CROSS-SITE GENERALIZATION FOR IMBALANCED EPILEPTIC CLASSIFICATION

ICASSP 2023

04.06.2023

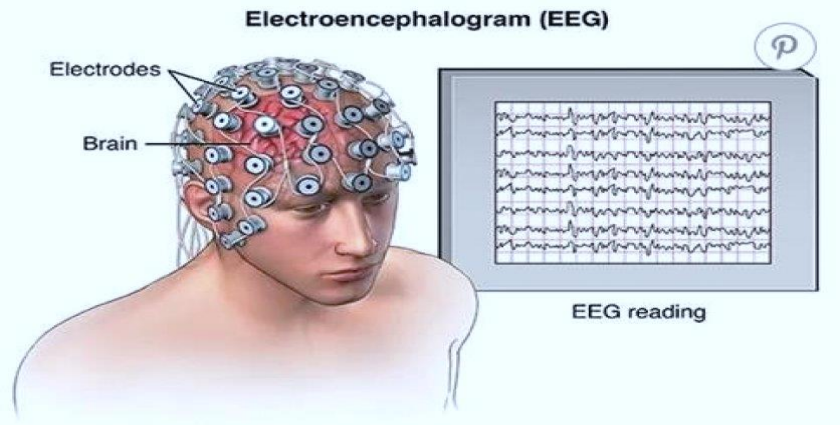
Tala Abdallah – LARIS – Angers – France
Nisrine Jrad – UCO – Angers – France
Fahed Abdallah – Lebanese University – Beirut – Lebanon
Anne Humeau-Heurtier – LARIS – Angers – France
Patrick Van Bogaert – CHU – Angers – France

Context:

Epilepsy

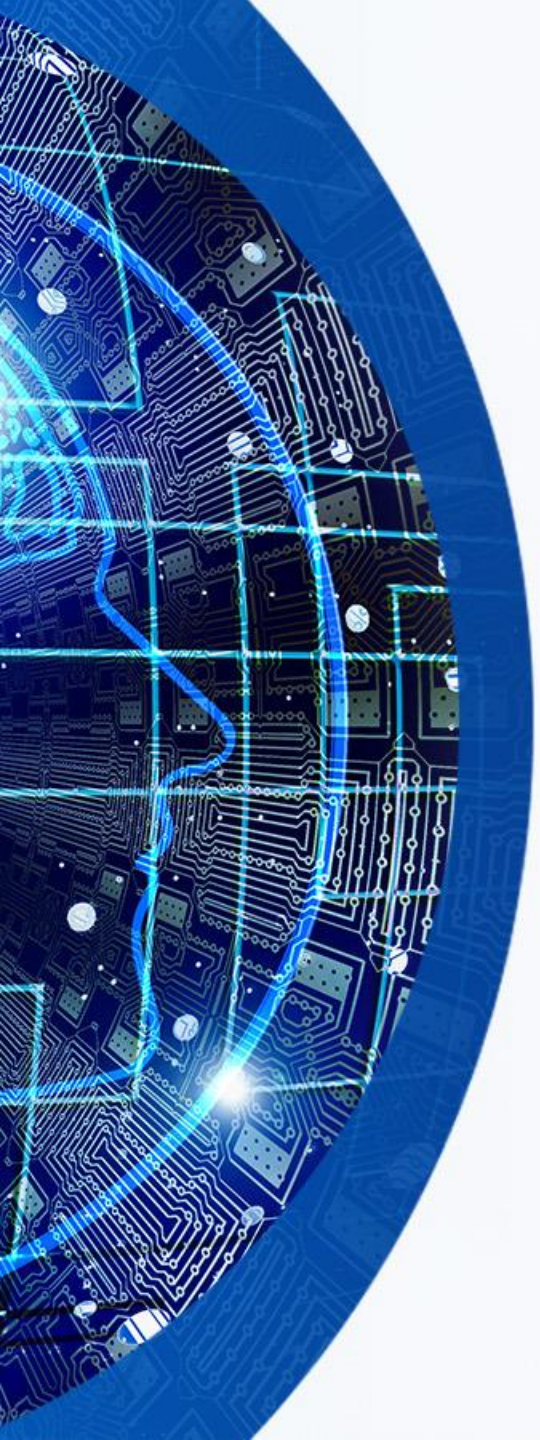
- Disorder of normal brain function
- Impacts 2% of the world's population
- It is defined by the recurrence of epileptic seizures (can lead to a sudden death!!)

EEG



→ Consuming time and effort

Solution → Automatic seizure detection



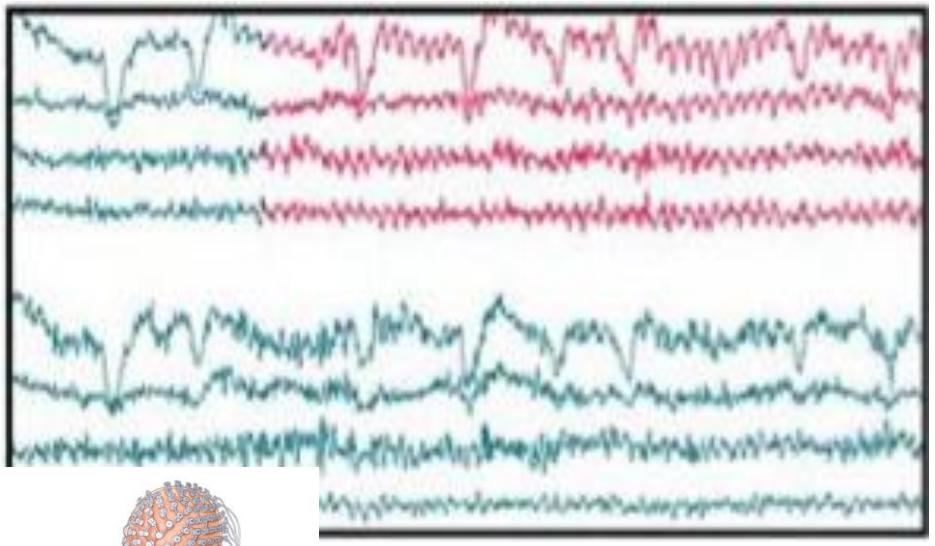
Challenges:

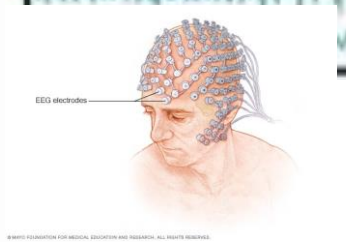
1)Non-stationary, noisy signals

2)Large amount of data

3) EEG signals are imbalanced

4)High inter-subject variability



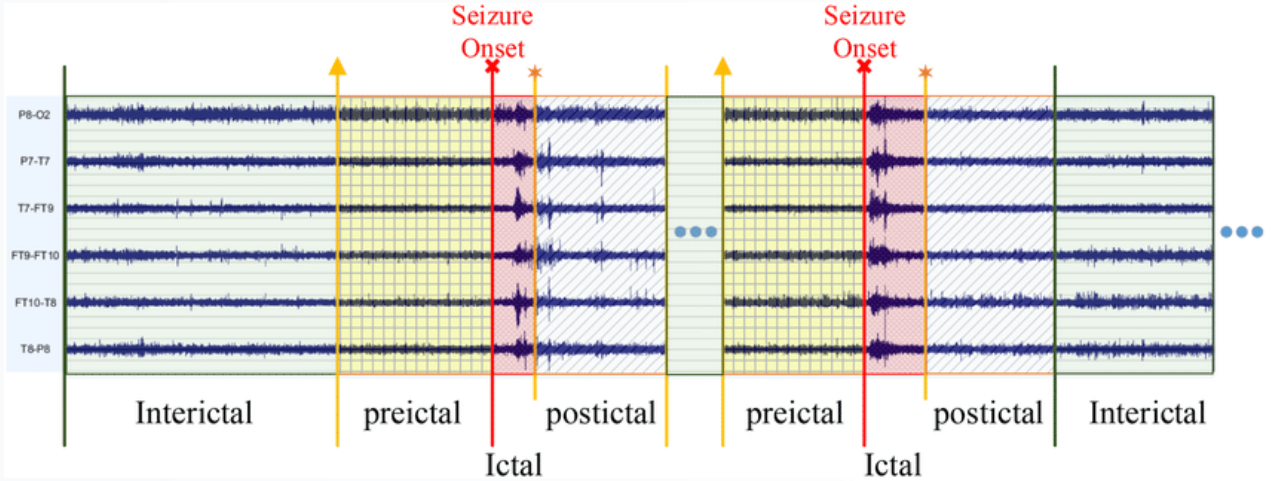


Originality and main contribution of the paper:

1. Applying data augmentation (DA) techniques on the seizure class followed by the widely used CNN-LSTM classifier.

2. Validation of the cross-site generalization ability of our DA-CNN-LSTM method.

3. The use of a local data set constitutes a challenge and an important originality of the study which can help other researchers to select widely applicable methods.



Children Hospital Boston data set (CHB-MIT):

- 24 cases (9 cases were selected for this study)
- Between 9 and 42 continuous .edf files for each case
- One hour of digitized EEG signals in each EDF file
- 18 common electrodes , bipolar montage
- Sampling Frequency of 256 HZ
- 664 EDF files (Records) in total
- 535 files with no seizures
- 129 files with one or more seizures
- 182 seizures

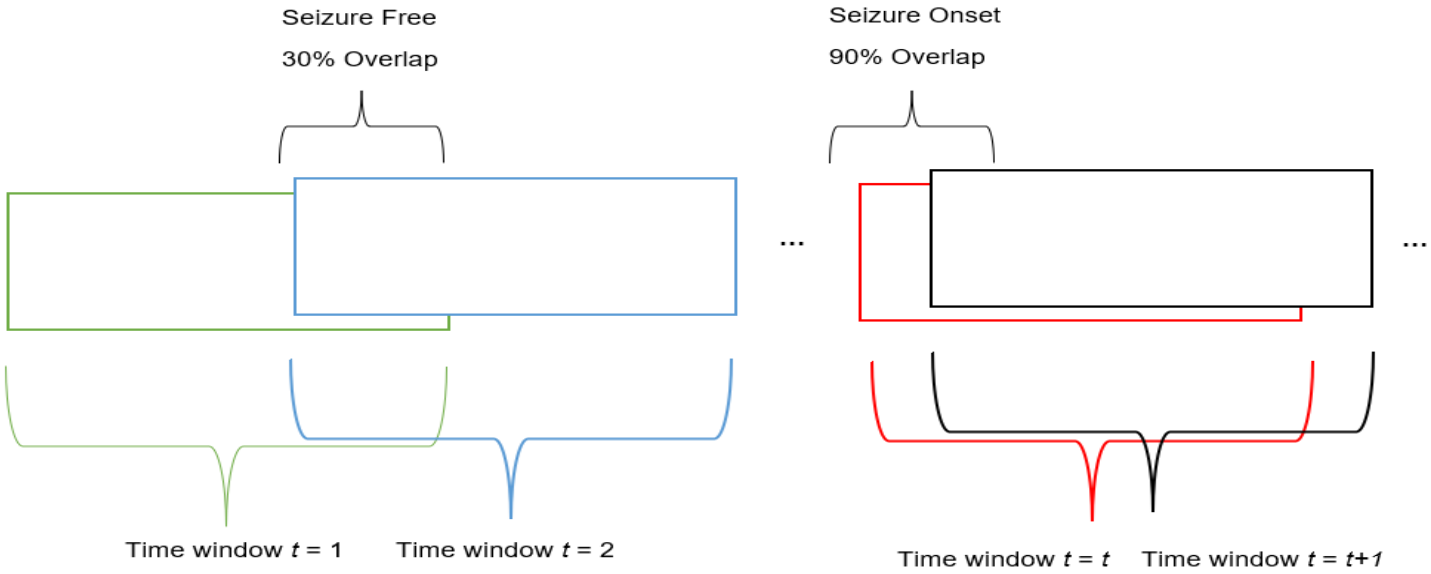
Subject	Gender	Age	# of seizure	Duration (hh: mm: ss)
chb01	F	11	7	40:33:08
chb02	M	11	3	35:15:59
chb03	F	14	7	38:00:06
chb04	M	22	4	156:03:54
chb05	F	7	5	39:00:10
chb06	F	1.5	10	66:44:06
chb07	F	14.5	3	67:03:08
chb08	M	3.5	5	20:00:23
chb09	F	10	4	67:52:18
chb10	M	3	7	50:01:24
chb11	F	12	3	34:47:37
chb12	F	2	27	20:41:40
chb13	F	3	12	33:00:00
chb14	F	9	8	26:00:00
chb15	M	16	20	40:00:36
chb16	F	7	10	19:00:00
chb17	F	12	3	21:00:24
chb18	F	18	6	35:38:05
chb19	F	19	3	29:55:46
chb20	F	6	8	27:36:06
chb21	F	13	4	32:49:49
chb22	F	9	3	31:00:11
chb23	F	6	7	26:33:30
chb24	-	-	16	21:17:47
Total			**185**	**979:56:07**

CHU of Angers:

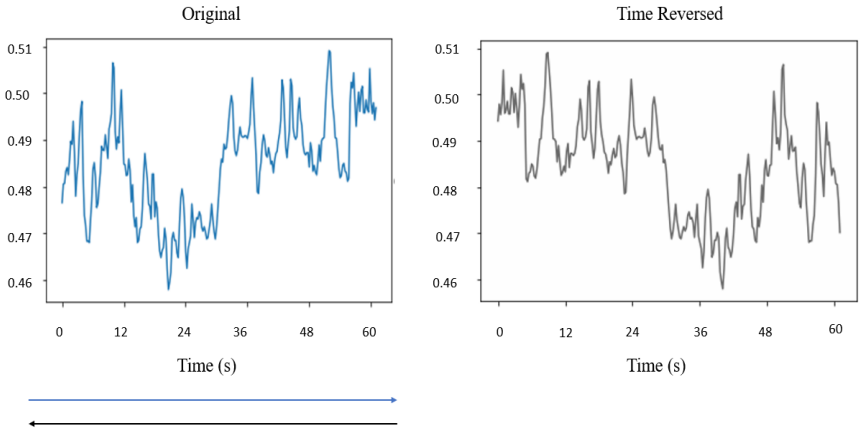
- Ethics approval by the agreement of the CHU Angers committee
- 20 cases aged between 1 and 17
- Between 20 and 45 continuous .edf files for each case
- 48 hours of digitized EEG signals in most of EDF files
- One to seven seizures for each patient
- Between 13 and 19 electrodes , monopolar montage
- Sampling Frequency of 256 HZ
- Around 960 hours of EEG recording

Data Augmentation (DA) Techniques for imbalanced data set

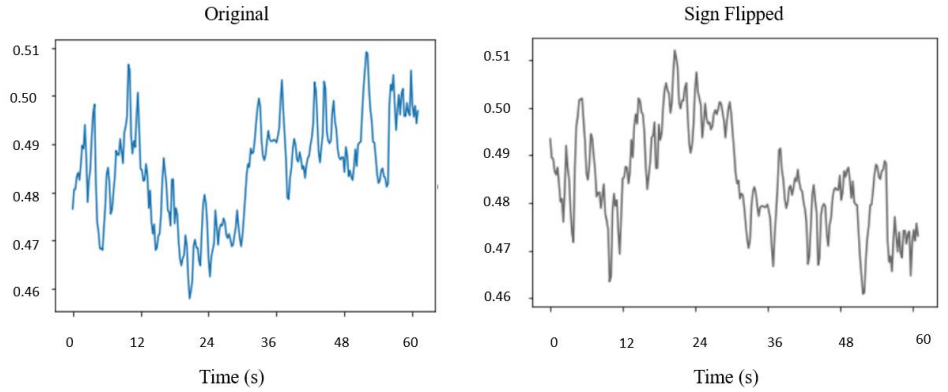
1- Sliding windows:



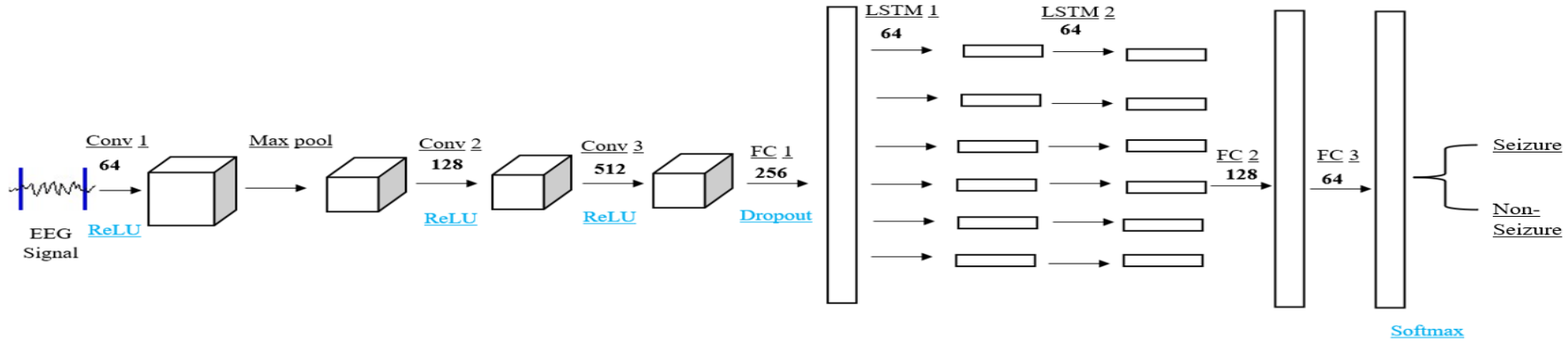
2- Time Reverse:



3-Flip Sign:



CNN-LSTM model



4 experiments:

Experiment 1: Train and test the model on the CHB-MIT data set without using DA techniques

Experiment 2: Same as the first one but using DA techniques

Experiment 3: Train the model on the CHB-MIT data set and test it on the CHU data set without using DA techniques

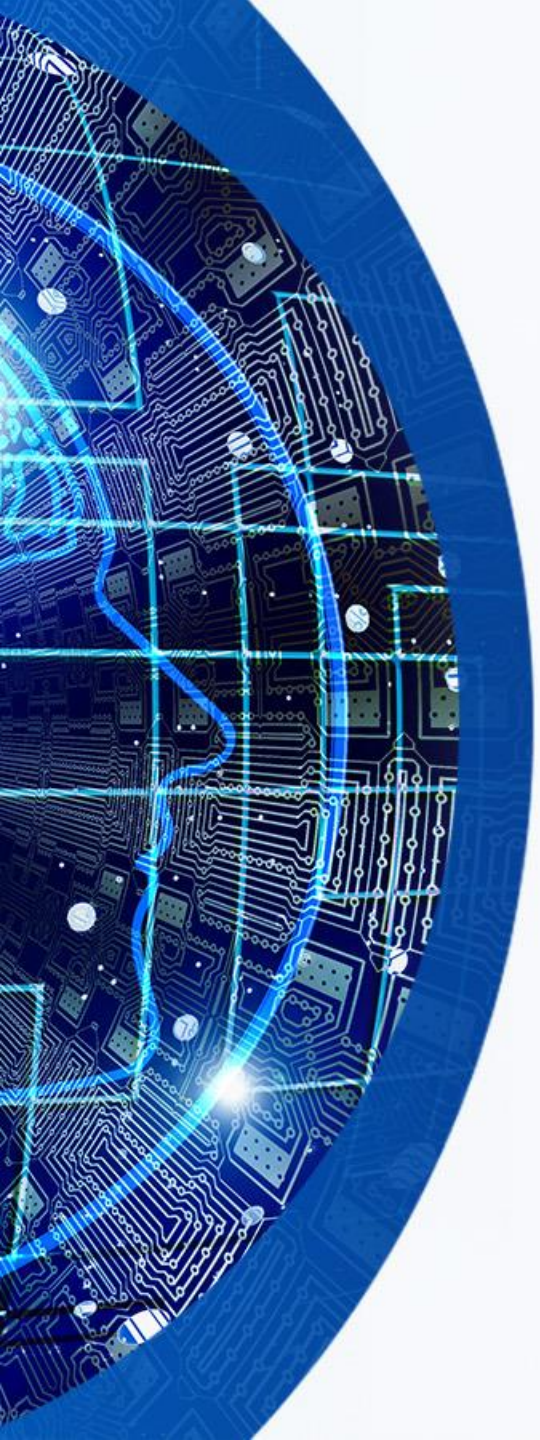
Experiment 4: Same as the third one but using DA techniques

Results

CNN-LSTM	Accuracy	Precision	Recall	F1-score
One site	98.3% ± 0.06	93.5% ± 0.08	93.5% ± 0.08	93.5% ± 0.08
One site - DA	98.7% ± 0.01	97.0% ± 0.01	96.8% ± 0.02	96.9% ± 0.01
Cross-site	87.3% ± 0.08	86.5% ± 0.07	87.0% ± 0.06	86.9% ± 0.07
Cross-site - DA	89.9% ± 0.05	89.1% ± 0.02	89.0% ± 0.04	88.9% ± 0.03

CNN	Accuracy	Precision	Recall	F1-score
One site	97.9% ± 0.12	92.5% ± 0.15	92.0% ± 0.15	92.2% ± 0.15
One site - DA	97.6% ± 0.10	96.0% ± 0.09	96.0% ± 0.10	96.0% ± 0.10
Cross-site	85.8% ± 0.20	85.3% ± 0.23	85.0% ± 0.25	85.1% ± 0.24
Cross-site - DA	89.7% ± 0.18	88.0% ± 0.17	87.0% ± 0.16	88.0% ± 0.17

LSTM	Accuracy	Precision	Recall	F1-score
One site	98.0% ± 0.12	93.0% ± 0.14	92.0% ± 0.13	92.5% ± 0.14
One site - DA	97.8% ± 0.09	96.5% ± 0.07	94.5% ± 0.06	95.5% ± 0.06
Cross-site	86.6% ± 0.21	86.0% ± 0.23	86.5% ± 0.21	86.2% ± 0.22
Cross-site - DA	89.7% ± 0.12	88.7% ± 0.14	88.0% ± 0.13	88.3% ± 0.14

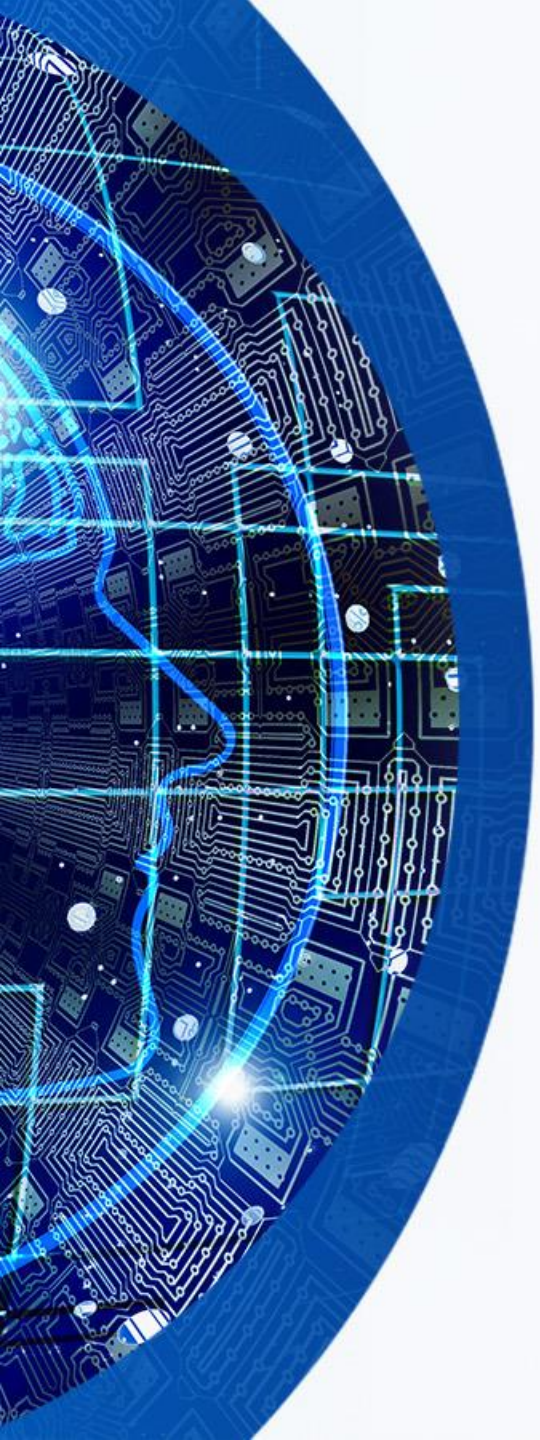


Comparison with state-of-the-art methods

Authors	Years	Methods	Accuracy	Precision	Recall	F1-score
Single-Site: Ozcan et al. [1]	2019	3D CNN	-	-	81.20%	-
Single-Site: Duan et al. [2]	2019	Bi-GRU	94.8%	-	91.7%	-
Single-Site: Sanguk Ryu [3]	2021	Dense-LSTM	93.3%	-	92.9%	92.3%
Cross-Site: Danielle et al.[4]	2021	CNN	-	-	89.3%	-
Our Cross-Site model	**2022**	**CNN-LSTM with DA**	**98.7%**	**97.0%**	**96.8%**	**96.9%**
Our Single-Site model	**2022**	**CNN-LSTM with DA**	**89.9%**	**88.5%**	**89.6%**	**89.0%**

References:

- [1] Ozcan, A. & Erturk, S. Seizure prediction in scalp EEG using 3D convolutional neural networks with an image-based approach. IEEE Transactions On Neural Systems And Rehabilitation Engineering. 27, 2284-2293 (2019)
- [2] Duan, L., Hou, J., Qiao, Y. & Miao, J. Epileptic seizure prediction based on convolutional recurrent neural network with multi-timescale. International Conference On Intelligent Science And Big Data Engineering. pp. 139-150 (2019)
- [3] Ryu, S. & Joe, I. A Hybrid DenseNet-LSTM model for epileptic seizure prediction. Applied Sciences. 11, 7661 (2021)
- [4] Currey, D., Hsu, D., Ahmed, R., Venkataraman, A. & Craley, J. Cross-site Epileptic Seizure Detection Using Convolutional Neural Networks.2021 55th Annual Conference On Information Sciences And Systems (CISS). pp. 1-6 (2021)



Conclusion and Future Work

- Very promising results
  - ➢ especially on imbalanced data sets
  - ➢ model robust to cross-site variability

- Future Work
  - ➢ develop reliable models that can be applied to different types of epilepsy using more sophisticated ML and DA methods.

For any question:

tala.abdallah@net.usj.edu.lb