Automatic Detection of Epileptic Seizure Onset and Offset in Scalp EEG

Presented by

Poomipat Boonyakitanont

Advised by

Assist. Prof. Jitkomut Songsiri, Ph.D.[1]

Assist. Prof. Apiwat Lek-uthai, Ph.D.[1]

Assist. Prof. Krisnachai Chomtho, M.D.[2]

[1]Department of Electrical Engineering, Faculty of Engineering Chulalongkorn University

[2]Department of Pediatrics, Faculty of Medicine, Chulalongkorn University

poomipatosk129@gmail.com

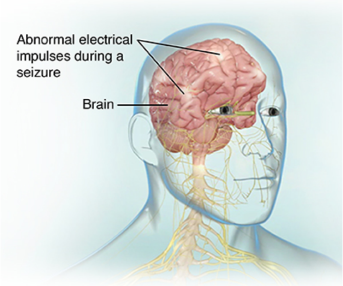
Aug 30, 2019

Outline

Epilepsy

- Neurological disorder due to an excessive amount of electrical discharges in the brain.
- Damage on the human brain and uncontrollable physical activities.

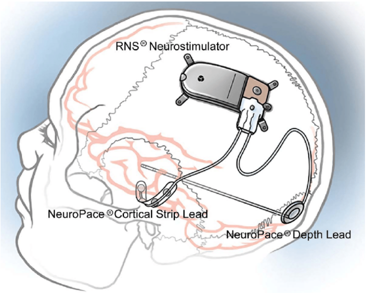
https://blogs.allizhealth.com/dealing- epileptic-seizure/



https://www.saintlukeshealthsystem.org/health-library/electroencephalogram-eeg

Epileptic Seizure Onset and Offset

- Responsive neurostimulator requiring almost simultaneous seizure alarm
- Length of seizure activity needed for proper treatment



https://www.researchgate.net/figure/Responsive-neurostimulation-device_fig1_309278949

https://pxhere.com/en/photo/566564

Objective and Scopes

Objective

- This work aims to provide **an offline detection** method of seizure activities and the indication of **the onsets and offsets** in multi-channel scalp EEG signals.

Scopes of work

- The proposed system needs to be trained before it will be employed in a specific subject.
- Data used in this work is multi-channel scalp EEG signals collected from an **online accessible database**. Seizure onset and offset points are given if seizure events appear in EEG signals, and training and testing data are required to have the **same montage**.
- Stages of training and testing are operated **offline**.
- We determine seizure onset and offset but do not specify types of seizure.

Benefits and Outcomes

Benefits

- Less efforts from neurologists are needed.
- Only one brain modality which is EEG is required.
- A pre-annotated data including onset and offset points can reduce time spent by neurologists on reviewing long EEG signals.

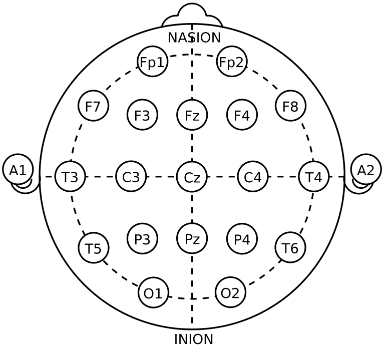
Outcomes

- We provide a scheme of automatic detection of epileptic seizures and the onsets and offsets.
- The algorithm of automatic detection of epileptic seizures and the onsets and offsets is given.

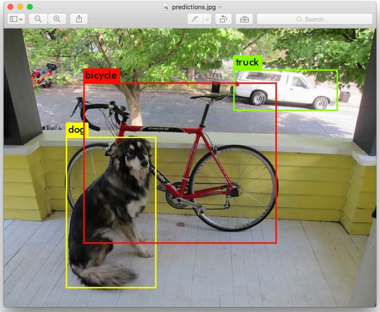
Background

Two basic backgrounds

- EEG and montages
- Convolutional neural networks (CNNs)



https://commons.wikimedia.org



https://pjreddie.com/darknet/yolo/

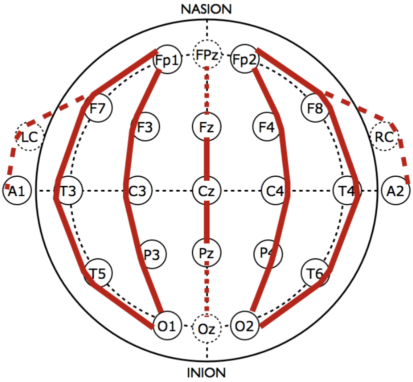
EEG and montages

EEG

- A way to observe electrical activities in the brain
- Fast temporary response but low spatial resolution

Montages

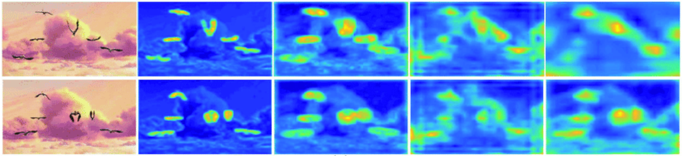
- Referential and bipolar montages

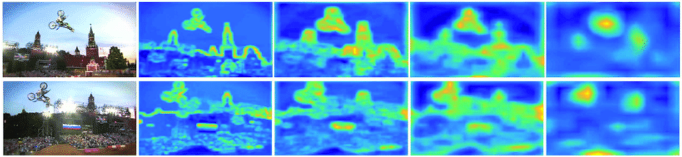


https://eegatlas-online.com/index.php/en/montages/bipolar/double-banana

Convolutional Neural Networks

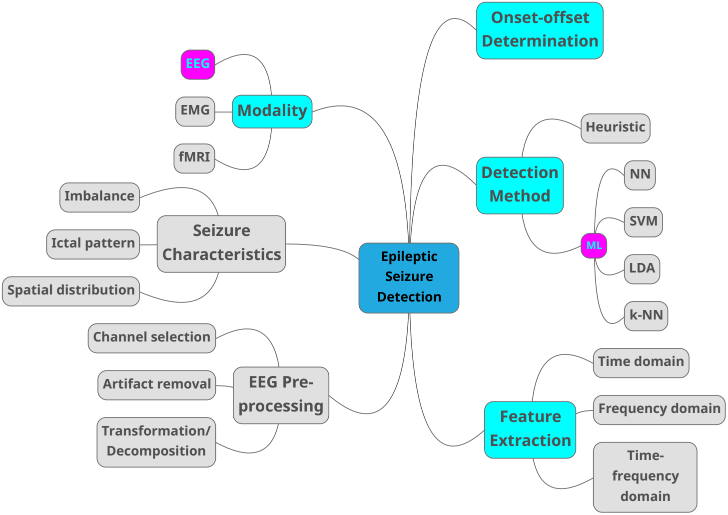
- Self-extract features in feature maps.
- Gain popularity from image detection and speech recognition [LBH15], *e.g.*, VGG16 [SZ15], InceptionNet [SVI+16], ResNet [HZRS16] in image classification.


(a)


(b)

Source: [LXD+18]

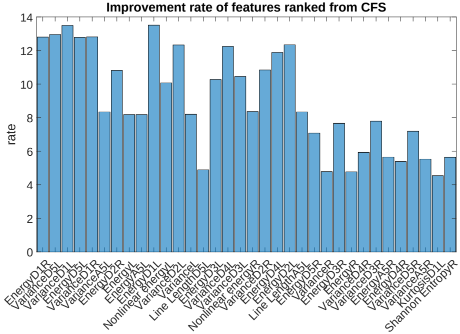
Literature Review



Feature Extraction

From our results in [BLuCS19]

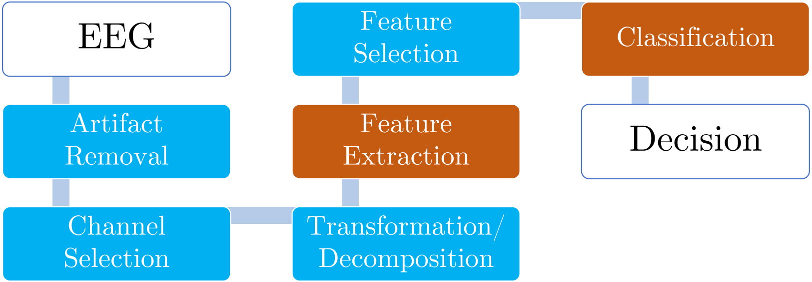
- **Statistical parameters, energy, entropies** were commonly selected features in time, frequency, and time-frequency domains.
- Features capturing changes in amplitude (energy) and uncertainties (entropies) were favorable and sometimes used individually, while statistical parameters were always applied jointly.



Epileptic Seizure Detection

From our review in [BLuCS19]

- Many studies considered epochs individually [ASS$^+$13, SG10, Jan17], while some technique was based on event [SLuC15].
- Machine learning techniques including ANN and SVM were favorable.
- Almost all of them did not use all data records.



Epileptic Seizure Onset and Offset

There have been only three studies using online accessible database.

- Shoeb et al. [SKS+11] applied signal **energies** and **SVM** to detect seizure termination. Only one offset was missed, and an average absolute latency of 10.30 ± 5.50 seconds.
- Orosco et al. [OCDL16] used **relative energy** computed from **stationary wavelet transform** coefficients with **LDA** and **ANN**. The method using **LDA** obtained 92.60% GDR and 0.3 FPR/h. Ranges of latencies were wide, 42.40 seconds for onset and 84.40 seconds for offset.
- Chandel et al. [CUFK19] employed **statistics-based features** extracted from **orthonormal triadic wavelet transform** coefficients. On average, using **LDA** obtained 100% GDR, 4.02 FPR/h. Ranges of onset and offset latencies were 19.67 and 58.67 seconds, respectively.

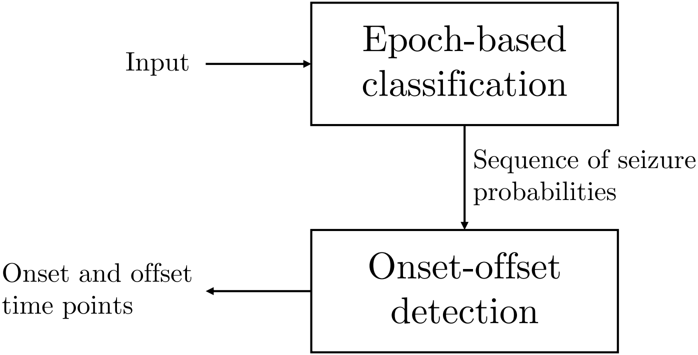
Missing Gap

There are only a few studies in this topic.
Existing approaches still suffer from having

- Large latency ranges
- High FPR/h
- Heuristic approach in onset-offset detection

Problem Statement

- Classifier detecting seizures based on EEG epochs
- Onset-offset detector identifying a seizure occurrence in long EEG from outputs of the classifier



Classifier

Consider $\mathcal{D} = \mathcal{X} \times \mathcal{Y}$ be a set of samples (x_i, y_i) drawn from a joint probability distribution $f_{xy}(x, y)$ where $\mathcal{Y} = \{0, 1\}$. A classifier h, also called a hypothesis, in a hypothesis space $\mathcal{H}$ is a function that tries to map x_i to its label y_i. To evaluate the classifier, the *true risk*, or the expectation of a loss function L, is theoretically used:

$$R_{\mathrm{true}}(h) = \mathbf{E}[L(h(x), y)] = \int_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} f_{xy}(x, y) L(h(x), y). \tag{1}$$

In practice, the classifier is evaluated via *empirical risk*:

$$R_{\mathrm{emp}}(h) = \sum_{(x,y) \in \mathcal{D}} P(x, y) L(h(x), y) = \frac{1}{|\mathcal{D}|} \sum_{(x,y) \in \mathcal{D}} L(h(x), y), \tag{2}$$

where $P(x, y)$ is the hypothetical joint probability.

Classifier

The goal is to find the optimal hypothesis $h^* \in \mathcal{H}$ such that

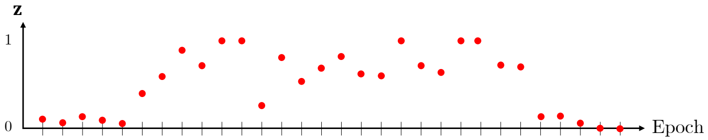
$$h^* = \operatorname*{argmin}_{h \in \mathcal{H}} R_{\mathrm{emp}}(h). \tag{3}$$

Example of $\mathcal{H}$

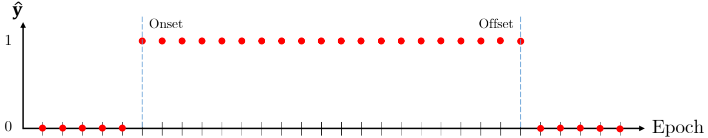
- Linear classifier
- Polynomial classifier
- Neural network
- Support vector machine

Onset-offset detector

The onset-offset detector is a function g transforming output vector $\boldsymbol{z}$ to vector of predicted class $\hat{\boldsymbol{y}}$.



(a) Output from the epoch-based seizure detection.



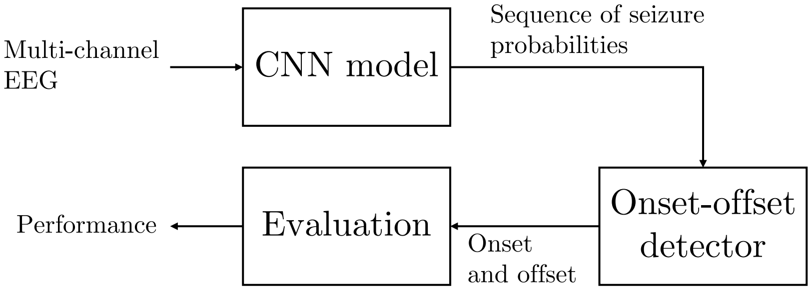
(b) Output of the onset-offset detection.

Research Methodology

- Review literature on data collection, pre-processing, feature extraction, classification and onset-offset detection.
- Propose epoch-based seizure detection method and technique to determine the onset and offset.
- Collect data from online database.
- Train a classifier on EEG epochs and verify the trained model on an unseen record from the same subject.
- Determine onset and offset from epoch-based classification results and compare the results of the onset-offset detection and of the classification.
- Conclude the detection performance, limitations, and future work.

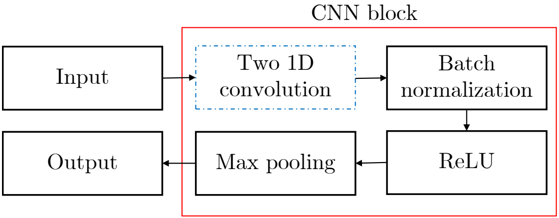
Proposed Method

- Classification
- Onset-offset detection
- Evaluation

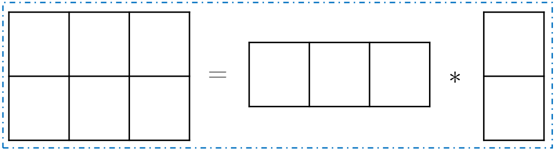


Classification

- Design blocks of layers in CNN stacked deeply for feature extraction.

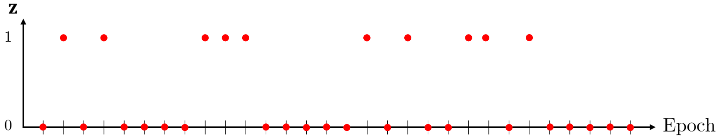


- Use filter factorization technique to reduce a number of parameter [SVI+16].

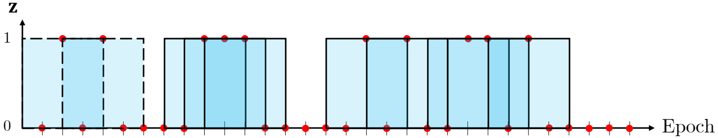


Onset-offset detection

- Cover all detected epochs by a window of size $2l+1$ centered at the epoch.



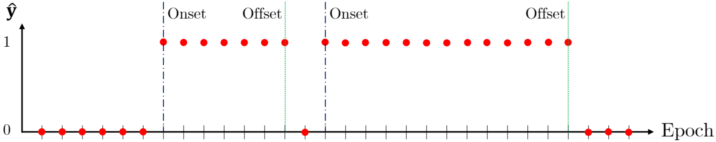
(a) Output of the CNN model.



(b) Rectangular windows covering seizure epochs.

Onset-offset detection

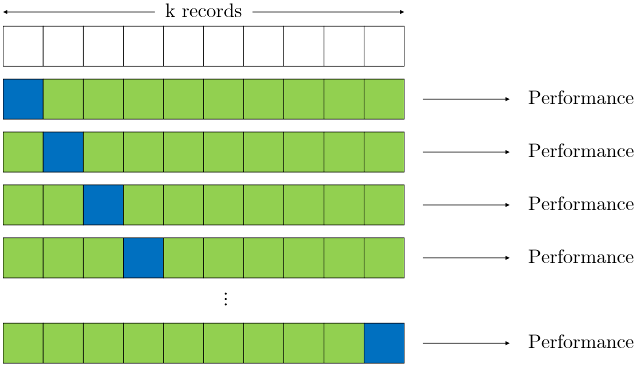
- Group p adjacent touching or overlapping windows into one and neglect the others.
- Onset and offset are the beginning and ending epochs of the event.



(c) Results of the onset-offset detector. The onset and offset are the first and last epoch of the predicted event.

Evaluation

- Evaluate classification performance of each patient independently
- Apply leave-one-record-out for validation
- Green records for training, and the blue one for testing



Performance Metrics

Epoch-based

- Sensitivity (Sen) $= (TP)/(TP + FN) \times 100\%$
- $\mathrm{F}_1 = (2TP)/(2TP + FN + FP) \times 100\%$

Event-based

- False positive rate per hour (FPR/h) $=$ FP per hour
- Good detection rate (GDR) $=$ percent of detected event

Time-based

- Onset and offset latencies
- Absolute onset and offset latencies

Experiment

Goal

- Compare improvement of using onset-offset detector from classifier.
- Compare performances with other studies.

Data

- CHB-MIT Scalp EEG database [GAG^{+}00]

Expected results

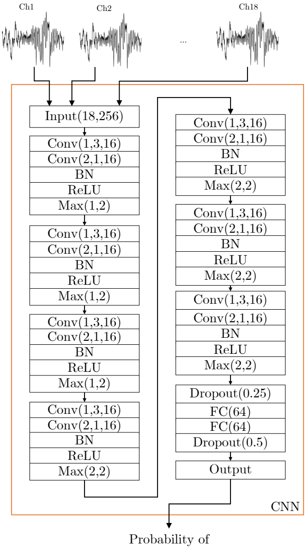
- Improve Sen, F_1
- Reduce FPR/h
- Maintain GDR

Data Description

- All records from CHB-MIT Scalp EEG database, containing 24 cases (`chb01-chb24`), were used. See more details in Appendix.
- Chosen channels were *FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-T3, T3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, T8-P8, P8-O2, FZ-CZ*, and *CZ-PZ*.
- The model input was one-second multi-channel EEG epoch represented as 18×256 dimensional matrix.
- This database is publicly and freely downloaded from PhysioNet (`https://physionet.org/physiobank/database/chbmit/`).

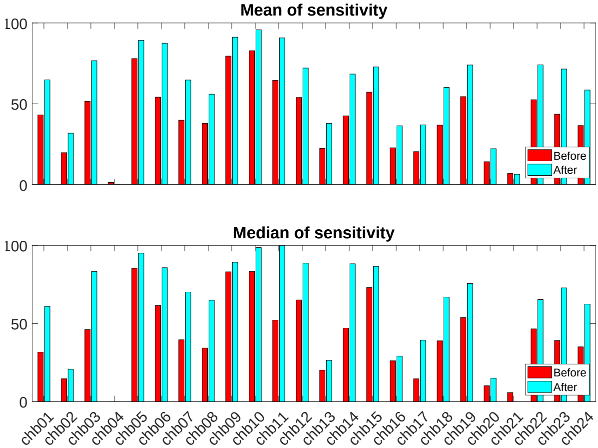
Model Description

- Hyperparameters were tuned by reducing the model size while the model still gave high GDR on the case chb24.
- Activation function: *ReLU*.
- Optimizer: *ADADELTA* [Zei12]
- Loss function: *binary cross entropy*.
- The mini-batch size: 100 epochs
- Stopping criteria: 100 iterations.
- An epoch was denoted as seizure when seizure probability was higher than 0.5.
- We selected $l = 2$ and $p = 3$ to cover the smallest seizure.



Preliminary Results

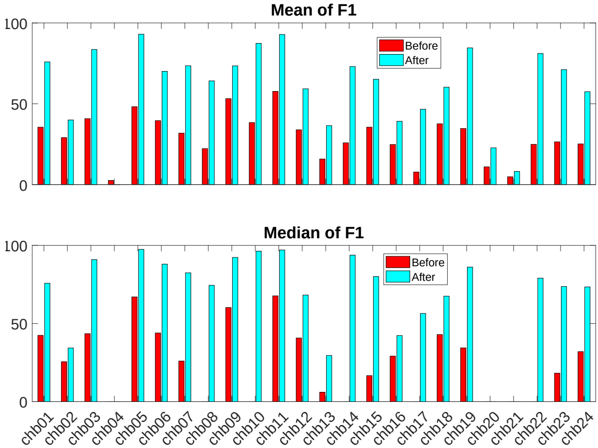
Before/After = before/after onset-offset detection.



- Sen significantly increased

Preliminary Results

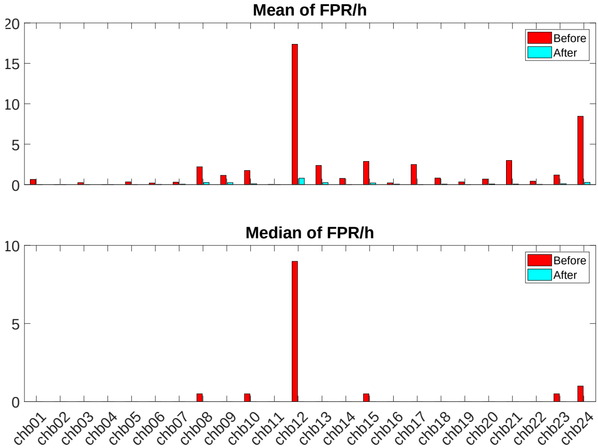
Before/After = before/after onset-offset detection.



- Like Sen, F_1 remarkably improved

Preliminary Results

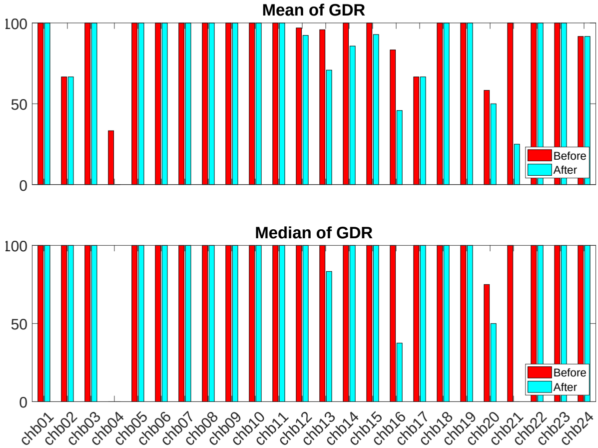
Before/After = before/after onset-offset detection.



- FPR/h after onset-offset detection considerably reduced

Preliminary Results

Before/After = before/after onset-offset detection.



- GDR in some cases decreased

Preliminary Results

		Our work	[OCDL16]	[CUFK19]
Onset latency	Min	-2	-28.00	-10.00
	Max	20	14.40	9.67
Offset latency	Min	-62.8	-24.20	-6.00
	Max	-0.20	60.20	52.67
Absolute onset latency	Mean	6.12	-	-
	Min	0.43	-	-
	Max	20	-	-
Absolute offset latency	Mean	13.95	-	-
	Min	0.33	-	-
	Max	62.8	-	-

- Even though cancellation of latency in other work occurs, the ranges of our work are comparable to that of others.
- Our results are more reasonable and applicable in practice.

Conclusion

- In this work, we used all records from the CHB-MIT Scalp EEG database to develop patient-specific seizure detection method.
- We designed the method based on deep learning techniques to detect seizures based on small epoch and heuristic approach to determine onset and offset.
- Onset-offset detector potentially reduced FPR/h and significantly Sen and F_1.
- Ranges of latencies from our method were comparable with others even though others personally selected data set.

Limitation and Future Work

Limitations

- When CNN cannot perform well enough, the onset-offset detector also cannot improve the performance significantly.
- The onset-offset detector is heuristic, not adaptive.

Future work

- Explore or design another epoch-based classifier that
  - Deals with imbalance data.
- Propose an onset-offset detection based on machine learning that
  - Takes imbalance properties into account.
- Analyze performance of onset-offset detection.

References I

[ASS+13] U.R. Acharya, S.V. Sree, G. Swapna, R.J. Martis, and J.S. Suri.
Automated EEG analysis of epilepsy: A review.
Knowledge-Based Systems, 45:147–165, 2013.

[BLuCS19] P. Boonyakitanont, A. Lek-uthai, K. Chomtho, and J. Songsiri.
A review of feature extraction and performance evaluation in epileptic seizure detection using EEG.
arXiv preprint arXiv:1908.00492, 2019.

[CUFK19] G. Chandel, P. Upadhyaya, O. Farooq, and Y.U. Khan.
Detection of seizure event and its onset/offset using orthonormal triadic wavelet based features.
IRBM, 40(2):103–112, 2019.

[GAG+00] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C. Peng, and H.E. Stanley.
PhysioBank, PhysioToolkit, and PhysioNet.
Circulation, 101(23):e215–e220, 2000.

[HZRS16] K. He, X. Zhang, S. Ren, and J. Sun.
Deep residual learning for image recognition.
In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

[IS15] S. Ioffe and C. Szegedy.
Batch normalization: Accelerating deep network training by reducing internal covariate shift.
In *Proceedings of the 32nd International Conference on Machine Learning*, pages 448–456, 2015.

[Jan17] S. Janjarasjitt.
Epileptic seizure classifications of single-channel scalp EEG data using wavelet-based features and svm.
Medical & Biological Engineering & Computing, 55(10):1743–1761, 2017.

References II

[LBH15] Y. LeCun, Y. Bengio, and G. Hinton.
Deep learning.
Nature, 521(7553):436–444, 2015.

[LXD+18] Y. Li, T. Xu, H. Deng, G. Shi, and J. Guo.
Adaptive correlation model for visual tracking using keypoints matching and deep convolutional feature.
Sensors, 18(2):653, 2018.

[OCDL16] L. Orosco, A.G. Correa, P. Diez, and E. Laciar.
Patient non-specific algorithm for seizures detection in scalp EEG.
Computers in Biology and Medicine, 71:128–134, 2016.

[SG10] A.H. Shoeb and J. Guttag.
Application of machine learning to epileptic seizure detection.
In *Proceedings of the 27th International Conference on Machine Learning*, pages 975–982, 2010.

[SHK+14] N. Srivastava, G. E. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov.
Dropout: a simple way to prevent neural networks from overfitting.
Journal of Machine Learning Research, 15(1):1929–1958, 2014.

[SKS+11] A.H. Shoeb, A. Kharbouch, J. Soegaard, S. Schachter, and J. Guttag.
An algorithm for detecting seizure termination in scalp EEG.
In *Proceedings of 2011 the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 1443–1446. IEEE, 2011.

[SLuC15] C. Satirasethawong, A. Lek-uthai, and K. Chomtho.
Amplitude-integrated EEG processing and its performance for automatic seizure detection.
In *Proceedings of the 2015 IEEE International Conference on Signal and Image Processing Applications*, pages 551–556. IEEE, 2015.

References III

[SVI+16] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna.
Rethinking the inception architecture for computer vision.
In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pages 2818–2826, 2016.

[SZ15] K. Simonyan and A. Zisserman.
Very deep convolutional networks for large-scale image recognition.
In *International Conference on Learning Representations*, 2015.

[Zei12] M. D. Zeiler.
ADADELTA: An adaptive learning rate method.
arXiv preprint arXiv:1212.5701, 2012.

.

Appendix

- **Convolutional Neural Networks**
- **CHB-MIT Scalp EEG database**

Convolutional Neural Networks

- Inspired by the mammal visual system
- Gaining popularity from image detection and speech recognition [LBH15], *e.g.*, VGG16 [SZ15], InceptionNet [SVI+16], ResNet [HZRS16]
- Mainly consisting of convolutional, activation, pooling, and fully-connected layers
- Optionally adding batch-normalization [IS15] and dropout [SHK+14] layers

Convolutional Layer

- Extract features by convolution operation of input and filter
- Example: vertical edge detection
- Train the filter weights by optimization

4	9	2	5	8
5	6	2	4	0
2	4	5	4	5
5	6	5	4	7
5	7	7	9	2

Input

$*$

1	0	-1
1	0	-1
1	0	-1

Filter

$=$

2	6	-4
0	4	0
-5	0	3

Output

Activation Layer

- Transform input to output with nonlinear function

2	6	-4
0	4	0
-5	0	3

σ

0.8808	0.9975	0.0180
0.5000	0.9820	0.5000
0.0067	0.5000	0.9526

ReLU

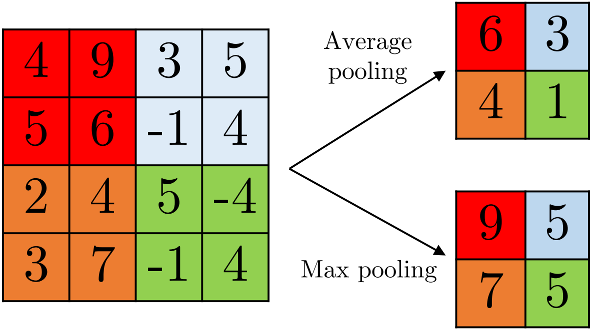
2	6	0
0	4	0
0	0	3

tanh

0.9640	1.0000	-0.9993
0.0000	0.9993	0.0000
-0.9999	0.0000	0.9951

Pooling Layer

- Down-sampling input by appropriate approach: average and max



Batch Normalization Layer

- Normalizing each feature map at each mini-batch [IS15]
- Consider a mini-batch $\mathcal{B} = \{x_1, x_2, \ldots, x_k\}$

Input : x over a mini-batch $\mathcal{B}$
Parameter : γ, β
Output : $\{y_i = \gamma x_i + \beta\}$

1 $\mu_{\mathcal{B}} \leftarrow \frac{1}{k} \sum_{i=1}^{k} x_i$ `// mean of mini-batch`

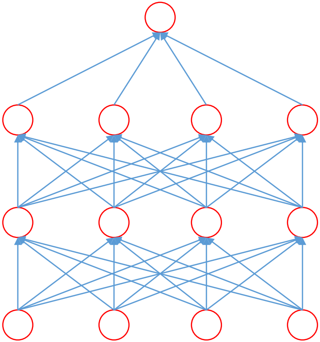
2 $\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{k} \sum_{i=1}^{k} (x_i - \mu_{\mathcal{B}})^2$ `// variance of mini-batch`

3 $\hat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}}$ `// normalization`

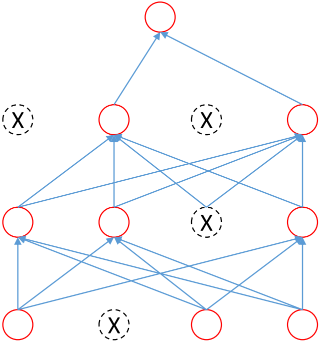
4 $y_i \leftarrow \gamma x_i + \beta$ `// scale and shift`

Dropout Layer

- Randomly and temporarily deleting neurons from the layer [SHK$^+$14]



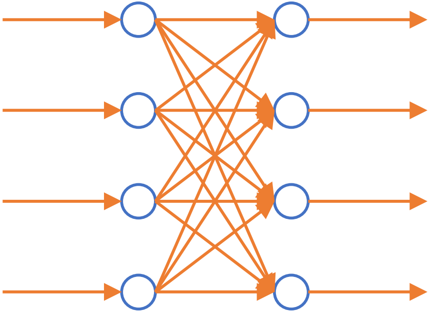
(a) Standard neural network.



(b) After employing dropout.

Fully-Connected Layer

- Every neuron is connected to all neurons in the adjacent layers.



CHB-MIT Scalp EEG database

Cases	Number of records	Total duration (sec)	Number of seizures	Total seizure duration (sec)
chb01	42	145,988	7	449
chb02	36	126,959	3	175
chb03	38	136,806	7	409
chb04	42	561,834	4	382
chb05	39	140,410	5	563
chb06	18	240,246	10	163
chb07	19	241,388	3	328
chb08	20	72,023	5	924
chb09	19	244,338	4	280
chb10	25	180,084	7	454
chb11	35	123,257	3	809
chb12	24	85,300	40	1,515
chb13	33	118,800	12	547
chb14	26	93,600	8	177
chb15	40	144,036	20	2,012
chb16	19	68,400	10	94
chb17	21	75,624	3	296
chb18	36	128,285	6	323
chb19	30	107,746	3	239
chb20	29	99,366	8	302
chb21	33	118,189	4	203
chb22	31	111,611	3	207
chb23	9	95,610	7	431
chb24	22	76,640	16	527
sum	**686**	**3,536,540**	**198**	**11,809**