

Deep Learning for multilabel classification of fNIRS data to predict cognitive workload.

06/19/2019

1. Data Format and Pre-processing

Data representing change in absorption of IR(>810nm) and Red(<710nm) optical streams by cerebral blood flow is recorded using a Hitachi ETG-4000 fNIRS device. Each of these 52-channels record raw optical densities observed at different locations on the subjects' brain as shown in Figure 1.

Cerebral blood flow change in terms of oxygenated and de-oxygenated hemoglobin (del Hb and del HbO) is obtained using Modified Beer Lambert Law [...]. The resulting del Hb and del HbO measurements are then individually reshaped into 5x22 arrays to replicate the sensor topology as shown in Figure 1, and in turn preserving the spatial dependency of the input data.

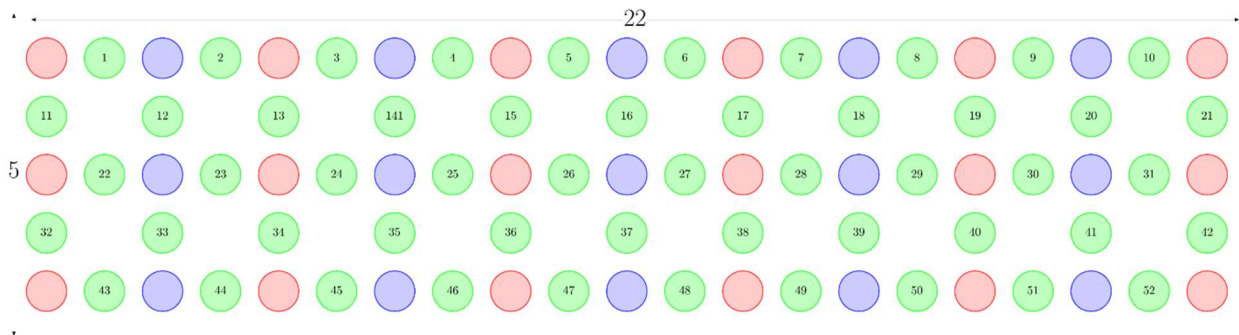


Figure 1: Sensor topology of fNIRS headset

These del Hb and del HbO values for an experiment lasting T seconds are then band pass filtered at 0.5 and 1.5 Hz, detrended using a GLM. The preprocessed del Hb and del HbO values are later horizontally stacked to form 5x44 dimensional arrays. The resulting dimensions of the pre-processed data are (T*sampling frequency)x5x44.

2. Deep Learning Models

Neural Networks (CNN/DNNs) are well known universal function approximators and as such they have a large design space. Moreover, our data has long term dependences across time domain as well as spatial dependencies w.r.t. the location of sensor probes Figure 1. With this knowledge at hand we decided to experiment with 2 families of deep neural networks viz.

volumetric convolutions i.e. 3D CNNs [1] and LSTMs with Temporal Pooling [2]. Each of the models are trained with Adam optimizer and a learning rate decay strategy which reduces learning by half if the training reaches a plateau before maximum amount of training epochs have been completed. In order to obtain confidence scores for multi-label classification of the inputs a Sigmoidal Layer at the final output layer has been employed.

3. Volumetric convolutions with 3D CNNs

3D CNNs are used to extract features from both spatial and temporal dimensions by utilizing alternating layers 3D convolutions and 3D pooling. Research has shown that it can take approximately 6-8 seconds for a change in HbO concentration indicative of brain activation [3]. Following upon that combined with 10-fold cross validation we concluded that pooling with 3D kernels of dimensions 65x1x2 resulted in the best performance. This added a constraint to the input data format causing the depth of an input to be a multiple of 65. In order to resolve this issue, we rounded the depth of input data to a multiple of 65 and duplicate the remainder number of rows from the beginning of the experiment.

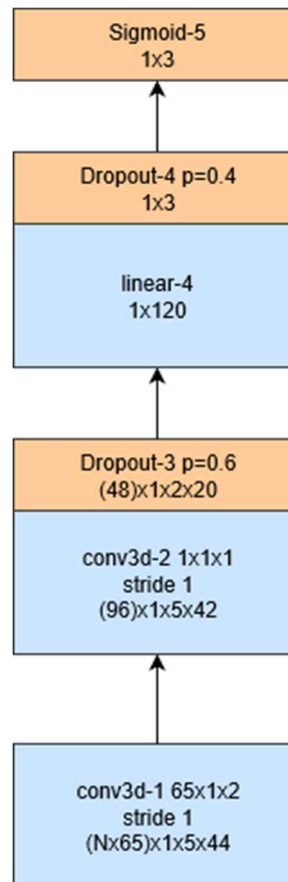


Figure 2: Architecture of 3D CNN model. first row: layer name and kernel size, second row: stride, third row: input dimensions

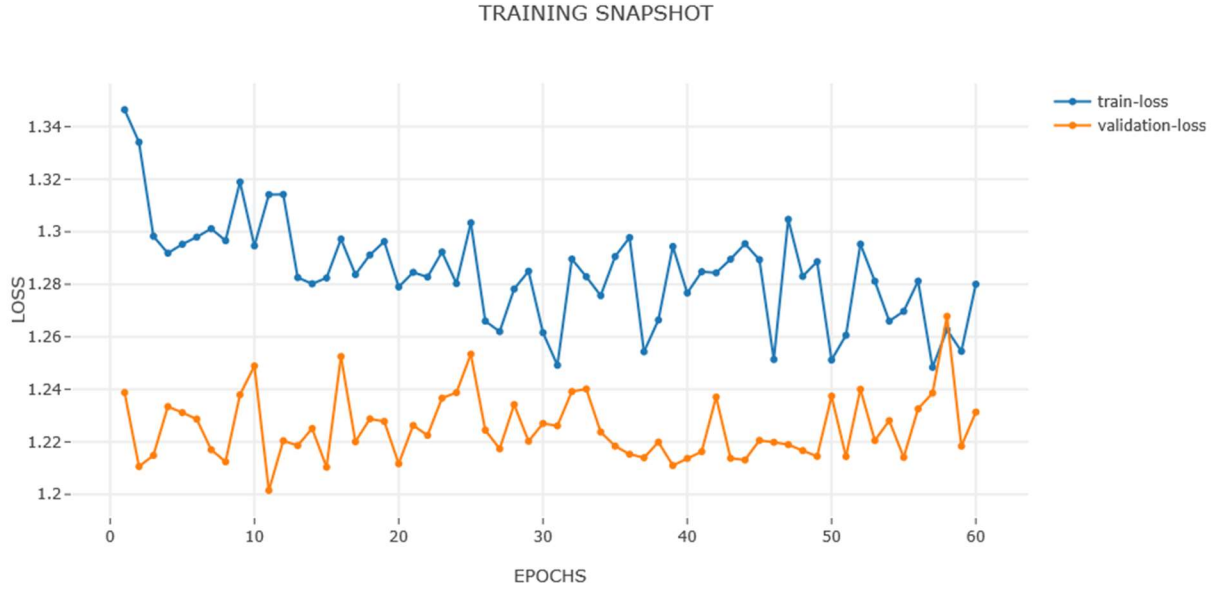


Figure 3: Training Snapshot of 3D CNN Model for 60 epochs

Conclusion and next steps

While a 3D CNN model does exhibit convergence during training, the model seems to perform poorly on the validation set, which is a strong indicator for overfitting. Our next steps are to experiment with **3D CNNs with pixel wise attention** [4] and LSTMs with fewer timesteps than our current 3D CNN model to prevent overfitting.

References

- [1] Learning Spatiotemporal Features with 3D Convolutional Network - *Computer Vision and Pattern Recognition 2014*, Du Tran, Lubomir Bourdev, Rob Fergus, Lorenzo Torresani, Manohar Paluri.
- [2] Action Recognition by Learning Deep Multi-Granular Spatio-Temporal Video Representation – *Microsoft Research*, Qing Li, Zhaofan Qiu, Ting Yao, Tao Mei, Yong Rui, Jiebo Luo.
- [3] Functional near-infrared spectroscopy. *IEEE engineering in medicine and biology magazine* 25(4): 54–62, Bunce SC, Izzetoglu M, Izzetoglu K, Onaral B and Pourrezaei K (2006).
- [4] SCA-CNN: Spatial and Channel-wise Attention in Convolutional Networks for Image Captioning. *Computer Vision and Pattern Recognition 2017*, Long Chen, Hanwang Zhang, Jun Xiao, Liqiang Nie, Jian Shao, Wei Liu, Tat-Seng Chua.