Detection of Epileptic Seizures with Tangent Space Mapping Features of EEG Signals

Fatih ALTINDİŞ
Electrical and Computer Engineering
Abdullah Gül University
Kayseri, Turkey
fatih.altindis@agu.edu.tr

Bülent YILMAZ
Electrical and Electronics Engineering
Abdullah Gül University
Kayseri, Turkey
bulent.yilmaz@agu.edu.tr

2021 Medical Technologies Congress (TIPTEKNO) | 978-1-6654-3663-2/21/$31.00 ©2021 IEEE | DOI: 10.1109/TIPTEKNO53239.2021.9632959

***Abstract*—Detection of epileptic seizures from EEG signals is well-studied topic for the last couple of decades. Lately, automated signal processing and machine learning methods were developed to detect epileptic seizures. However, most of the methods are tailored to subjects and require fine tuning of many parameters. In this study, we proposed to use Riemannian geometry-based signal processing method that already showed superior performance on brain-computer interface problems, to extract features. We showed that tangent space mapping features of EEG signals can be used to detect seizures with high accuracy and precision.**

Keywords—Riemannian geometry, Tangent space mapping, EEG, Seizure detection

I. INTRODUCTION

According to the 2019 World Health Organization report, approximately 50 million people diagnosed with epilepsy, one of the most common neurological disease. It is also reported that proper diagnose and treatment can prevent 70% of the patients from having seizures [1]. Epilepsy patients have two or three times more premature death risk than healthy people [1], which highlights importance of diagnose and treatment of epilepsy. Detection of epileptic seizures and localization of seizures have vital importance for proper treatment [2].

Electroencephalography (EEG) is the most common way to monitor brain activity thanks to its non-invasiveness, low equipment cost and high temporal resolution [3]. EEG signals are recorded from scalp surface with electrodes, and it represents electrical potential differences between cortical surfaces. This allows to understand how different cortical areas behave during a task or event, making EEG the most preferred technique to monitor and capture beginning and end of the epileptic seizures as well as locating the affected area on the brain [4]. By monitoring EEG signals, it can be understood whether specific region of the brain is involved with seizure, or it is widespread.

In general, hours of EEG recordings are taken from patients, and then experts visually examine all the recorded data to detect seizures and annotate them. This is an inefficient, time-consuming, and exhausting process for both patients and experts. On the other hand, considering the duration of recordings, artifact and noise caused by daily activities of the patient, are present in EEG signals making it difficult to detect seizures automatically [5]. Further, seizure characteristics vary across subjects that prevent standardization in the signal processing and machine learning steps of automation [6]. Yet, for the last couple of decades, many research efforts have been put forward to automate seizure detection process [7, 8]. Time domain and frequency domain methods are used to extract features from EEG signals and neural networks are utilized for classification [9]. Though existing methods reach high accuracy in terms of detecting seizures, performance is highly dependent to fine tuning of parameters, resulting a low generalization capability across different subjects and recordings [4].

Here in this study, we used Riemannian geometry-based analysis which is becoming a *state-of-the-art* method in brain-computer interface field, for feature extraction from EEG signals to detect whether a seizure is present or not. Rest of the paper organized as follows: in Section 2 we briefly introduced how to use Riemannian geometry on EEG signals. Section 3 explains epilepsy dataset and signal processing steps we followed. Signal processing and classification results were shared in Section 4, and we discussed outcomes of the study in Section 5.

II. RIEMANNIAN GEOMETRY

Riemannian geometry deals with smoothly curved (differentiable) spaces that behave like Euclidean spaces. EEG signals can be mapped into this geometric space with a suitable metric to create manifolds [10, 11]. For every point on the manifold, linear approximation of Riemannian manifold called 'tangent space' that behaves like Euclidean space exists. Earth is the real-life example of this concept, where locally any chosen point has Euclidean properties but in large scale point to point metrics becomes smooth curves.

Consider an EEG signal epoch with M channels and N time samples in it, where any i^{th} epoch represented as $\boldsymbol{X}_i \in \mathbb{R}^{M\times N}$. Such a signal can be used to estimate sample covariance matrix (SCM) as in (1), after preprocessing each channel in order to center around zero. Any SCM of EEG signals then will be symmetric along the diagonal have positive values and resulting matrices have $M(M+1)/2$ different elements in it.

$$\boldsymbol{P}_i = \frac{1}{N-1}\boldsymbol{X}_i\boldsymbol{X}_i^T \in \mathbb{R}^{M\times M} \qquad (1)$$

978-1-6654-3663-2/21/$31.00 ©2021 IEEE

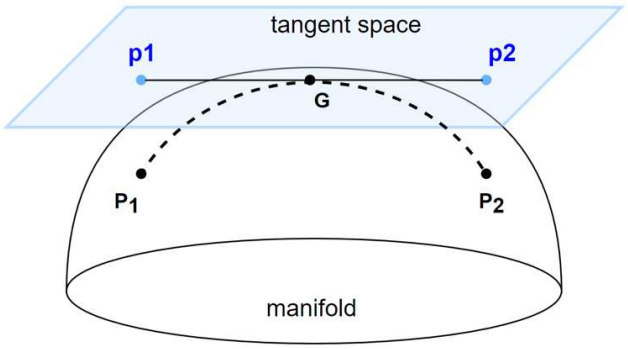

Figure 1. Two black points represent P1 and P2 in Riemannian manifold. Mean of these points represented as point G. Their projection to tangent space represented as p1 and p2 blue points.

Any two point $\mathbf{P}_1$and $\mathbf{P}_2$ that are belong to two EEG signal epochs, can be represented on Riemannian manifold as in Figure 1. Riemannian distance between these two points calculated as follows [12];

$$\delta_R(\boldsymbol{P}_1, \boldsymbol{P}_2) = \|\mathrm{Log}(\boldsymbol{P}_1^{-1}\boldsymbol{P}_2)\|_F$$
$$= \left\|\mathrm{Log}(\boldsymbol{P}_1^{-1/2}\boldsymbol{P}_2\boldsymbol{P}_1^{-1/2})\right\|_F = [\textstyle\sum_{i=1}^{M} \log^2 \lambda_i]^{1/2} \quad (2)$$

where $\|\cdot\|_F$ is Frobenius norm of matrix and λ_i is the eigenvalue of $\mathbf{P}_1^{-1}\mathbf{P}_2$ and $\mathrm{Log}(\cdot)$ is matrix logarithm in (2). Riemannian distance $\delta_R(\cdot)$ returns the shortest curve between given two points and it is called 'geodesic'. Geodesics can be used to find mean or 'center of mass' of the manifold, which is important and useful for two reasons; first, it allows to use minimum distance to Riemannian mean (MDRM) for classification [13]. Second, mean of manifolds can be used as reference points to project manifold back to Euclidean space with tangent space mapping (TSM). Riemannian manifolds can be projected to tangent space at a chosen reference point. Riemannian mean chosen as a reference point for tangent space mapping, since Riemannian mean of all the SPD matrices in manifold is the result of optimization problem in (3), where point $\boldsymbol{G}$ will be the Riemannian mean that minimizes the sum of distances of K points of the manifold [14].

$$\arg\min_{\boldsymbol{G}} \frac{1}{K}\textstyle\sum_{i=1}^{K} \delta_R^2(\boldsymbol{P}_i, \boldsymbol{G}) \quad (3)$$

Once the reference point $\boldsymbol{G}$ for tangent space is chosen, each SPD matrix can be mapped to tangent space $\mathcal{T}(M)$ as follows;

$$\mathcal{T}(M) = \boldsymbol{s}_i = upper\left(\mathrm{Log}\left(\boldsymbol{G}^{-1/2}\boldsymbol{X}_i\boldsymbol{G}^{-1/2}\right)\right),$$
$$\boldsymbol{s}_i \in \mathbb{R}^{M(M+1)/2} \quad (4)$$

where $upper(\cdot)$ is vectorization operator that takes upper triangular part of the matrix. Figure 1 shows tangent space mapping of two points $\mathbf{P}_1$ and $\mathbf{P}_2$ with respect to reference point $\boldsymbol{G}$.

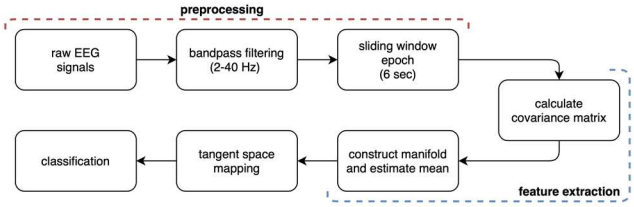

Figure 2. Signal processing pipeline of tangent space mapping for raw EEG recordings.

III. Materials and Methods

A. Dataset and Software

In this study, we used an EEG dataset named as 'TUH EEG Seizure Corpus' provided by Temple University Hospital [15]. EEG recordings of this dataset were taken from 692 different subjects and total duration of the records is more than 1000 hours. Although available channels and sampling rate varies across subjects, majority of the recordings were taken with 256 Hz sampling rate and from at least 20 channels located according to international 10-20 system. All the records were annotated by experts with two labels: '*background*' and '*seizure*' and label information provided with a text file. In total, 7.4% of the recordings were belong to seizure class and rest was belong to background in other saying non-seizure class. Among all the available channels, we chose following 18 channels that are available across all the recordings: FP1, FP2, F7, F3, FZ, F4, F8, C3, CZ, C4, P3, P4, T5, T3, T4, T6, O1, O2. This helped to extract features from same locations and to have same feature space for all the subjects.

For the Riemannian geometry analysis, python library named '*PyRiemann*' was used [16]. Multilayer perceptron learner of '*sklearn*' machine learning library in python was trained and tested with extracted features [17].

B. Feature Extraction and Classification

The Signal processing has two main steps: preprocessing and feature extraction. For the preprocessing, we utilized 5th order IIR bandpass filter with 2-40 Hz cutoff frequencies, resulting EEG signals centered around zero and does not contain any powerline noise. Next, we used 6 seconds-wide sliding windows with 3 seconds overlap to divide EEG signals into epochs.

For the feature extraction, each epoch of 18 channel EEG signal was used to calculate sample covariance matrix provided in (1). Riemannian manifolds were constructed, and Riemannian mean of the manifold was estimated using covariance matrices. Mean of the manifold was chosen as reference point for tangent space mapping for all the epochs. After projecting all the covariance matrices to estimated tangent space, feature vector created by using (4), resulting 171 features for each epoch. Main steps of the signal processing and feature extraction are summarized in figure 2.

Additionally, shrinkage covariance matrices were calculated to create another manifold. Shrinkage covariance matrix helps to reduce outliers in covariance estimations and provide regularized covariance matrices that are centered to identity matrix. Considering diversity in the dataset where EEG recordings were taken from hundreds of subjects, introducing a

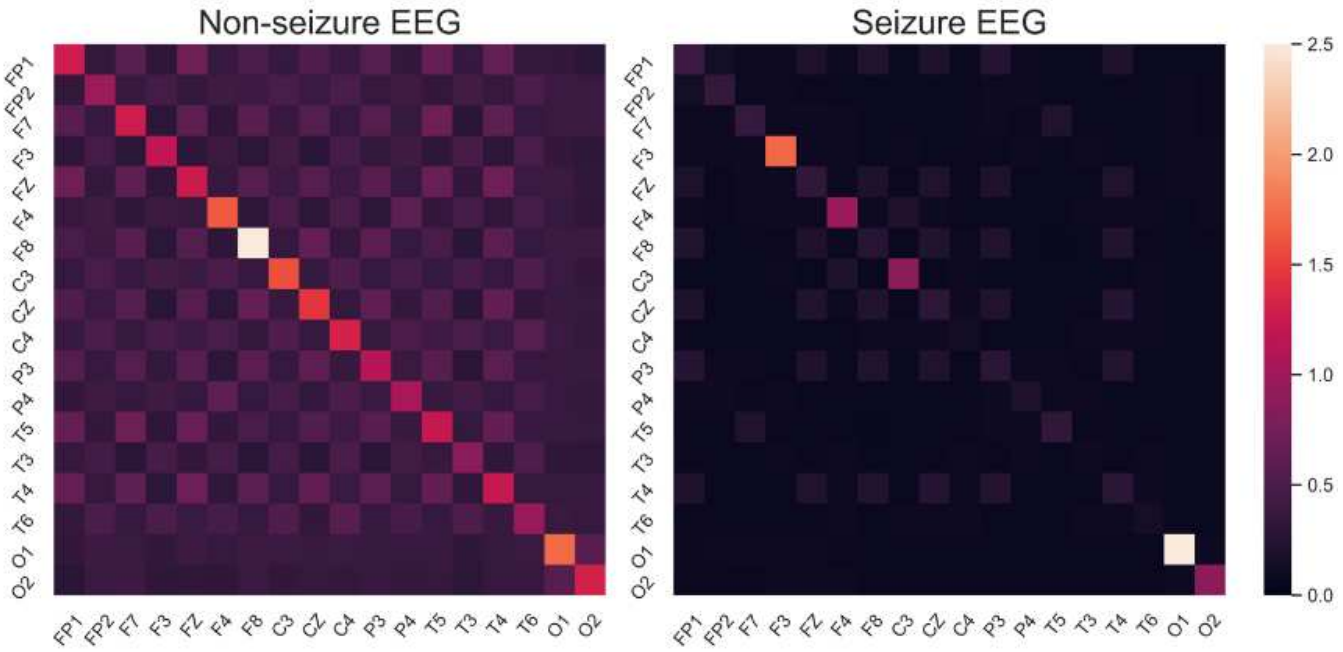

Figure 3. Heatmaps of average sample covariance matrices of non-seizure and seizure epochs. While brighter colors indicate higher covariance values, darker colors indicate lower covariance values as shown in colorbar at right-most.

regularization coefficient to covariance matrices would be helpful to exclude artefacts and outliers [18]. Shrinkage estimations on the covariance matrices were completed using (5), where $\boldsymbol{I}_M$ is identity matrix, $\boldsymbol{P}_{shrunk}$ is shrunk covariance of $\boldsymbol{P}$ and α is shrinkage coefficient.

$$\boldsymbol{P}_{shrunk} = (1-\alpha)\boldsymbol{P} + \alpha \frac{tr(\boldsymbol{P})}{M} \boldsymbol{I}_M,$$
$$\boldsymbol{I}_M, \boldsymbol{P} \text{ and } \boldsymbol{P}_{shrunk} \in \mathbb{R}^{M \times M} \qquad (5)$$

IV. Results

Here, we first compared covariance matrices of non-seizure and seizure epochs. Figure 3 shows heatmaps of average covariance matrices for both classes. Seizure related epochs have lower covariance values compared to non-seizure EEG epochs. This indicated that seizure related epochs have lower covariance between channels compared to non-seizure related epochs.

TABLE I. Classification Performance of Tangent Space Mapping Features with Different Covariance Matrices

Covariance Matrices	*Accuracy*	*Kappa*
SCM	**98.91%**	**0.913**
SCM with feature selection	98.80%	0.907
Shrinkage Cov. Mat.	98.63%	0.889
Shrinkage Cov. Mat. with feature selection	97.76%	0.817

Extracted TSM features are used for training and test of the classifier. Multi-layer perceptron classification results in table 1 show accuracy and kappa score versus two different covariance matrix estimations. Sample covariance matrices yielded better results than shrinkage covariances. Furthermore, variance threshold as a feature selection method was applied to reduce number of features. Feature numbers were reduced to 55 from 171, but it did not improve classification performance.

We have used three different reference points to extract TSM features which are mean of all the points, mean of seizure class points and mean of non-seizure class points. Classification performances were similar as they were shown in table 2. Considering the imbalance between seizure and non-seizure class, results showed that TSM can be used to detect seizures with high accuracy and precision.

TABLE II. Classification Performance of Tangent Space Mapping Features with Different Reference Points

Reference point	*Accuracy*	*Kappa*
Mean of the manifold	98.91%	0.913
Mean of the non-seizure epochs	**98.94%**	**0.916**
Mean of the seizure epochs	98.93%	0.916

V. Discussion

In this study, we have adapted tangent space mapping approach of Riemannian geometry analysis to epileptic seizure EEG data in order to detect seizures. Our first aim was to investigate applicability and performance of Riemannian geometry analysis on seizure data. Thus, all the signal processing steps were kept as simple as possible. Once all the covariance matrices were calculated, they were used to estimate class means and Riemannian mean of the manifold. Extracted TSM features were used in classification step without any scaling or normalization operation. Classification performance reached a high kappa score as 0.9.

When there is small number of epochs available, using shrinkage covariance matrices yields better decoding performance, because outliers and artefacts can be recentered to

identity matrix [18]. However, as we shared in table 1, use of shrinkage covariance matrices resulted in lower kappa scores. Since we have hundreds of different subjects with hours of recording, shrinking covariance matrices increased similarity between seizure and non-seizure epochs taken from the same subjects and resulted in lower classification performance.

We have used three different reference points for TSM of the manifold. If we assume that data points were clustered into two separate groups in Riemannian manifold, choosing Riemannian mean of the manifold as a reference point gave better projection on tangent space. In addition, considering the imbalance between seizure and non-seizure we have estimated mean of both classes and used them as reference points for TSM. Although there were slight improvements on precision and kappa score when class means were used for TSM, classification performances were not significantly different ($p>0.05$) than each other. Since we have not employed any regularization parameter to align and recenter cross-subject epochs on Riemannian manifold, these results were expected. There are transfer learning methods available in Riemannian manifold to reduce cross-subject variabilities and increase separability between classes. It is worthwhile to utilize similar approaches on this dataset in future studies [19].

The dataset we employed in this study, was used in a signal processing challenge back in 2020. One of the participant groups extracted time and spectral features and utilized an oversampling technique to balance classes. They used decision tree-based ensemble machine learning for the classification and reached 67% overall accuracy [20]. Winner of the challenge employed bidirectional long short term memory network and applied blind source separation to extract features from EEG signals [21]. Additionally, rather than using unipolar channel data, participants of the challenge preferred to create bipolar channels from unipolar recordings. Here, we used unipolar raw EEG signals as they recorded and only filtered with 2-40 Hz band-pass filter. Compared to these studies, Riemannian geometry analysis offers simpler signal processing pipeline. Despite having an imbalance between seizure and non-seizure epochs, no resampling or down sampling operations were employed. Still, classification with TSM features reached precision rate that is higher than any of the participants of the challenge.

This study showed that tangent space mapping features of Riemannian manifold can be used to detect whether given EEG epoch has a seizure or not. Despite having many different subjects and varied recording qualities, high kappa scores were achieved using TSM features.

REFERENCES

[1] "Epilepsy." World Health Organization. https://www.who.int/news-room/fact-sheets/detail/epilepsy (accessed 26/07/2021, 2021).

[2] D. Buck, G. A. Baker, A. Jacoby, D. F. Smith, and D. W. Chadwick, "Patients' experiences of injury as a result of epilepsy," (in English), *Epilepsia,* vol. 38, no. 4, pp. 439-44, Apr 1997, doi: 10.1111/j.1528-1157.1997.tb01733.x.

[3] H. J. Hwang, S. Kim, S. Choi, and C. H. Im, "EEG-Based Brain-Computer Interfaces: A Thorough Literature Survey," (in English), *Int J Hum-Comput Int,* vol. 29, no. 12, pp. 814-826, Dec 2 2013, doi: 10.1080/10447318.2013.780869.

[4] I. Ullah, M. Hussain, E. U. Qazi, and H. Aboalsamh, "An automated system for epilepsy detection using EEG brain signals based on deep learning approach," (in English), *Expert Systems with Applications,* vol. 107, pp. 61-71, Oct 1 2018, doi: 10.1016/j.eswa.2018.04.021.

[5] U. R. Acharya, S. V. Sree, G. Swapna, R. J. Martis, and J. S. Suri, "Automated EEG analysis of epilepsy: A review," (in English), *Knowl-Based Syst,* vol. 45, pp. 147-165, Jun 2013, doi: 10.1016/j.knosys.2013.02.014.

[6] T. McShane, "A clinical guide to epileptic syndromes and their treatment," *Archives of Disease in Childhood,* vol. 89, no. 6, p. 591, 2004.

[7] T. Zhang, W. Z. Chen, and M. Y. Li, "AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier," (in English), *Biomed Signal Proces,* vol. 31, pp. 550-559, Jan 2017, doi: 10.1016/j.bspc.2016.10.001.

[8] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," (in English), *Expert Systems with Applications,* vol. 32, no. 4, pp. 1084-1093, May 2007, doi: 10.1016/j.eswa.2006.02.005.

[9] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis," (in English), *Ieee T Inf Technol B,* vol. 13, no. 5, pp. 703-10, Sep 2009, doi: 10.1109/TITB.2009.2017939.

[10] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Riemannian Geometry Applied to BCI Classification," in *Latent Variable Analysis and Signal Separation*, Berlin, Heidelberg, V. Vigneron, V. Zarzoso, E. Moreau, R. Gribonval, and E. Vincent, Eds., 2010// 2010: Springer Berlin Heidelberg, pp. 629-636.

[11] M. Congedo, A. Barachant, and A. Andreev, "A New Generation of Brain-Computer Interface Based on Riemannian Geometry," p. arXiv:1310.8115.

[12] R. Bhatia, *Positive Definite Matrices*. Princeton University Press, 2009.

[13] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten, "Multiclass brain-computer interface classification by Riemannian geometry," *IEEE Trans Biomed Eng,* vol. 59, no. 4, pp. 920-8, Apr 2012, doi: 10.1109/TBME.2011.2172210.

[14] M. Congedo, A. Barachant, and R. Bhatia, "Riemannian geometry for EEG-based brain-computer interfaces; a primer and a review," (in English), *Brain-Comput Interfa,* vol. 4, no. 3, pp. 155-174, 2017, doi: 10.1080/2326263x.2017.1297192.

[15] I. Obeid and J. Picone, "The Temple University Hospital EEG Data Corpus," *Front Neurosci,* vol. 10, p. 196, 2016, doi: 10.3389/fnins.2016.00196.

[16] A. Barachant, *pyRiemann v0.2.2*. Zenodo, 2015.

[17] F. Pedregosa *et al.*, "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research,* vol. 12, pp. 2825-2830, 2011 2011.

[18] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain-computer interfaces: a 10 year update," (in English), *J Neural Eng,* vol. 15, no. 3, Jun 2018, doi: 10.1088/1741-2552/aab2f2.

[19] P. L. C. Rodrigues, C. Jutten, and M. Congedo, "Riemannian Procrustes Analysis: Transfer Learning for Brain-Computer Interfaces," (in English), *Ieee T Bio-Med Eng,* vol. 66, no. 8, pp. 2390-2401, Aug 2019, doi: 10.1109/Tbme.2018.2889705.

[20] L. Wei and C. Mooney, "Epileptic Seizure Detection in Clinical EEGs Using an XGboost-based Method," in *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 5-5 Dec. 2020 2020, pp. 1-6, doi: 10.1109/SPMB50085.2020.9353625.

[21] C. Chatzichristos *et al.*, "Epileptic Seizure Detection in EEG via Fusion of Multi-View Attention-Gated U-Net Deep Neural Networks," in *2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, 5-5 Dec. 2020 2020, pp. 1-7, doi: 10.1109/SPMB50085.2020.9353630.