Proceedings of the
10th Graz Brain-Computer Interface Conference 2026

10.3217/978-3-99161-093-9-014

Riemannian Geometry-Preserving Variational Autoencoder for MI-BCI Data Augmentation

Viktorija Poļaka[1], Ivo Pascal De Jong[1], Andreea Ioana Sburlea[1]

[1]Faculty of Science and Engineering, University of Groningen, Groningen, The Netherlands

E-mail: victoria_polaka@proton.me

ABSTRACT: This paper addresses the challenge of generating synthetic electroencephalogram (EEG) covariance matrices for motor imagery brain-computer interface (MI-BCI) applications. *Objective*. We aim to develop a generative model capable of producing high-fidelity synthetic covariance matrices while preserving their symmetric positive-definite nature. *Approach*. We propose a Riemannian geometry-preserving variational autoencoder (RGP-VAE) integrating geometric mappings with a composite loss function combining Riemannian distance, tangent space reconstruction accuracy and generative diversity. *Results*. The model generates valid, representative EEG covariance matrices, while learning a subject-invariant latent space. Synthetic data proves practically useful for MI-BCI, with its impact depending on the paired classifier. *Contribution*. This work introduces and validates the RGP-VAE as a geometry-preserving generative model for EEG covariance matrices, highlighting its potential for signal privacy, scalability and data augmentation.

INTRODUCTION

While Riemannian geometry-based classifiers currently dominate MI-BCI competitions, their advancement towards mainstream applications is hindered by data scarcity and inter-subject variability, which necessitates lengthy calibration sessions [1–4]. Deep learning alternatives have yet to surpass these geometric pipelines, possibly explained by the limited availability of subject-level data [3]. To overcome these limitations, we propose a novel data augmentation framework tailored to the specific geometric properties of EEG covariance matrices, which are symmetric positive-definite (SPD), i.e., symmetric with strictly positive eigenvalues.

Previous work exploring data augmentation directly on the SPD manifold by geometrically interpolating between existing covariance matrices of the same class has successfully boosted BCI classification accuracy in data-scarce scenarios for SSVEP and ERP tasks [5]. However, this approach is fundamentally limited to the convex hull of the original data and thus cannot generate plausible variations that exist in unexplored regions of the manifold. A variational autoencoder (VAE) [6], which can learn a latent representation of a manifold [7], may offer an alternative to overcome this convex-hull limitation and generate diverse yet plausible synthetic samples that extend beyond the convex hull.

However, standard VAEs assume Euclidean geometry, creating a conflict when working with the Riemannian SPD manifold structure of EEG covariance matrices; applying standard Euclidean operations on this curved manifold causes geometric distortions (e.g., the "swelling effect") [5]. We address this by proposing the Riemannian geometry-preserving VAE (RGP-VAE) designed to preserve geometric integrity, utilizing parallel transport [8] to align data and thus enable the model to learn subject-invariant features. The focus is specifically on the challenging and practically relevant problem of cross-subject generalization, with the aim to reduce the need for extensive calibration [1]. Accordingly, this paper aims to: (1) establish if a Riemannian geometry-preserving VAE can generate valid synthetic EEG covariance matrices, and (2) evaluate whether this synthetic data improves cross-subject MI-BCI performance.

METHODS

Data and Preprocessing: We use the dataset from Faller et al. [9], containing 13-channel EEG recordings from 12 subjects performing a two-class motor imagery task (right hand versus both feet). The data loading procedure, using the "Mother of All BCI Benchmarks" framework [10], resulted in a total of 5572 trials across the 12 subjects (398 or 597 trials per individual).

The EEG trials are bandpass filtered (8–30 Hz) to capture sensorimotor rhythms. The raw voltage signals were then scaled to microvolts (10^6). To address non-stationarity, exponential moving standardization (EMS) [11] is applied (decay factor = 0.01, initial block size = 1000, epsilon = 1e-6). EMS can be seen applied to data for training deep learning models as these models tend to be sensitive to the input scale [12]. Finally, trials are converted into spatial covariance matrices in $\mathbb{R}^{13\times13}$ using the oracle approximating shrinkage estimator [13], yielding well-conditioned SPD matrices.

To address inter-subject EEG variability, which manifests as geometric differences in their location on the Riemannian manifold, parallel transport [8] was applied. This technique geometrically transports matrices from each subject reference mean to a global (class) reference mean via a congruence transformation.

10.3217/978-3-99161-093-9-014

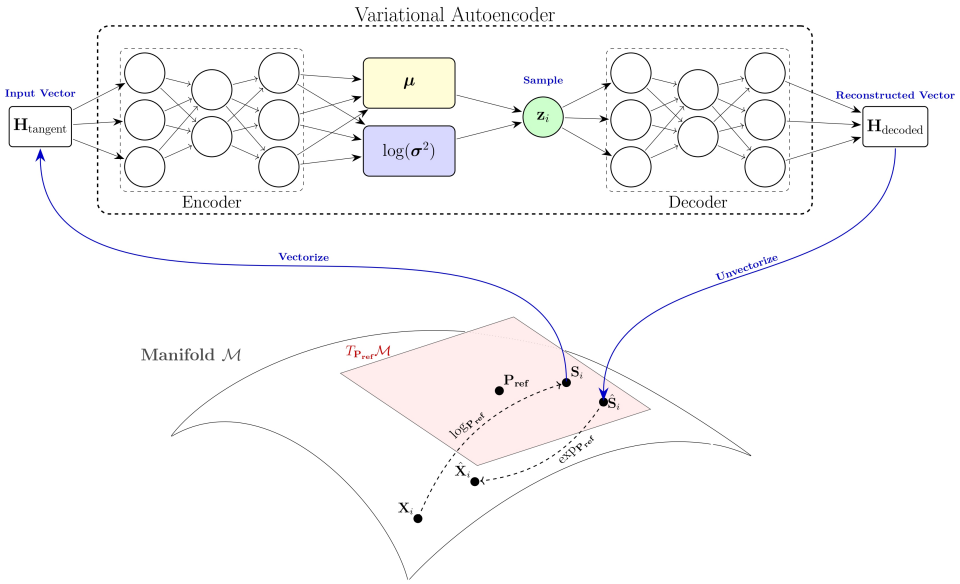

Figure 1: An overview of the proposed RGP-VAE, illustrating the integration of a standard VAE with geometric operations on the SPD manifold. An input SPD matrix $\mathbf{X}_i$ is first projected onto the tangent space at a reference point $\mathbf{P}_{\text{ref}}$ using the logarithmic map $\log_{\mathbf{P}_{\text{ref}}}$ (Eq. 3). This tangent representation $\mathbf{S}_i$ is then vectorized to serve as the encoder input $\mathbf{H}_{\text{tangent}}$. The encoder maps this input to a latent distribution parameterized by μ and $\log(\sigma^2)$, from which a latent vector $\mathbf{z}_i$ is sampled and passed to the decoder to produce the reconstructed vector $\mathbf{H}_{\text{decoded}}$. The vector is unvectorized back into a tangent space representation $\hat{\mathbf{S}}_i$, which is finally mapped back onto the SPD manifold via the exponential map ($\exp_{\mathbf{P}_{\text{ref}}}$) (Eq. 4) to produce the reconstructed SPD matrix $\hat{\mathbf{X}}_i$.

Model Architecture: Conceptually building on prior work on Riemannian variational autoencoders for manifold-valued data [14], the modified VAE (Fig. 1) learns a latent representation $\mathbf{z}$ from SPD matrices by bridging the curved manifold $\mathscr{M}$ and the Euclidean space required by neural networks. The manifold of symmetric positive-definite matrices is defined as $\mathscr{M} = \{\mathbf{X} \in \mathbb{R}^{N\times N} \mid \mathbf{X} = \mathbf{X}^\top, \mathbf{X} \succ 0\}$, where N is the number of EEG channels and $X \succ 0$ indicates that the matrix is composed of strictly positive eigenvalues [2]. The proposed architecture relies on a class-specific reference point $\mathbf{P}_{\text{ref}}$, calculated as the Riemannian Fréchet mean of the training class. Unlike arithmetic mean, which may not yield a valid SPD matrix, the Fréchet Mean guarantees a valid point on the manifold by minimizing the sum of squared Riemannian distances to all other matrices in a set $\{\mathbf{X}_i\}_{i=1}^{M_{\text{set}}}$:

$$\mathbf{G} = \arg\min_{\mathbf{P}\in\mathscr{M}} \sum_{i=1}^{M_{set}} d_r^2(\mathbf{P}, \mathbf{X}_i) \tag{1}$$

where $\mathbf{P}$ is the candidate SPD matrix over which the minimization occurs and $d_r(\mathbf{P}, \mathbf{X}_i)$ is the affine-invariant Riemannian metric (AIRM)[5, 15, 16] which defines the distance between two SPD matrices as:

$$d_r(\mathbf{P}_1, \mathbf{P}_2) = \| \log(\mathbf{P}_1^{-1/2}\mathbf{P}_2\mathbf{P}_1^{-1/2}) \|_F \tag{2}$$

The model processes a batch of aligned SPD covariance matrices $\mathscr{X} = \{\mathbf{X}_1, \ldots, \mathbf{X}_B\}$, with batch size $B = 128$, selected as a balance between computational load and the requirements for the diversity loss. Each $\mathbf{X}_i$ is projected to the tangent space—a local Euclidean approximation—at the class-specific reference point $\mathbf{P}_{\text{ref}}$ using the logarithmic map [1, 2]:

$$\mathbf{S}_i = \log_{\mathbf{P}_{\text{ref}}}(\mathbf{X}_i) = \mathbf{P}_{\text{ref}}^{1/2} \log\left(\mathbf{P}_{\text{ref}}^{-1/2}\mathbf{X}_i\mathbf{P}_{\text{ref}}^{-1/2}\right)\mathbf{P}_{\text{ref}}^{1/2}. \tag{3}$$

This is implemented via batched whitening followed by the matrix logarithm to support numerical stability by centring operations around the identity.

The resulting batch $\mathscr{S} = \{\mathbf{S}_1, \ldots, \mathbf{S}_B\}$ consists of symmetric matrices in the tangent space at $\mathbf{P}_{\text{ref}}$ and each $\mathbf{S}_i$ is vectorized by using only the upper-triangular elements forming the batch of vectors $\mathbf{H}_{\text{tangent}} \in \mathbb{R}^{B\times D_{\text{spd}}}$ (with $D_{\text{spd}} = N(N+1)/2$) as input to the encoder.

The encoder maps this batch of vectors to the parameters $(\mathbf{M}, \log\sigma^2)$ of the latent distribution, where $\mathbf{M} = \{\mu_1, \ldots, \mu_B\}$ and $\log\sigma^2 = \{\log\sigma_1^2, \ldots, \log\sigma_B^2\}$. The encoder consists of five sequential blocks (linear $\rightarrow$ batch normalization [17] $\rightarrow$ LeakyReLU [18]) with dimensions $D_{\text{spd}} \rightarrow 32 \rightarrow 64 \rightarrow 16 \rightarrow 32 \rightarrow 64$, followed by two separate linear projections to produce $\mu_i, \log\sigma_i^2 \in \mathbb{R}^{D_{\text{lat}}}$ where $D_{\text{lat}} = 64$. Batch normalization stabilizes training [17], while LeakyReLU activations preserve representational capacity by preventing inactive neurons [18]. The batch of latent vectors $\mathbf{Z} = \{\mathbf{z}_1, \ldots, \mathbf{z}_B\} \in \mathbb{R}^{B\times D_{\text{lat}}}$ is sampled via the reparameterization trick: $\mathbf{z}_i = \mu_i + \varepsilon_i \odot \exp(0.5 \cdot \log\sigma_i^2)$, where $\varepsilon_i \sim \mathscr{N}(\mathbf{0}, \mathbf{I})$, allowing gradients to flow back through $\mathbf{M}$ and $\log\Sigma^2$ during training.

The decoder MLP mirrors the encoder structure, mapping the batch of latent vectors $\mathbf{Z}$ back to a batch of decoded vectors $\mathbf{H}_{\text{decoded}} \in \mathbb{R}^{B\times D_{\text{spd}}}$, which is subsequently unvectorized into a batch of symmetric matrices $\hat{\mathbf{S}}' \in \mathbb{R}^{B\times N\times N}$. The decoder output $\hat{\mathbf{S}}_i'$ is explicitly re-symmetrized via $\hat{\mathbf{S}}_i'' = (\hat{\mathbf{S}}_i' + (\hat{\mathbf{S}}_i')^T)/2$ to eliminate any asymmetries. To return to the manifold, we apply the Exponential Map to

10.3217/978-3-99161-093-9-014

each matrix:

$$\hat{\mathbf{X}}_i = \exp_{\mathbf{P}_{\text{ref}}}(\hat{\mathbf{S}}'_i) = \mathbf{P}_{\text{ref}}^{1/2} \exp\left(\mathbf{P}_{\text{ref}}^{-1/2} \hat{\mathbf{S}}'_i \mathbf{P}_{\text{ref}}^{-1/2}\right) \mathbf{P}_{\text{ref}}^{1/2} \quad (4)$$

Numerical instability caused by floating-point arithmetic can violate the strict SPD constraints, therefore validity is enforced throughout the model architecture and parallel transport. During the matrix exponential computation, eigenvalues are conditionally scaled (threshold $T = 20$) to prevent overflow: if $\lambda_{\max} > T$, all eigenvalues are scaled by $T/\lambda_{\max}$. Throughout all geometric operations, we maintain a numerical threshold $\varepsilon = 10^{-6}$. If the minimum eigenvalue $\lambda_{\min}$ of any intermediate or output matrix falls below ε, we add $(\varepsilon - \lambda_{\min})\mathbf{I}$ to shift all eigenvalues above the threshold, ensuring positive-definiteness.

Training and Optimization: The network is optimized using a loss function L_{total} balancing reconstruction accuracy, latent space regularization, and diversity:

$$L_{\text{total}} = (L_{\text{manifold}} + L_{\text{tangent}}) + \beta L_{\text{KL}} + \gamma L_{\text{diversity}} \quad (5)$$

The reconstruction term combines L_{manifold} which enforces geometric fidelity using the AIRM distance (Eq. 2):

$$L_{\text{manifold}} = \frac{1}{B} \sum_{i=1}^{B} d_r(\mathbf{X}_i, \hat{\mathbf{X}}_i) \quad (6)$$

and L_{tangent}, which minimized the normalized Euclidean error between original and decoded tangent vectors:

$$L_{\text{tangent}} = \frac{1}{B} \sum_{i=1}^{B} \frac{\sum_j (h_{\text{decoded},i,j} - h_{\text{tangent},i,j})^2}{\sum_j h^2_{\text{tangent},i,j} + \varepsilon} \quad (7)$$

where $\varepsilon = 10^{-6}$ for numerical stability. Meanwhile, latent space regularization is achieved through KL divergence L_{KL} toward a standard Gaussian prior:

$$L_{\text{KL}} = \frac{1}{B} \sum_{i=1}^{B} \left[-0.5 \sum_{k=1}^{D_{\text{lat}}} (1 + \log \sigma^2_{i,k} - \mu^2_{i,k} - \exp(\log \sigma^2_{i,k})) \right] \quad (8)$$

We apply KL cost annealing [19], linearly increasing β from 0.0001 to 0.2 during training to prevent posterior collapse while maintaining reconstruction fidelity.

The diversity loss $L_{\text{diversity}}$ encourages sample diversity by maximizing the geometric volume of generated tangent vectors. Since the determinant of a covariance matrix quantifies the generalized variance (i.e., the volume spanned by data points), maximizing it promotes wider spatial coverage in the tangent space. The loss minimizes the negative log-determinant of the batch covariance of decoded tangent space vectors $\mathbf{H}_{\text{decoded}} \in \mathbb{R}^{B \times D_{\text{spd}}}$:

$$L_{\text{diversity}} = -\log\det(\text{Cov}(\mathbf{H}^T_{\text{decoded}}) + \varepsilon_{\text{cov}}\mathbf{I}) \quad (9)$$

with $\varepsilon_{\text{cov}} = 10^{-6}$ for numerical stability, weighted by an empirically determined $\gamma = 0.035$.

The AdamW optimizer [20] is employed for a fixed 100 epochs with an empirically found learning rate of 1×10^{-4} and a weight decay parameter of 1×10^{-6}. Training is further regularized with gradient clipping (max norm=1.0) and learning rate reduction (factor=0.5) after 20 epochs of stagnation.

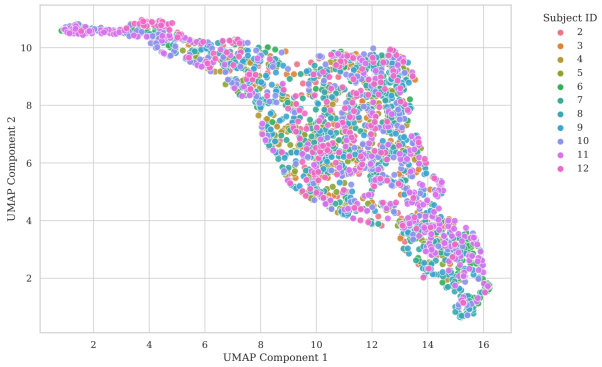

Figure 2: 2D UMAP visualization of the latent space of the RGP-VAE for right-hand movement data. Points are colored by Subject ID; their significant overlap indicates the learning of a subject-invariant representation.

Data Generation and Evaluation Protocol: leave-one-subject-out cross-validation (LOSO-CV) is employed; in each fold, class-specific RGP-VAEs are trained on aligned data from $N-1$ subjects to test generalization to unseen individuals. Two synthetic generation strategies are evaluated: Posterior sampling encodes each training matrix $\mathbf{X}_i$, samples $\mathbf{z}_i$ via reparameterization, and decodes to create variations preserving core characteristics of each sample (1:5 real-to-synthetic ratio). Prior sampling draws $\mathbf{z} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ directly to generate novel samples beyond the training convex hull (5000 per class).

Three Riemannian classifiers—minimum distance to mean (MDM), k-nearest neighbours (KNN; k = 5), and support vector classifier (SVC; C = 1.0)—are trained and evaluated on held-out test subjects under three conditions: (1) baseline using only the original training data, (2) augmented with synthetic data, and (3) synthetic-only training to assess standalone quality. Balanced accuracy, averaged across all folds, serves as the primary metric.

Synthetic data quality is assessed by verifying SPD properties (symmetry and positive-definiteness), comparing statistical variance (element-wise and global) between real and synthetic matrices, and measuring geometric spread via mean pair-wise Riemannian distances within each class. A scrambled-label diagnostic test confirms that performance degrades to chance level, indicating no spurious correlations.

Source code and additional information can be found at `https://641e16.github.io/RGP-VAE/`.

RESULTS

We validate the proposed RGP-VAE through an assessment of the generated synthetic data fidelity, addressing the fundamental question of whether the model can produce valid and realistic covariance matrices, and with a comparison against a standard VAE approach. Our final analysis quantifies the impact of this data on cross-

10.3217/978-3-99161-093-9-014

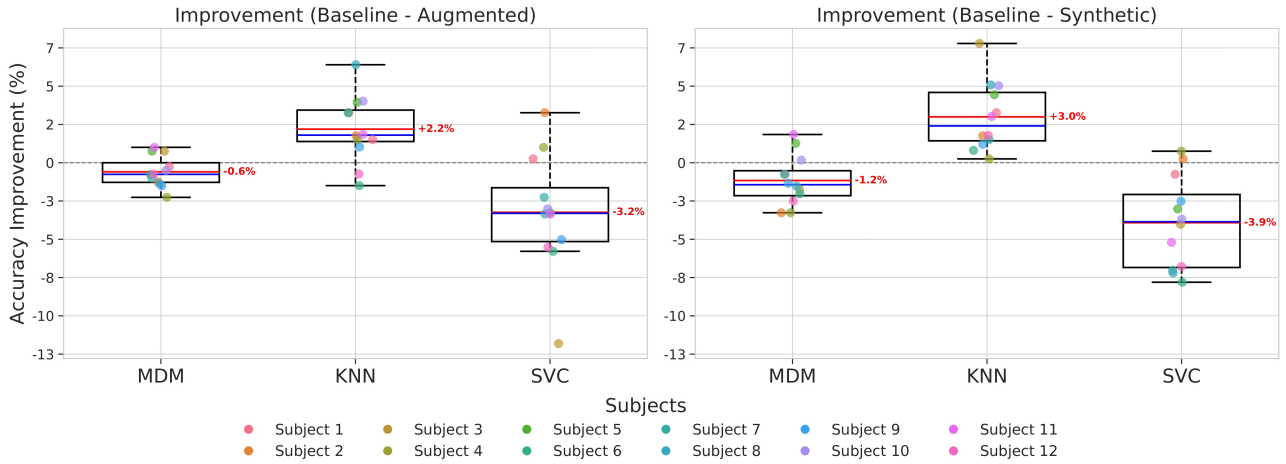

Figure 3: Distribution of accuracy improvement for each classifier using the prior generator. The plot shows the percentage point difference between the 'Augmented' and 'Synthetic-Only' conditions relative to the 'Baseline' across all subjects. The red line signifies the mean whilst the blue line is the median.

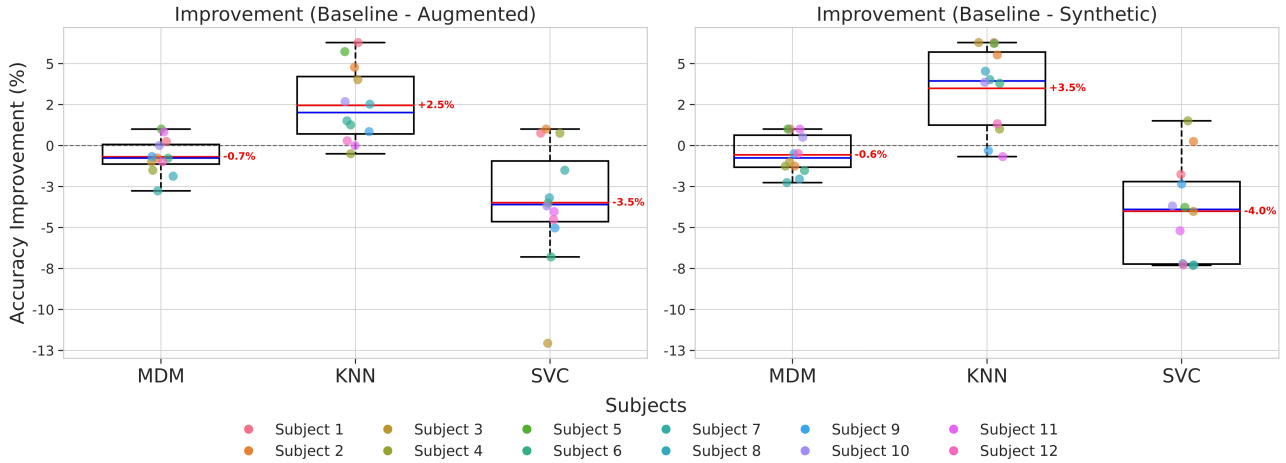

Figure 4: Distribution of accuracy improvement for each classifier using the posterior generator, showing similar trends to the prior generator but with more pronounced fluctuations.

subject classification performance to determine the practical value of the proposed method.

Latent Space Structure: UMAP [21] visualization (Fig. 2) reveals that latent codes organize into a unified structure where subjects are heavily intermingled rather than clustered by individual. This suggests the model learned a largely subject-invariant representation—a critical property enabled by parallel transport alignment—implying generated samples will reflect generalized task patterns rather than subject-specific details.

Fidelity Assessment: Across all folds, 100% of synthetic matrices from both prior and posterior generators passed symmetry and positive-definiteness verification checks, confirming the effectiveness of the architecture's geometric constraints and numerical stabilisation steps. Tab. 1 compares the statistical variance and geometric spread of synthetic data relative to the original. The chosen γ maintained statistical variance close to the original. To address the lower geometric diversity of synthetic samples, the noise vector is scaled by $\varepsilon_i = 2.2$ during generation, increasing the mean intra-class Riemannian distance to ≈ 1.95, closely matching the original data's spread (2.03) without distorting statistical properties.

Cross-Subject Classification Performance: The impact of data augmentation was evaluated by comparing classification accuracies under different augmentation conditions using Wilcoxon signed-rank tests with Bonferroni correction ($p < 0.0083$). As detailed in Tab. 2, data augmentation produced divergent effects. For the KNN classifier, augmentation consistently and significantly improved performance. Posterior-based synthetic-only training yielded the largest gain (+3.49%, $p = 0.002$), while augmented training provided +2.45% ($p = 0.002$). Prior generation produced similar but slightly smaller significant benefits (+3.00% synthetic-only, $p < 0.001$; +2.19% augmented, $p = 0.003$). In contrast, SVC performance significantly degraded with augmentation (up to -4.01%, $p = 0.002$), while MDM remained largely unaffected. Figs. 3 and 4 illustrate subject-wise distributions, revealing high variability: KNN augmentation yielded

10.3217/978-3-99161-093-9-014

Table 1: Fidelity analysis of synthetic data averaged across 12 folds. The table compares the statistical variance ratio and the mean intra-class Riemannian distance, showing that the synthetic data distribution is valid.

Generator	**Statistical Variance**		**Geometric Diversity**	
	Original	**Synthetic (Ratio)**	**Original**	**Synthetic**
Prior	0.208	0.221 (1.061)	2.032	1.946
Posterior	0.208	0.221 (1.063)	2.032	1.918

Table 2: Average balanced accuracy (%) across 12 subjects for all training conditions and generators with corresponding p-values.

Generator	**Classifier**	**Baseline**	**Augmented Scenario**			**Synthetic-Only Scenario**		
		Acc. (%)	**Acc. (%)**	**Improvement**	**p-value**	**Acc. (%)**	**Improvement**	**p-value**
Prior	MDM	59.52 ± 5.52	58.92 ± 5.40	-0.59%	0.092	58.36 ± 5.03	-1.16%	0.043
	KNN	53.19 ± 4.00	55.38 ± 4.17	+2.19%	**0.003**	56.19 ± 4.19	+3.00%	$<$ **0.001**
	SVC	60.67 ± 5.33	57.43 ± 6.32	-3.24%	0.016	56.75 ± 6.37	-3.92%	**0.002**
Posterior	MDM	59.52 ± 5.52	58.83 ± 5.29	-0.69%	0.092	58.95 ± 5.51	-0.57%	0.151
	KNN	53.19 ± 4.00	55.64 ± 4.13	+2.45%	**0.002**	56.68 ± 4.06	+3.49%	**0.002**
	SVC	60.67 ± 5.33	57.18 ± 6.57	-3.48%	**0.007**	56.66 ± 6.25	-4.01%	**0.002**

gains up to +7.8% for subject no.3 (prior generation, synthetic only condition). A scrambled label test confirmed classifiers learned meaningful features, yielding chance-level average accuracy on scrambled data (prior: MDM = 50.2%, KNN = 49.9%, SVC = 52.3%; posterior: MDM = 46.6%, KNN = 50.5%, SVC = 47.2%).

Standard VAE Comparison: To validate the Riemannian framework, we compared the proposed RGP-VAE against a standard Euclidean VAE. The standard VAE failed to generate valid data, with $> 40\%$ of outputs in every fold violating positive-definiteness. Furthermore, augmenting with the valid portion of its data significantly degraded MDM performance (-9.49%, $p < 0.001$) and offered no statistically significant benefit to KNN or SVC. This confirms that the proposed architecture's geometric constraints are essential for generating valid and useful SPD matrices in this domain.

DISCUSSION

This study investigated whether the proposed RGP-VAE could generate high-fidelity EEG covariance matrices to improve cross-subject MI-BCI classification.

Generative Fidelity and Validity: A primary contribution of this work is confirming that the RGP-VAE framework inherently generates valid SPD matrices—a non-trivial task where standard Euclidean VAEs failed (producing $> 40\%$ invalid matrices). This success is attributable to the underlying Riemannian geometry that enforces the SPD constraint by design. While Parallel Transport enabled a subject-invariant latent space for cross-subject generalization, computing the target-subject reference mean still requires a minimal calibration step prior to online deployment.

While valid, the synthetic data exhibited a slightly elevated statistical variance (ratio ≈ 1.06) but reduced geometric diversity (ratio ≈ 0.95). With the chosen parameters ($\gamma = 0.035$, $\varepsilon_i = 2.2$), the model generated more prototypical samples concentrated near the class geometric means rather than spanning the full outlier range of real-world data.

Classifier-Dependent Utility: The impact of the synthetic data was highly divergent, revealing that data augmentation utility is not universal but classifier-dependent. Augmentation yielded statistically significant improvements for the KNN classifier, with posterior sampling boosting performance up to $+3.49\%$ ($p = 0.002$). KNN likely benefits because the prototypical synthetic samples densify the class manifolds, creating more dense and reliable local neighbourhoods for distance-based classification. Conversely, performance significantly degraded for the SVC (up to -4.01%, $p = 0.002$). The reduced diversity of synthetic data likely caused the SVC to learn decision boundaries too narrowly fitted around class centres, reducing generalization to boundary-case real samples. Meanwhile, performance for the MDM classifier remained stable—a positive result compared to the standard VAE, which caused a massive degradation (-9.49% under posterior, synthetic-only condition). Unlike the naive Euclidean approach that failed to even generate valid SPD matrices, the RGP-VAE preserved the SPD validity and successfully learnt Riemannian class means. Beyond immediate classification impacts, synthesizing this data holds broader practical value; it provides a mechanism to test pipeline scalability, mitigates data scarcity—possibly for data hungry models—and enables privacy protection by avoiding raw signal sharing.

Future Research Directions: This study provides a foundational proof of concept that opens avenues for future research. Building on these findings, future work may explore advanced manifold sampling techniques, such as Riemannian Hamiltonian VAEs to capture complex latent distributions more faithfully [22]. Additionally, as demonstrated by vEEGNet [23], integrating the RGP-VAE's geometric constraints and subject-invariance with discriminative frameworks could potentially yield

10.3217/978-3-99161-093-9-014

latent spaces that are simultaneously geometrically valid, subject-invariant and class-discriminative. Finally, while relying solely on training metrics and geometric spread to determine parameters successfully prevented data leakage, future work would benefit from a systematic hyperparameter search.

CONCLUSION

This paper developed and validated a novel Riemannian Geometry-Preserving VAE (RGP-VAE) for generating synthetic EEG covariance matrices in the challenging cross-subject MI-BCI context. The RGP-VAE is not only capable of consistently generating valid SPD matrices—overcoming the limitations of standard VAEs—but also closely matches the original data diversity. The high-fidelity synthetic data can maintain or even significantly improve classification performance for specific classifiers. However, divergent classifier results highlight generative capabilities on the SPD manifold do not guarantee universal downstream improvements.

REFERENCES

[1] Congedo M, Barachant A, Bhatia R. Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review. Brain-Computer Interfaces. 2017;4(3):155–174.
[2] Yger F, Berar M, Lotte F. Riemannian approaches in brain-computer interfaces: A review. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2017;25(10):1753–1762.
[3] Chevallier S et al. The largest eeg-based bci reproducibility study for open science: The moabb benchmark. 2024. arXiv: 2404.15319 [eess.SP].
[4] Blankertz B, Dornhege G, Krauledat M, Müller KR, Curio G. The non-invasive Berlin Brain-Computer Interface: fast acquisition of effective performance in untrained subjects. NeuroImage. 2007;37(2):539–550.
[5] Kalunga E, Chevallier S, Barthélemy Q. Data augmentation in Riemannian space for Brain-Computer Interfaces. In: STAMLINS 2015 proceedings. Lille, France, Jun. 2015.
[6] Kingma DP, Welling M. Auto-encoding variational bayes. 2013. arXiv: 1312.6114 [stat.ML]. [Online]. Available: https://arxiv.org/abs/1312.6114.
[7] Shao H, Kumar A, Fletcher PT. The riemannian geometry of deep generative models. CoRR. 2017;abs/1711.08014.
[8] Yair O, Ben-Chen M, Talmon R. Parallel transport on the cone manifold of SPD matrices for domain adaptation. IEEE Transactions on Signal Processing. 2019;67(7):1797–1811.
[9] Faller J, Vidaurre C, Solis-Escalante T, Neuper C, Scherer R. Autocalibration and recurrent adaptation: Towards a plug and play online ERD-BCI. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2012;20(3):313–319.
[10] Aristimunha B et al. *Mother of all BCI Benchmarks*. Version 1.0.0. 2023. [Online]. Available: https://github.com/NeuroTechX/moabb.
[11] Schirrmeister RT et al. Deep learning with convolutional neural networks for brain mapping and decoding of movement-related information from the human EEG. CoRR. 2017;abs/1703.05051.
[12] Zhu H, Forenzo D, He B. On the deep learning models for EEG-based brain-computer interface using motor imagery. IEEE Transactions on Neural Systems and Rehabilitation Engineering. 2022:1–1.
[13] Chen Y, Wiesel A, Eldar YC, Hero AO. Shrinkage algorithms for MMSE covariance estimation. IEEE Transactions on Signal Processing. 2010;58(10).
[14] Miolane N, Holmes SP. Learning weighted submanifolds with variational autoencoders and riemannian variational autoencoders. 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2019:14491–14499.
[15] Fletcher PT, Lu C, Pizer SM, Joshi S. Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Transactions on Medical Imaging. 2004;23(8):995–1005.
[16] Moakher M. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM Journal on Matrix Analysis and Applications. 2005;26(3):735–747.
[17] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. CoRR. 2015;abs/1502.03167.
[18] Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. In: Proceedings of the 30th International Conference on Machine Learning. 2013.
[19] Bowman SR, Vilnis L, Vinyals O, Dai A, Jozefowicz R, Bengio S. Generating sentences from a continuous space. In: Proceedings of the 20th SIGNLL Conference on Computational Natural Language Learning. Association for Computational Linguistics: Berlin, Germany, Aug. 2016, 10–21.
[20] Loshchilov I, Hutter F. Fixing weight decay regularization in adam. CoRR. 2017;abs/1711.05101.
[21] McInnes L, Healy J, Melville J. Umap: Uniform manifold approximation and projection for dimension reduction. Journal of Open Source Software. 2018;3(29):861.
[22] Chadebec C, Thibeau-Sutre E, Burgos N, Allassonnière S. Data augmentation in high dimensional low sample size setting using a geometry-based variational autoencoder. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2023;45(3):2879–2896.
[23] Zancanaro A, Cisotto G, Zoppis I, Manzoni SL. Veegnet: Learning latent representations to reconstruct eeg raw data via variational autoencoders. In: Information and Communication Technologies for Ageing Well and e-Health. Springer Nature Switzerland: Cham, 2024, 114–129.