. v. d. Berg, S. v. Donkelaar, and M. Alimardani. Inner speech classification using EEG signals: A deep learning approach. In *2021 IEEE 2nd International Conference on Human-Machine Systems (ICHMS)*, pages 1–4, 2021. doi: 10.1109/ICHMS53169.2021.9582457.

A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, volume 30, 2017.

D. Vivancos and F. Cuesta. Mindbigdata 2022: A large dataset of brain signals. Technical report, arXiv preprint arXiv:2212.14746, 2022.

Y. Wang, C. M. Ting, X. Gao, and H. Ombao. Exploratory analysis of brain signals through low dimensional embedding. In *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 997–1002, March 2019.

Proceedings Track

M. Weber, E. Saucan, and J. Jost. Characterizing complex networks with forman-ricci curvature and associated geometric flows. *Journal of Complex Networks*, 5(4):527–550, 2017.

J. Wu, X. Li, S. Xu, H. Yuan, H. Ding, Y. Yang, X. Li, J. Zhang, Y. Tong, X. Jiang, and B. Ghanem. Towards open vocabulary learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, February 2024.

J. Yin, A. Liu, C. Li, R. Qian, and X. Chen. A GAN guided parallel CNN and transformer network for EEG denoising. *IEEE Journal of Biomedical and Health Informatics*, May 2023.

K.-S. Yoo. Motion prediction using brain waves based on artificial intelligence deep learning recurrent neural network. *J. Exerc. Rehabil.*, 19(4):219–227, August 2023. doi: 10.12965/jer.2346242.121.

H. Zhang, M. Zhao, C. Wei, D. Mantini, Z. Li, and Q. Liu. EEGdenoiseNet: A benchmark dataset for deep learning solutions of EEG denoising. *J. Neural Eng.*, 18:056057, 2021.

Appendix A. Manifold Learning Details

A.1. Laplacian Eigenmap

For our initial dimensionality reduction, we apply a Laplacian eigenmap on networks of FFT-augmented, AT-AT-denoised channels. The adjacency matrix $\mathbf{A}$ captures pairwise relationships between rows based on specific regions of interest. Links are created between temporal-parietal (TP) cortical regions and frontal-parietal (FP) cortical regions, resulting in a symmetric matrix.

The graph Laplacian $\mathbf{L}$ is computed as:

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

where $\mathbf{D}$ is the degree matrix, defined as:

$$D_{ii} = \sum_{j=1}^{n} A_{ij}$$

The spectral embedding is performed by solving the eigenvalue problem:

$$\mathbf{L}\mathbf{v} = \lambda \mathbf{v}$$

where λ are the eigenvalues and $\mathbf{v}$ are the corresponding eigenvectors. The smallest non-zero eigenvalues correspond to the most informative dimensions of the data. We retain the first 256 eigenvectors to form a 256-column matrix $\mathbf{Y}$, which provides the lower-dimensional embedding:

$$\mathbf{Y} = f(\mathbf{L})$$

where $f(\mathbf{L})$ is the spectral embedding function.

Proceedings Track

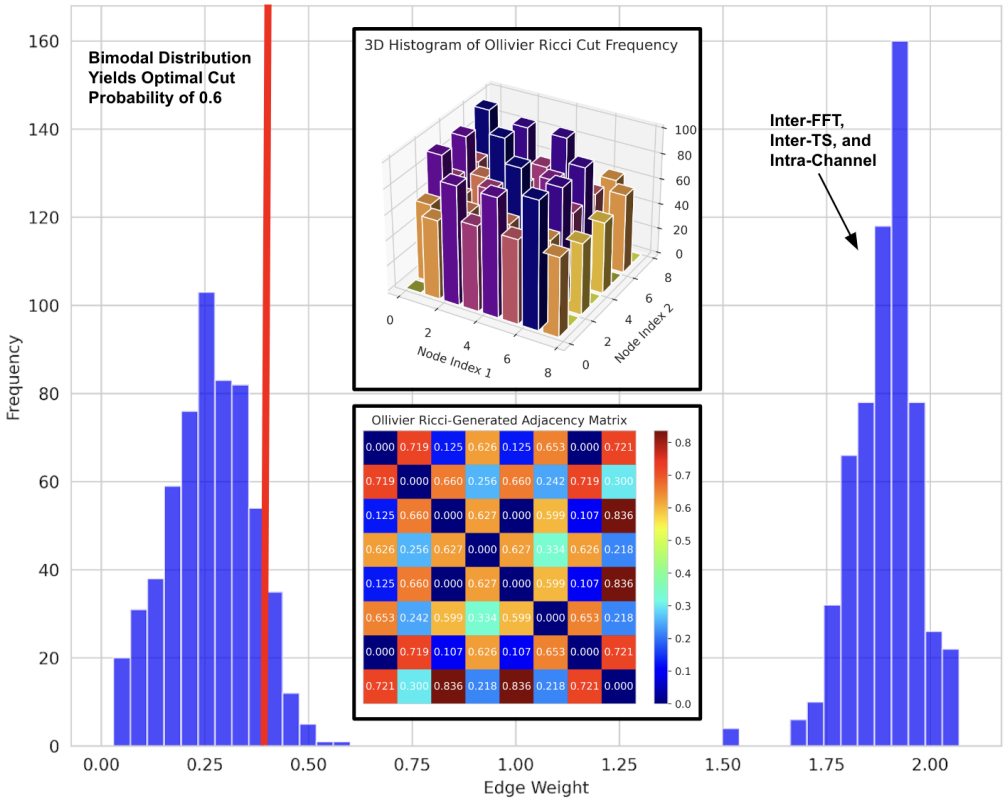


Figure 4: Summary of Ollivier-Ricci curvature pipeline metrics.

As mentioned, before performing the Laplacian eigenmap-driven dimensionality reduction, we also integrate the Fourier transform of the spectral embedding. Given a signal $\mathbf{x}(t)$, its Fourier transform is:

$$\mathcal{F}\{\mathbf{x}(t)\} = \mathbf{X}(f) = \int_{-\infty}^{\infty} \mathbf{x}(t) e^{-i2\pi f t} dt$$

The rows of $\mathbf{Y}$ and $\mathbf{Z}$ are then interleaved to form an augmented representation $\mathbf{W}$ upon which we perform the spectral embedding:

$$\mathbf{W} = [\mathbf{Y}_1; \mathbf{Z}_1; \mathbf{Y}_2; \mathbf{Z}_2; \ldots; \mathbf{Y}_n; \mathbf{Z}_n]$$

A.2. Ricci Flow

Algorithm 1 describes the mechanics of our Ricci flow algorithm for optimal graph derivation. Our custom metric capturing spatial and frequency relationships is described in Algorithm 1, **1(b)**. Figure 4 depicts summary distributions and statistics for our Ricci flow-driven algorithm on the imagined digit classification deployment case, with the bimodal nature of edge weight distribution (as mentioned in Section 2) evident. Edge cut frequencies and weight aggregations are displayed in the central panels of Figure 4.

Algorithm 1: Discrete Ricci Flow Algorithm for Neural Graph Evolution

1. **Graph Initialization**
   (a) Define the graph $G = (V, E, w)$, where V is the set of vertices, E is the set of edges, and w is the weight function $w : E \to \mathbb{R}^+$.
   (b) Compute the initial distances between node pairs using the distance function:

$$d(u,v) = \begin{cases} \frac{1}{2}(\|f(u) - f(v)\| + \mathrm{FFT}\Delta(u,v)) \\ \|f(u) - f(v)\| \\ \frac{1-\rho(\mathrm{FFT}(f(u)),\mathrm{rs}(f(v),256))}{2} \end{cases}$$

   (c) Construct the graph by assigning weights to edges:

$$w(u,v) = \frac{1}{d(u,v)}, \quad \forall (u,v) \in E \text{ with } d(u,v) \neq 0.$$

2. **Ricci Flow Evolution**
   (a) Compute the Ollivier-Ricci curvature for each edge:

$$\kappa(u,v) = 1 - \frac{W_1(\mu_u, \mu_v)}{d(u,v)},$$

   where W_1 is the Wasserstein distance and μ_u, μ_v are distributions over neighbors of u and v.
   (b) Evolve edge weights over time based on Ricci curvature until convergence:

$$w_{uv}(t+1) = w_{uv}(t) \cdot e^{-\alpha\kappa(u,v)}, \quad \alpha > 0.$$

3. **Post-Processing**
   (a) Identify and remove the top $\rho\%$ of edges by weight:

$$E_{\text{cut}} = \{(u,v) \in E : w_{uv}(t) \text{ in top } \rho\% \text{ of } E\}.$$

   (b) Compute the average adjacency matrix over all iterations:

$$\bar{A} = \frac{1}{N}\sum_{i=1}^{N} A_i, \quad A_i \text{ is the adjacency matrix of } G_i.$$

A.3. Graph Convolutional Network

Architecturally, the GCN consists of two graph convolutional layers; the relatively shallow architecture helps mitigate *oversmoothing*, a phenomenon where the node representations can become indistinguishable from each other after extended iterative computation (Peng et al., 2024). The optimized adjacency matrix $\mathbf{A}$ is derived via the Ricci flow-based algorithm, which preserves the geometric structure of the graph while enhancing the meaningful relationships between the nodes. The model is optimized with Adam with a learning rate of 1×10^{-3}. To preserve geometric relationships, the GCN is trained to maintain original distances via the objective function in the embedding space. We also implement skip connections at each layer to prevent oversmoothing, in accordance with Huang et al. (2020).

The first graph convolution layer transforms the input features $\mathbf{X}$ as:

$$\mathbf{H}^{(1)} = \sigma\left(\mathbf{A}\mathbf{X}\mathbf{W}^{(1)}\right) + \mathbf{X}\mathbf{W}^{(1)}_{\text{skip}}$$

The second graph convolution layer reduces the dimensionality of each node network to a vector embedding in $\mathbb{R}^{256}$ for final 1D-CNN downstream classification:

$$\mathbf{H}^{(2)} = \sigma\left(\mathbf{A}\mathbf{H}^{(1)}\mathbf{W}^{(2)}\right) + \mathbf{H}^{(1)}\mathbf{W}^{(2)}_{\text{skip}}$$

Appendix B. Further Results

The final 1D CNN classification model was evaluated on a lock box 100-sample test set comprising 20% of the initial classification data. The lightweight, 5.2K-parameter classifier accepts the GCN-reduced $\mathbb{R}^{256}$ output as input and applies two convolutional layers with max pooling and ReLU activation. The 1D CNN was optimized via Adam with a learning rate of 0.01 and trained over 70 epochs with a batch size of 80. The lightweight structure is intentional, as the model is meant to merely output a final classification with a robust, regularizable fit—the "hard work" of classification is all handled upstream by the geometric machine learning pipeline. Indeed, the 1D CNN was assigned a fairly trivial classification task, as evident in Figure 3, which depicts a t-SNE projection of the manifold learning-generated GCN embeddings. In an attempt to ensure the integrity and reproducibility of our results, we did not alter any of the default scikit-learn settings for any of the t-SNE plots used in this study, thereby maintaining the standard configuration in order to provide unbiased representations of the data.

In addition to the statistics reported in Section 3, more granular details regarding the final performance of the 1D CNN classifier—as a reflection of the upstream geometric ML processing—are described in Figure 5. By combining the final 1D-CNN with the bidirectional RNN architecture from Mahapatra and Bhuyan (2023) to determine final inter-digit granular classification, we can achieve an estimated 90% accuracy across all individual digits in an open-set learning environment. Note that this estimate, however, involves a balanced digit- versus non-digit dataset distribution, which may not generalize to real-world settings.

Proceedings Track

Figure 5: Estimated true performance on the imagined digit recognition problem with our geometric machine learning pipeline and a downstream bidirectional RNN from Mahapatra and Bhuyan (2023).

Appendix C. AT-AT Details

We provide further detail on the upstream autoencoder-targeted adversarial transformer (AT-AT) denoising model below. As previously mentioned, the autoencoder architecture (Choi et al., 2024; Leite et al., 2018) follows a convolutional structure with batch normalization and dropout layers to stabilize training and prevent overfitting. Training is conducted via a correlation coefficient-driven objective function. The encoder starts with an input layer that accepts a 1D signal consisting of 512 frames, representing the EEG signal from a single channel. The encoder includes two convolutional layers: the first Conv1D layer contains 32 filters with a kernel size of 3, followed by batch normalization and ReLU activation. The second Conv1D layer has 64 filters with similar activation and normalization steps. Max pooling with a factor of 2 is applied after each convolution to reduce the temporal dimension. The bottleneck layer of the encoder has 128 filters, maintaining the same structure of batch normalization and ReLU activation. On the decoder side, the architecture mirrors the encoder, with upsampling layers replacing the pooling layers. Two convolutional layers reconstruct the signal, and a final Conv1D layer with a single filter and sigmoid activation

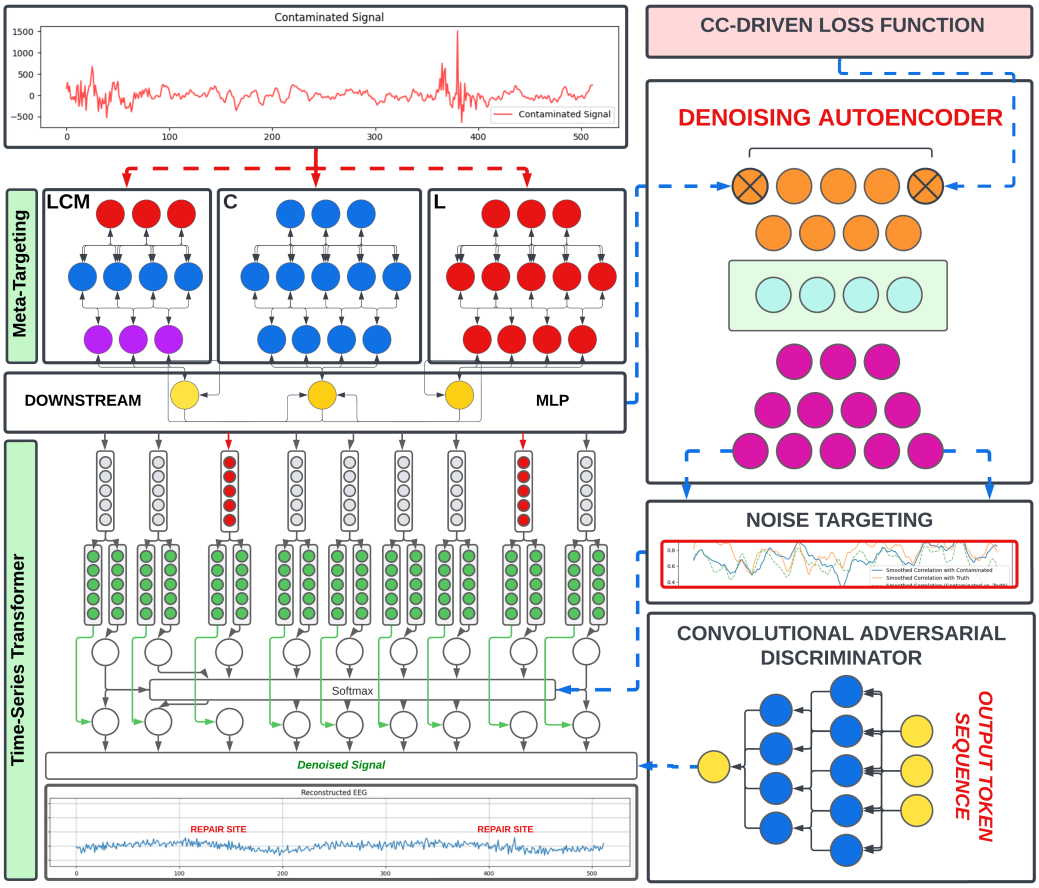


Figure 6: Architectural diagram of the autoencoder-targeted adversarial transformer.

outputs the final reconstructed signal. The autoencoder is trained via Adam with a learning rate of 1×10^{-4} and a batch size of 20 across ten epochs.

As mentioned in Section 2, we trained an upstream pre-processing model to infer the appropriate SNR target level before deploying AT-AT, following previously demonstrated abstraction methods (Choi et al., 2024). This model selects the suitable iteration of AT-AT based on the detected SNR level, enabling more tailored processing. Pre-processing accounted for 32.9% of the 249-second training time. The model integrates a hybrid architecture combining CNN, LSTM, and a meta-classifier to classify synthetic EEG and EMG data labeled with various SNR levels. The CNN pathway reshapes the EEG input into 2D blocks and processes them through Conv2D layers with ReLU, batch normalization, max-pooling, and dropout. The LSTM pathway captures temporal dependencies from the raw EEG data through two LSTM layers, which are flattened. The LSTM-CNN-MLP pathway involves two LSTM layers reshaped for convolutional processing via Conv2D, followed by a dense MLP layer. These outputs are concatenated into a meta-classifier with two fully connected layers and a softmax output, which predicts the appropriate SNR level across the -7 to 2 dB range. The training process spans 100 epochs with a batch size of 100. Inferring

Proceedings Track

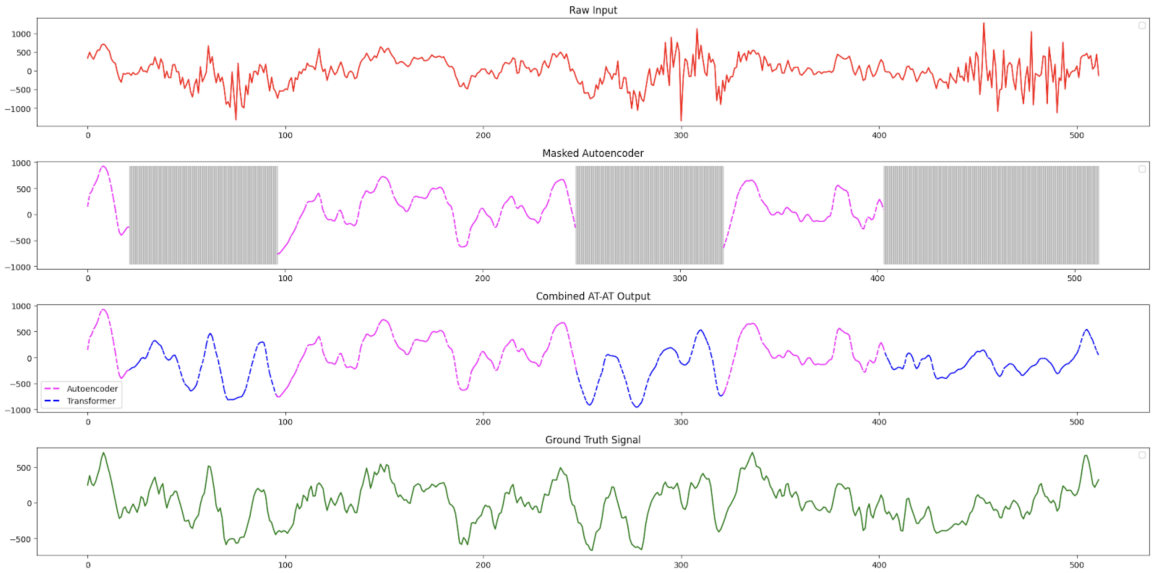


Figure 7: AT-AT model workflow. From top to bottom: (1) raw input signal, (2) initial autoencoder filtration pass with high-noise target site masking, (3) time-series transformer-based reconstruction of target sites, (4) ground truth signal.

relative SNR interference has been shown to be a relatively trivial task, with past accuracies demonstrated at 98% (Soroush et al., 2022); our upstream LCL model was able to correctly infer SNR across all 100 test cases. (This pre-processing layer is used to toggle between downstream models and does not directly interact with the reconstruction process.)

The adversarial model uses the previously described generative adversarial network (GAN) architecture, combining a transformer-based generator and a convolutional discriminator to perform EEG denoising. The generator utilizes a transformer encoder with two encoder layers, multi-head attention (with four heads), and a feedforward network of size 128. Input features, which consist of two channels, are embedded into a 16-dimensional space before entering the transformer. The output from the transformer passes through a convolutional smoothing layer (Conv1D with a kernel size of 3) to ensure smoothness in the generated signal. Finally, a fully connected layer transforms the output into a 1D signal. The discriminator is a Conv1D-based model designed to distinguish between real and generated EEG signals. It contains two convolutional layers: the first Conv1D layer has 64 filters and LeakyReLU activation, followed by dropout, while the second layer has 128 filters and the same activation and dropout sequence. A fully connected layer reduces the representation to a scalar, which is processed by a sigmoid activation function to classify the signal as authentic or generated. The GAN training process alternates between generator and discriminator updates over five cycles in each iteration, with both models using binary cross-entropy loss for optimization. The Adam optimizer is used for both the generator and discriminator, with a learning rate of 1×10^{-4}. The adversarial model is trained over ten epochs with a batch size of 20.