Qeios

Peer Review

Review of: "STEAM-EEG: Spatiotemporal EEG Analysis with Markov Transfer Fields and Attentive CNNs"

Levy Boccato[1]

1. Universidade Estadual de Campinas, Brazil

Summary: This work proposes STEAM-EEG, a framework aimed at processing electroencephalogram (EEG) signals and solving different tasks, such as epilepsy and sleep disorder analysis. The proposed model has two main branches: in the top branch, the EEG signals are decomposed into trend and seasonal components, which are processed by a 1D-CNN with attention layers; on the other hand, the bottom branch processes the images produced by MTF (Markov Transition Fields) via a 2D-CNN, also using attention layers. Then, all feature maps are combined and finally fed into an MLP classifier. STEAM-EEG was applied to some EEG-related tasks (10 datasets), and the obtained results were quite promising. In general, the article is well-written and brings an interesting combination of ideas to achieve an effective EEG processing pipeline.

Comments and suggestions

Figure 1: For completeness, the authors could insert the accuracy values obtained by the proposed approach in each corner.

Section 2: I believe it is more common to write SVM as “support-vector machines.”

Section 3:

- Even though the subsections after Figure 2 bring the detailed description of each part of the proposed approach, I believe it would be useful to expand the caption of that figure by incorporating a brief explanation of the framework in order to facilitate the interpretation of the diagram.

doi.org/10.32388/CCZMQX

- Additional comments on Figure 2:
  - What does "*Cardinal*" mean?
  - Fix the name of the “Co-split Attention” block.
  - The authors should highlight that the upper and lower blocks encompassing Co-Split Attention and Convolutional Layers are 1D and 2D, respectively.
  - It is not clear what happens in the fusion layer. Are the feature maps (or vectors) simply concatenated? Is there any pooling (like Global Average Pooling) before fusion?
  - Additionally, the authors could connect the overall idea of the proposed approach with the notion of *late fusion*, a strategy for combining in a late stage feature representations derived from multimodal data: each modality is processed individually, and only close to the end of the pipeline are they brought together.
  - How many MTF images are produced from the set of C EEG channels (or signals)?
  - How many branches of Cardinal + Co-split Attention are necessary?
  - It seems that the title "multi-level feature extraction" applies only to the lower block. Is it correct?

Section 3.1

- “*capturing the underlying temporal structure and separating different components*” – please include here a reference to detailed material about SSA.
- x_1, ..., x_N are scalars or vectors?
- What is the definition of *significance* and *similarity* in the context of SSA? How exactly are the singular values grouped?
- “*By applying SSA to the raw EEG time series, we obtain three reconstructed component series: trend, seasonal, and noise.*” – Does SSA always yield three groups of singular values/eigenvectors (and, thus, three reconstructed series)? The authors could expand a little bit on the exposition of SSA to fulfill this gap.
- Finally, it is not clear if SSA is simultaneously applied to each EEG signal (channel); in fact, in Section 5, the authors mention that the work was focused on single-channel EEG analysis, but (if so) such information should be provided earlier in the document.

Section 3.2

- This type of attention mechanism – Eqs. (5) and (6) – can be found in other works (e.g., with 2D-CNNs in computer vision tasks), or is it entirely new?

Section 3.3

- The authors could include an example or an illustration of the 2D topographic map with a relatively simple interpretation. Hence, readers who are not familiar with MTF can catch a glimpse of the intuition and the information captured by such MTF images.

Section 3.4

- "*The 2D-ResNet architecture is well-suited for processing image data and has demonstrated excellent performance in various computer vision tasks*" – please include a few references.
- Figure 3 was not rendered. The authors should check this issue.
- Correct Equation (10):

$$X_{l-1}$$

- "*we introduce a cross-channel split attention mechanism*." -> Perhaps it would be better to say that an attention mechanism similar to that applied in the 1D-CNNs is explored within the 2D-ResNets.

Section 4.1

- Suggestion: insert a table with the main characteristics of each dataset (number of samples (training, validation, and test), number of classes, duration of EEG signals, sampling rate, etc.).

Section 4.2

- References [43] and [44] correspond to the same work. Please correct [44] to [1].
- The four selected methods are based on deep/convolutional neural networks, so it is necessary to correct the last sentence: "*These baselines represent a diverse range of approaches in EEG signal analysis, from traditional machine learning methods to advanced deep learning architectures.*"
- Although the datasets are all EEG-based, they are related to classification scenarios with their own particularities and literature. Some of the chosen methods were designed for a given task/scenario and may not be strong competitors in the other cases. For instance, EEGNet was originally formulated and applied to BCI under different paradigms (SSVEP, P300, and motor imagery). But even in the BCI context for motor imagery, I am not sure if EEGNet still represents the state-of-

the-art (SOTA). Moreover, its competitiveness probably depends on the underlying experimental protocol and the related dataset. So, I think the selection of the models should be a little more justified.

- In some tasks, like motor imagery, I feel that other largely employed datasets could be explored (instead of that contained in the UCR Archive), and would allow a faster comparison with the literature (by reporting the best results found in these cases) – see, for instance, BCI Competition IV 2a/2b, MOABB (https://moabb.neurotechx.com/docs/index.html) and a more recent dataset (https://www.nature.com/articles/sdata2018211). Therefore, there might be other methods (and datasets) that better represent the SOTA.

Section 4.4

- Experimental setup and hyperparameters: even though the code is available, I believe the manuscript should present at least the main configurations of the model and of the training procedure.
- The authors could also mention how sensitive the approach is with respect to some parameters (for instance, kernel size, learning rate, etc.).
- Table 1: are the obtained values associated with a single or several executions of each method? It would be a nice improvement to include standard deviation, boxplots, and probably run pairwise statistical tests to indicate the significance of the differences among performances.

Section 4.5

- "*Notably, the MTF Imaging component exhibited heightened importance in the MotorImagery and SelfRegulationSCP datasets*" – from Table 2, I am not sure if these would be the only datasets mostly impacted by MTF. This observation was based on other information or experimental perceptions the authors had?

Section 4.6

- Figure 4:
  - The authors should explain a little bit what is exhibited in Figure 4: what does each subtitle indicate? Why 4 subfigures?
  - The x and y axes should have titles (and, if valid, the unit).
  - Time should appear in seconds.

- Figure 5: please indicate the dataset / EEG-based task involved in this analysis.

Section 5

- "*Our current methodology focuses primarily on single-channel EEG analysis, which may limit its applicability to more complex EEG datasets*" – In which sense the study was limited to single-channel EEG analysis? I think expanding this comment would enlighten the limitation and it should be clear since the beginning of the text.

References

1. ^*Vernon J Lawhern, Amelia J Solon, Nicholas R Waytowich, Stephen M Gordon, et al. (2018). EEGNet: a compact convolutional neural network for EEG-based brain–computer interfaces. J. Neural Eng., vol. 15 (5), 056013. doi:10.1088/1741-2552/aace8c.*

Declarations

Potential competing interests: No potential competing interests to declare.