A GENERALIZABLE MODEL FOR Seizure Prediction BASED ON DEEP LEARNING USING CNN-LSTM ARCHITECTURE

Mohamad Shahbazi¹, Hamid K. Aghajan¹,²

¹Dept. of Electrical Engineering, Sharif University of Technology
²imec, Ghent University, Ghent, Belgium
Motivation

Epilepsy
- The fourth most common neurological disorder
- Affecting about 65 million people around the world
- Sudden seizures

Seizure Prediction
- Uncontrollable seizures in about $\frac{1}{3}$ of the patients
- The importance of seizure prediction systems
- EEG signals

Information obtained from [www.epilepsy.com/learn/about-epilepsy-basics](http://www.epilepsy.com/learn/about-epilepsy-basics)
Outline

- Introduction
- Literature Review
- Methods and Material
- Results and Discussion
- Conclusion
Outline

- Introduction
- Literature Review
- Methods and Material
- Results and Discussion
- Conclusion
Epilepsy Temporal States

- Preictal
- Interictal
- Ictal
- Post-ictal

The diagram illustrates the temporal states of epilepsy with arrows showing the transition between these states.
Challenges

- Complexity and variability preictal patterns
  - Different patients
  - Different seizures of the same patient

- Preictal Labels
Epileptic EEG Signals

https://www.epilepsydiagnosis.org/seizure/absence-typical-eeg.html
Deep Neural Networks

- Proved to be powerful in many areas
- Convolutional Neural Networks (CNN)
  - Extracting the best features from the best channels using trainable filters
- Recurrent Neural Networks (RNN)
  - Sequences
  - Long Short-term Memory (LSTM)
Outline

- Introduction
- Literature Review
- Methods and Material
- Results and Discussion
- Conclusion
Literature Review

- **Studies based on hand-crafted Features [1]**
  - Complex and time consuming feature extraction and selection

- **Studies based on CNNs [2],[3],[4]**
  - 2D images constructed from EEG segments as input
  - Mediocre performance

- **Hand-crafted Features + RNN [5]**
  - Suffers from the problems of hand-crafted feature extraction
The Whole Picture

- Preprocessing
- CNN-LSTM Network
- Post-processing
CHB–MIT Dataset*
- Non-invasive continuous EEG recordings
- 22 patients, 23 cases
- 23 channels (most cases)
- Annotation contains the start and the end of each seizure

Extra annotation for this work
- Preictal: up to 30 minutes before each seizure onset
- Interictal: recordings at least 2 hours away from seizures and their annotated preictal state

* Dataset is available at [https://www.physionet.org/pn6/chbmit/](https://www.physionet.org/pn6/chbmit/)
Preprocessing

- Split EEG recordings into sequence of segments
  - Sequences of six 10-second overlapping segments

- Short-Time Fourier Transform
  - 1-second sliding window with 75% overlap
  - Removing DC frequency and frequencies related to power line noise
  - Standardizing each frequency along the time axis
Preprocessing

An example of a standardized STFT image extracted from a 10-second EEG segment
Proposed CNN-LSTM Architecture
CNN Architecture
Network Training

- Patient-specific training
- Pre-training of the CNN weights
- Train and test sets
  - Preictal data
    - Leave one seizure out
  - Interictal Data
    - 40% of non-seizure files as test set
Post-Processing

Interictal Interictal Preictal Interictal Preictal Preictal Preictal

Seizure Prediction Alarm

8 predictions out of 10 predictions
**Evaluation**

- **Sensitivity** = \( \frac{TP}{\# \text{seizures}} \times 100 \)
- **FPR** = \( \frac{FP}{\# \text{hours}} \times 100 \)
- **Seizure prediction Horizon (SPH)**
  - 30 minutes

![Graph showing True Prediction, False Prediction, and Seizure Onset with SPH highlighted between them.](image-url)
Outline

- Introduction
- Literature Review
- Methods and Material
- Results and Discussion
- Conclusion
### Results

**Sensitivity:** 98.21%

**FPR:** 0.13 /h

**Prediction Time:** 44.74 min.

<table>
<thead>
<tr>
<th>Case</th>
<th>No. seizures</th>
<th>Sen. (%)</th>
<th>FPR (/h)</th>
<th>Pred. Time (Minutes)</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>5</td>
<td>100.00</td>
<td>0.08</td>
<td>29.50</td>
</tr>
<tr>
<td>02</td>
<td>2</td>
<td>100.00</td>
<td>0.06</td>
<td>50.00</td>
</tr>
<tr>
<td>03</td>
<td>4</td>
<td>75.00</td>
<td>0.00</td>
<td>32.00</td>
</tr>
<tr>
<td>05</td>
<td>3</td>
<td>100.00</td>
<td>0.00</td>
<td>35.00</td>
</tr>
<tr>
<td>07</td>
<td>3</td>
<td>100.00</td>
<td>0.16</td>
<td>49.00</td>
</tr>
<tr>
<td>09</td>
<td>3</td>
<td>100.00</td>
<td>0.00</td>
<td>103.00</td>
</tr>
<tr>
<td>10</td>
<td>7</td>
<td>100.00</td>
<td>0.50</td>
<td>32.00</td>
</tr>
<tr>
<td>17</td>
<td>3</td>
<td>100.00</td>
<td>0.22</td>
<td>43.00</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>100.00</td>
<td>0.13</td>
<td>37.00</td>
</tr>
<tr>
<td>19</td>
<td>2</td>
<td>100.00</td>
<td>0.00</td>
<td>46.00</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>100.00</td>
<td>0.00</td>
<td>26.00</td>
</tr>
<tr>
<td>21</td>
<td>4</td>
<td>100.00</td>
<td>0.50</td>
<td>51.92</td>
</tr>
<tr>
<td>22</td>
<td>3</td>
<td>100.00</td>
<td>0.18</td>
<td>40.00</td>
</tr>
<tr>
<td>23</td>
<td>2</td>
<td>100.00</td>
<td>0.18</td>
<td>52.00</td>
</tr>
</tbody>
</table>
## Comparison with Related Works

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors</th>
<th>Dataset</th>
<th>Method</th>
<th>Sen. (%)</th>
<th>FPR (/h)</th>
<th>Pred. Time (min)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009</td>
<td>Mirowski et al [13]</td>
<td>Freiburg 15 cases</td>
<td>Bivariate features + CNN</td>
<td>71</td>
<td>0</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>Truong et al [14]</td>
<td>CHB-MIT 13 cases</td>
<td>STFT + CNN</td>
<td>81.2</td>
<td>0.16</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>Khan et al [15]</td>
<td>CHB-MIT 13 cases</td>
<td>Wavelet + CNN</td>
<td>83.3</td>
<td>0.14</td>
<td>5.81</td>
</tr>
<tr>
<td>2018</td>
<td>Tsiouris et al [16]</td>
<td>CHB-MIT 24 cases</td>
<td>Hand-crafted features + LSTM</td>
<td>99.8</td>
<td>0.02</td>
<td>-</td>
</tr>
<tr>
<td>2018</td>
<td>This work</td>
<td>CHB-MIT 14 cases</td>
<td>STFT + CNN-LSTM</td>
<td>98.2</td>
<td>0.13</td>
<td>44.74</td>
</tr>
</tbody>
</table>
Outline

- Introduction
- Literature Review
- Methods and Material
- Results and Discussion
- Conclusion
Conclusion

◉ A novel method based on CNN-LSTM architecture
  ○ Outperforming studies based on CNN
  ○ Learning time-frequency features without human supervision

◉ Future Work
  ○ Optimal Preictal length for each patient
  ○ Unsupervised methods based on temporal clustering
References


Thank You!

Any questions?

M.shahbazi72@gmail.com