Graph-CSPNet

Artyom Matveev

Moscow Institute of Physics and Technology

matveev.as@phystech.edu

November 11, 2023

Model's structure[1]

1. SPDNet[2]. A neural network that operates on SPD matrices.
    - BiMap layer. It transforms the input SPD matrices to new SPD matrices by a bilinear mapping.
    - ReEig layer. It rectifies the SPD matrices by tuning up their small positive eigenvalues.
    - LOG layer. It maps an SPD matrix $\mathbf{S}$ onto its tangent space at identity matrix $\mathbf{I}$.
2. Riemannian Batch Normalization[3].

[1] **Ju, C.**, & Guan, C. Graph Neural Networks on SPD Manifolds for Motor Imagery Classification: A Perspective From the Time–Frequency Analysis. IEEE Transactions on Neural Networks and Learning System. 2023

[2] **Huang, Z.**, & Van Gool, L. A Riemannian Network for SPD Matrix Learning. AAAI-2017

[3] **Brooks, D.**, Schwander, O., et al. Riemannian batch normalization for SPD neural networks. NeurIPS 2019

Graph construction

1. $\mathbf{X} \in \mathbb{R}^{n_C \times n_T}$ – an EEG signals trial
2. $\mathbf{S} = \mathbf{X}\mathbf{X}^\top \in \mathcal{S}_{++}$ – an SPD matrix
3. $\mathbf{S} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{U}^\top$ – an SPD matrix representation using eigenvalue decomposition
4. $d_{g^{\text{AIRM}}}(\mathbf{S}_1, \mathbf{S}_2) = d_{g^{\text{AIRM}}}(\mathbf{W}\mathbf{S}_1\mathbf{W}^\top, \mathbf{W}\mathbf{S}_2\mathbf{W}^\top)$, where $\mathbf{W}$ is weight matrix of BiMap transformation <u>with the full-row rank</u>
5. $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ – a time-frequency graph:
   - $\mathcal{V}(\mathcal{G}) := \{\mathbf{S}_i = S(\Delta t_i \times \Delta f_i)\}$, where $\{S(\Delta t_i \times \Delta f_i)\}_{i \in \mathcal{I}}$ is the set of SPD matrices under the specific time and frequency constraints.
   - $$\mathcal{E}(\mathcal{G}) := \mathbf{A} = \begin{cases} e^{-d^2_{g^{\text{AIRM}}}(\mathbf{S}_i, \mathbf{S}_j)/t}, & \text{if } \mathbf{S}_i \text{ and } \mathbf{S}_j \text{ are adjacent} \\ 0, & \text{others} \end{cases}$$

     where $e^{(\cdot)}$ is the RBF kernel and preset Gaussian kernel width $t > 0$.

Graph BiMap Layer

Each GNN layer updates the following way:

$$H^{(l+1)} \leftarrow \mathrm{RBN}\left(\mathrm{ReEig}\left(\mathbf{W}^{(l)}(\bar{\mathbf{D}}^{-1}\bar{\mathbf{A}}^{(l)})H^{(l)}\mathbf{W}^{(l)^\top}\right)\right),$$

where $\bar{\mathbf{A}}^{(l)} := \mathbf{A}^{(l)} + \mathbf{I}_N$, $\bar{\mathbf{D}}_{ii} := \sum_j \bar{\mathbf{A}}_{ij}^{(l)}$, $H^{(l)} \in \mathbb{R}^{|\mathcal{V}|\times n_C^2}$, $\bar{\mathbf{A}}^{(0)}$ is the adjacency matrix of the time-frequency graph, and $\bar{\mathbf{A}}^{(l)} := \mathbf{I}_N$, for $l \geq 1$.

ReEig layer

This layer performs $\mathbf{U}\max(\epsilon\mathbf{I}, \mathbf{\Sigma})\mathbf{U}^\top$, where ϵ is a rectification threshold, and $\mathbf{I}$ denotes an identity matrix.

LOG layer

This layer maps matrix $\mathbf{S}$ onto its tangent space at identity matrix $\mathbf{I}$ using $\mathbf{U}\log(\mathbf{\Sigma})\mathbf{U}^\top$

SPDNet and GraphCSP-Net illustrations