Technical Report

Sultan[1]

[1]Affiliation not available

September 7, 2018

Limitation

There are many challenges dealing with fMRI data: First, high dimensionality of the feature compared with the small number of samples, which usually yield to overfitting, the noisiness of the signal. Further adding to this complexity, each human brain has a different response to the same cognitive task. So in order for the machine learning model to learn the mapping between multi-voxel activity and the cognitive state, it is required to find a way to minimize these complexities. In this study, we have reduced the number of features by focusing on a subset of the brain then transforming the fMRI signal to other domain or by extracting a new discernible features that represent the voxels pattern.

Methods

ROIs Selection

The dataset consist of 25 anatomically defined regions of interests (ROIs), only seven of these regions were involved with the starplus experiment.[1]. In this work, after analyzing the performance of different regions, it has been shown that "CALC" is the most discriminative region.

Signal Processing Techniques

In order to help the model to discriminate the features successfully, we have applied different feature creation methods to transform the raw signal into new features that have higher predictive power. For within-subject classification, the created features should reflect the activity evoked by each stimulus.

Fourier Transform

Previous work used Fourier transform for a hypothesis that the signal might have impulse response pattern correspond to each event [1]. Here, we have used the phase information of the Discrete Fourier Transform (DFT) for each trial to generate the phase difference between the time-series signal. The DC component is generated by taking the first value of the fourier transformed signal. Furthermore, we generated two features by selecting the maximum phase and the frequency of the maximum phase of the DFT.

Wavelet Transform

To extract the wavelet features, we have used the Maximal Overlap Discrete Wavelet Packet Transform (MODWPT) with three vanishing moments.

Semi-Classical Signal Analysis

SCSA is used to decompose the fMRI signal into a set of eigenfunctions and eigenvalues of the Schrodinger operator. Using SCSA, we extract five different features from the seven ROIs after performing the following steps:

1. Re-scale each trial between [0,1] using min-max normalization.
2. Take the average voxels over the seven ROIs speretaly to yield seven supervoxels for each snapshot.
3. Concatenate the snapshots in each trial to form 112 features.

Voxel Weight-based Matrix (VWM)

The VWM is a representation of the voxel pattern in fMRI signal. It helps to generate a new set of powerful discriminative features that describe the probability of the voxel signal for each stimulus. The generation steps of the VWM is done by the following:

- Quantize the continuous voxels signals into a discrete set of predefined levels, the choice of the intervals is based on the standard deviation of the frequency distribution for a single subject.
- Construct the Voxel Probability Matrix (VPM) (one for each category) by dividing each quantized integer by the number of levels. Hence, we obtain two matrices of size number_of_levels*number_voxels where each column in the VPM can be viewed as a multinomial distribution.
- The quantized matrices then scan over the VPMs to convert each level in the matrix to its weighted score.
- By summing up all the weighted scores we get the two features, for this study we have two features since we have only two labels.

We should note that the VWM assumes independence between the test and training samples by excluding the test sample from the construction of VPM.

Model Training and Testing

After the extraction of the features, we need to judge how well these features will perform. There are many methods to evaluate the data, and it usually depends on the data to be tested. In fMRI analysis, Gaussian Naive Bayes (GNB) and Logistic Regression (LR) are commonly used classifiers. In our case, they have resulted in the highest accuracy.

Since we have only a few samples in our dataset, we should select an appropriate validation method to make sure we have an unbiased estimation. Consequently, we have split the data using a leave-one-out cross-validation scheme where every sample will be tested.

References

[1]Just, M., and T. Mitchell. “StarPlusfMRI data.” URL http://www. cs. cmu. edu/afs/cs. cmu. edu/project/theo-81/www 39.62 (2001): 103.[Accessed: May 30, 2018].

[2] Li, Yung-hui, and Tien-ho Lin. “Neural Signal Processing for fMRI Data”.