EEG Seizure Analysis

Badea Adrian Catalin, Alexandru-Vlad Adam

University of Bucharest

Introduction

In this project, we address two tasks. The first one is discriminating between **seizures** and **non-seizure time** series on EEG (electroencephalogram) from multi-channel signals.

The second task is discriminating between **pre-seizure** time series and other samples. Some genuine results here would signify the possibility to predict if a seizure is going to happen before it even happens.

The analysis was done on a random subset of the **CHB-MIT** dataset, using convolutional neural networks and deep learning. Multiple approaches were tested on the dataset. [2]

- Training on pure scaled signal
- Using feature extraction and manipulation

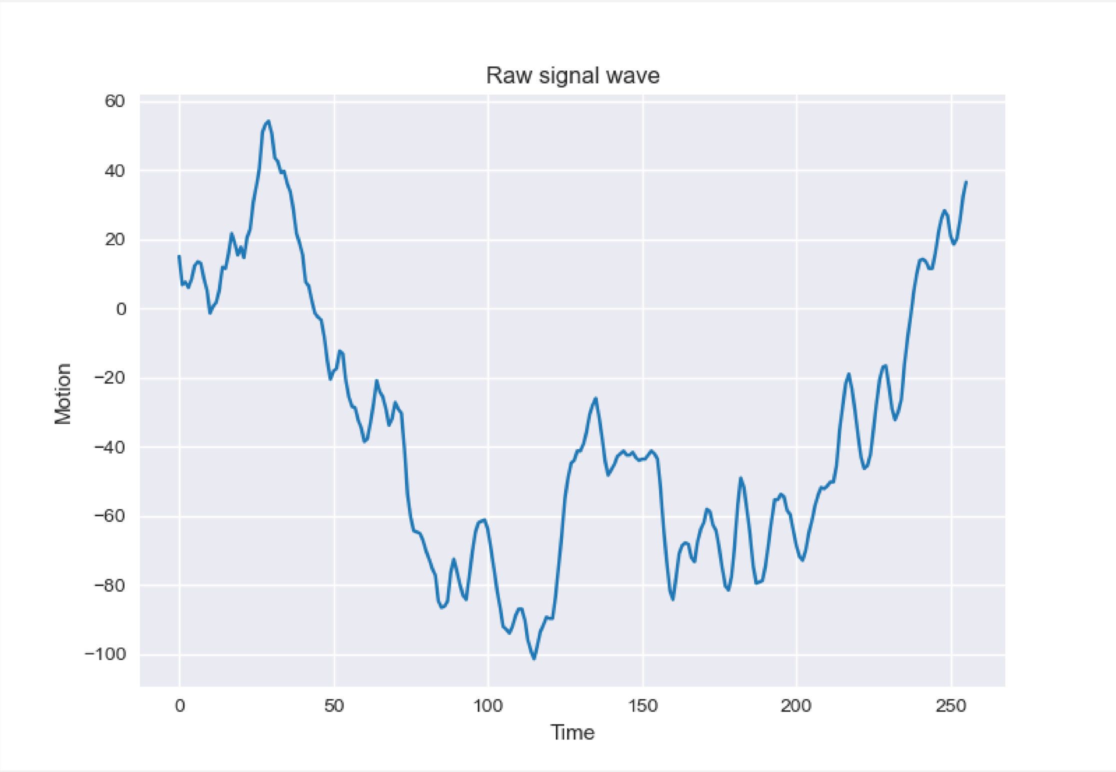
Dataset Description

This database, collected at the Children's Hospital Boston, consists of EEG recordings from pediatric subjects with intractable seizures.

Recordings were grouped in 23 cases. Each case contains more continuous signal taken from patients. All signals were sampled at 256 samples per second with 16-bit resolution. Most files contain 23 EEG signals (24 or 26 in a few cases). All in all, almost one thousand hours of recordings are provided for the purpose of this dataset.

As the initial dataset was too large for the purpose of this binary classification task, the non-seizure and pre-seizure time series were subsampled in two different datasets.

Figure: Raw Signal



Kaggle dataset

Seizure Analysis dataset.

The signals from the original EDF files were split according to the given information about the seizure segments of time.

The ratio of non-seizures to seizures timestamps was, initially 300:1, and thus, we decided to reduce it by a factor by a factor of 60. We also reduced every sample to only 23 channels.
Overall, 54 thousands samples were chosen. These were split in train, test and validation.

After subsampling, approximately 80% of train, test and validation data are negative samples. So, the dataset was still unbalanced but can provide an accurate insight into models. On the next tables we will address it as "unb." [1]

We have also created a balanced dataset for validation by eliminating random samples of negatives until they were 1:1.
Predictive Seizure Analysis.
For the second dataset, we have chosen samples for every seizure with 10-20 seconds before, and random samples not from these intervals.
Then, we have also created a balanced validation set for this dataset, to determine true accuracy

Architectures

The chosen loss for training this binary classificator is NLL.
We have experimented with three different model approaches, with different types of input data, three optimizers (SGD, Adam and AdaBound) and distinctive learning rates.

- **1D-Small-CNN with FCC**
  For this baseline model, we used two 1-dimensional convolutions and four fully connected at l
- **1D-CNN on Raw Signal or Short** Fourier Transformation.
  The model performs multiple 1D convolutions on STFT or signal data. The convolution data is then passed into a fully connected neural network.
- **2D-CNN on MFCCs or Melspectograms.**
  The model is based on the VGG16 CNN architecture. The signals are preprocessed into MFCCs, and are passed into the model, with multiple 2D convolutions transforming the data. Finally, the computed d-vector is fed through a fully-connected neural network.

We moved from simple models (1D CNNs) with raw data to preprocessed data often used in signal engineering (such as spectrograms or MFCCs) that required stronger models in order to predict accurately. 2D-CNNs were starting to show better results, but the best ones were obtained on a VGG16-like architecture.

Observation

Training on raw signals is not the most efficient, nor does it give the best results. Every raw signal is characterised by a bigger input than data after feature extraction.
We have started from simple models with simple convolutions. Then, we have experimented with batch normalization and max pooling, and dropout.
The idea of the deep neural network comes from the co-domain VGG, whose many layers and well-designed

Feature extraction methods

For every feature extractor method, we have used Librosa from Python.

Features	Features per channel	Best unb. results	Best b.
Raw Input	256	0.91	0.881
STFT	269	0.925	0.91
MFCC	20	**0.961**	**0.937**
Melspectogram	128	0.897	

Table: Features

- **STFT**
  The STFT represents a signal in the time-frequency domain by computing discrete Fourier transforms (DFT) over short overlapping windows.
- **Melspectogram**
  A mel spectrogram is a spectrogram where the frequencies are converted to the mel scale
- **MFCC**
  The envelope of the time power spectrum of the speech signal is representative of the vocal tract and MFCC (which is nothing but the coefficients that make up the Mel-frequency cepstrum) accurately represents this envelope.

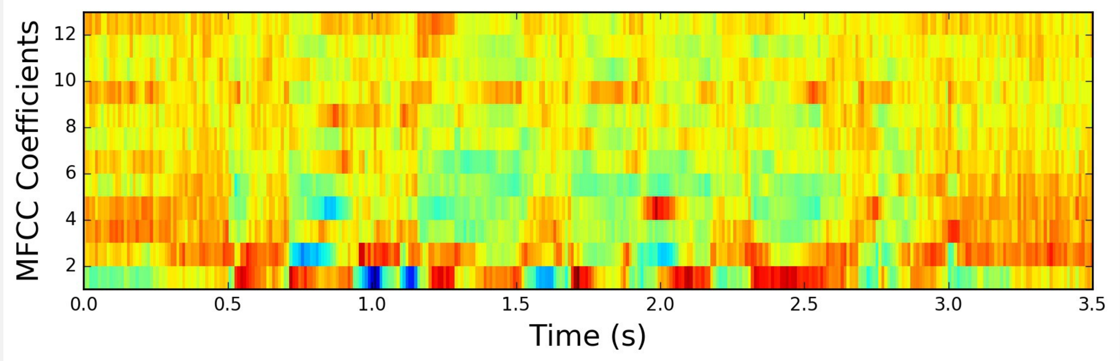


Figure: MFCC

Results

We have compiled a summary table and graph showing the different results from our experiments.

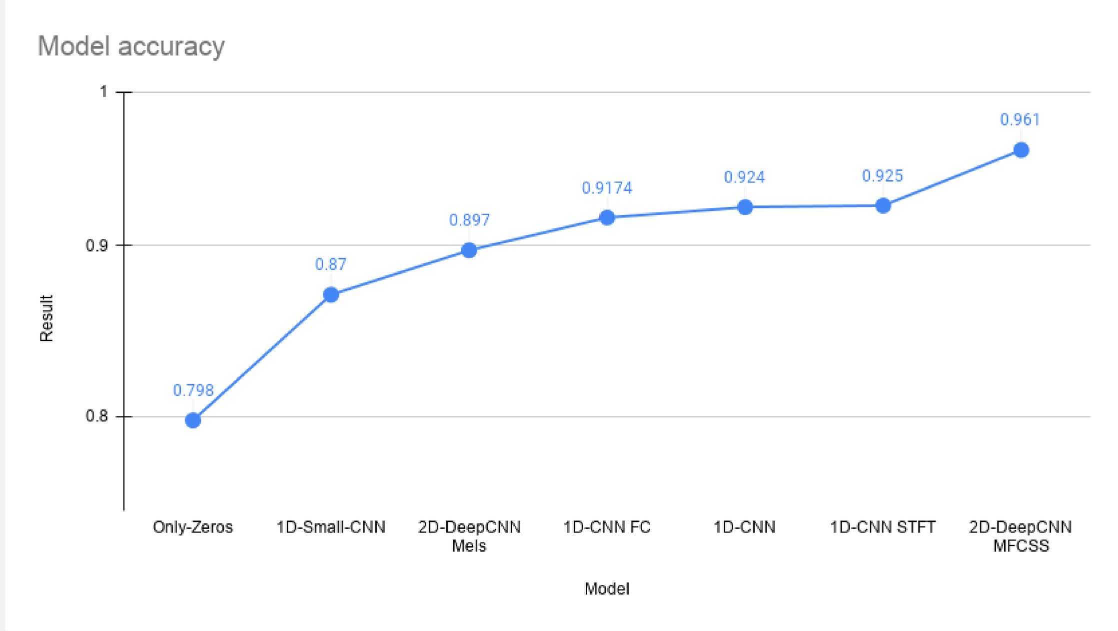


Figure: Accuracy

The 2D-CNNs had the best outcome, followed closely by the 1D-CNNs.

Model	Epochs	Opt.	Best unb.	Best b.
Only-Zeros-Baseline	-	-	0.798	0.50
1D-Small-CNN-Baseline	50	Adam	0.87	0.854
1D-CNN with fully connected	50	SGD	0.9174	0.881
1D-CNN (scaled data)	65	SGD	0.924	
1D-CNN of STFT features	50	Adam	0.925	
2D-DeepCNN on MFCSS	75	Adam	**0.961**	**0.937**
2D-DeepCNN on Melspectograms	50	Adam	0.897	

Table: Experimental results

Training

All of the models train successfuly, and reduce their loss accordingly. We regularized the models with Batch Normalization, Dropout etc. and scaled the data in order to improve accuracy. Once the peaks for each model were reached, we investigated more powerful models such as VGG16 on MFCC data. The models were modified for the task at hand.

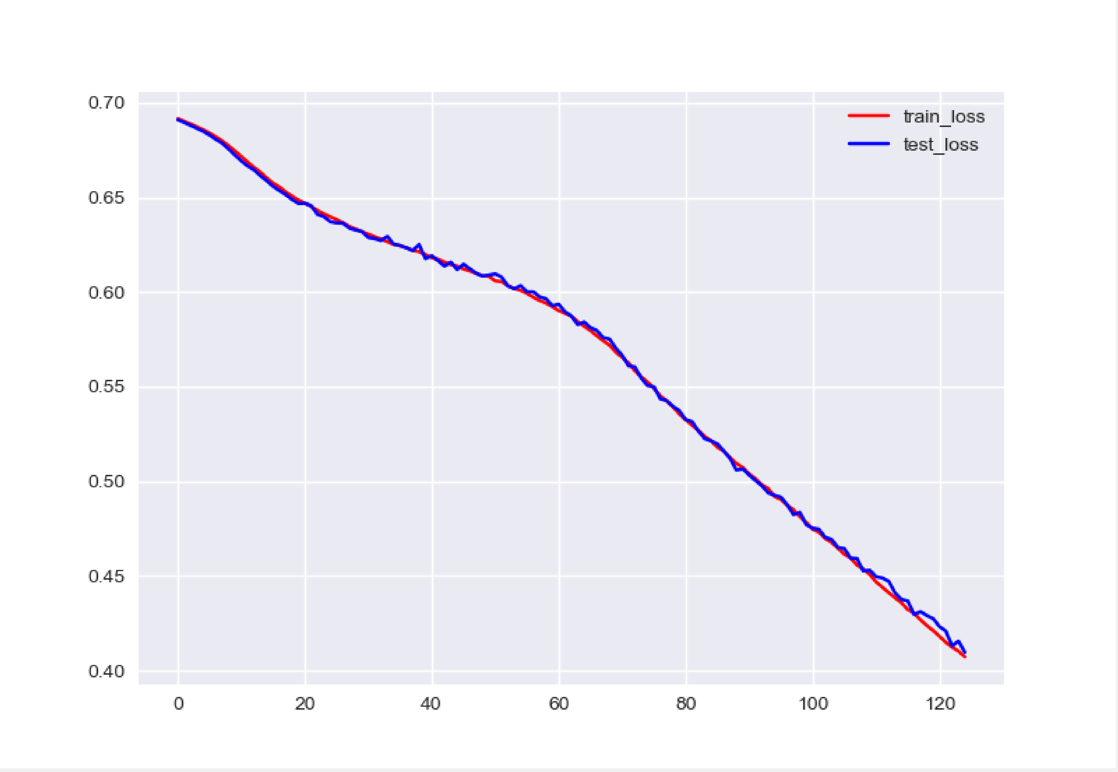


Figure: Training vs validation loss on baseline

Predictive Analysis

For this secondary task, we have addressed the problem of predicting if a seizure is supposed to start in the next 10 seconds. We trained with the best classifier technique for the previous

Model	Best on predictive
Small-CNN-Baseline on raw input	0.713
CNN on raw input	0.731
DeepCNN on MFCC	**0.8125**

Table: Experimental results predictive

References

[1] Adam A. Badea A. *EEG Seizure Analysis Dataset*. `https://www.kaggle.com/adibadea/chbmitseizuredataset`. 2021.

[2] A. Goldberger et al. "Components of a new research resource for complex physiologic signals. Circulation [Online].". In: *Journal title* 101.23 (Mar. 2000), e215–e220.

[3] Liangchen Luo et al. "Adaptive Gradient Methods with Dynamic Bound of Learning Rate". In: *Proceedings of the 7th International Conference on Learning Representations*. New Orleans, Louisiana, May 2019.

[4] Karen Simonyan and Andrew Zisserman. *Very Deep Convolutional Networks for Large-Scale Image Recognition*. 2015. arXiv: `1409.1556 [cs.CV]`.

sample