A Multi-View Deep Learning Framework for EEG Seizure Detection

PAPER BY : YE YUAN, GUANGXU XUN, KEBIN JIA AND AIDONG ZHANG

PRESENTED BY : ABDUR RAFAY

SUPERVISED BY : DR.-ING. TANUJ HASIJA

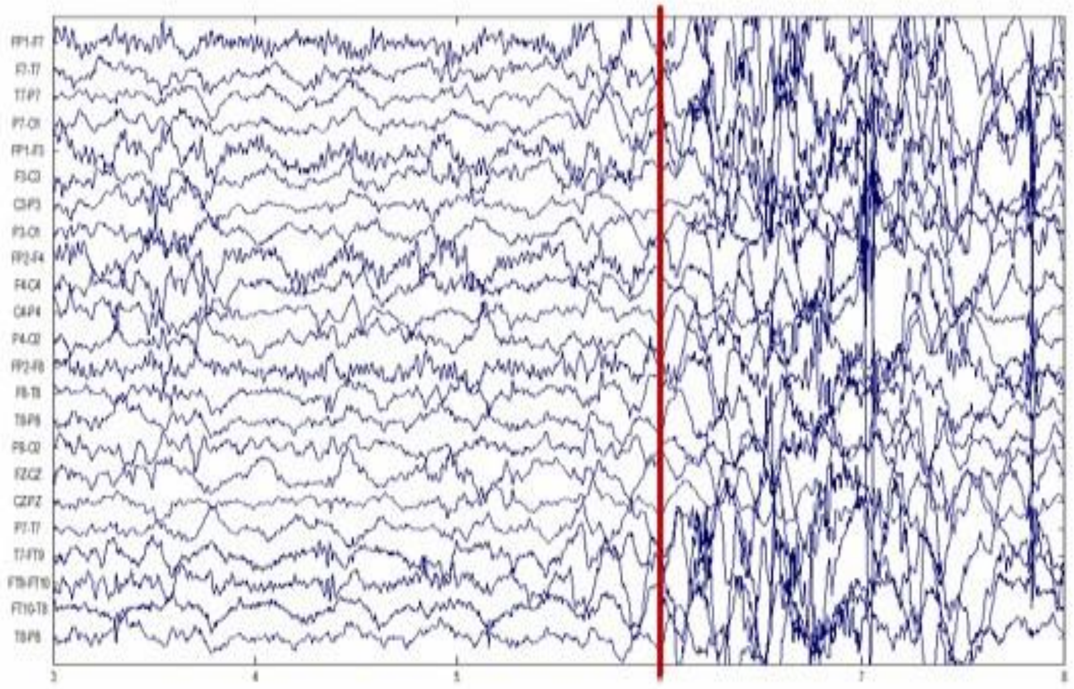
DATE : JANUARY 22, 2025

FACULTY OF ELECTRICAL ENGINEERING, INFORMATION TECHNOLOGY AND MATHEMATICS (EIM), SIGNAL AND SYSTEM THEORY GROUP (SST)

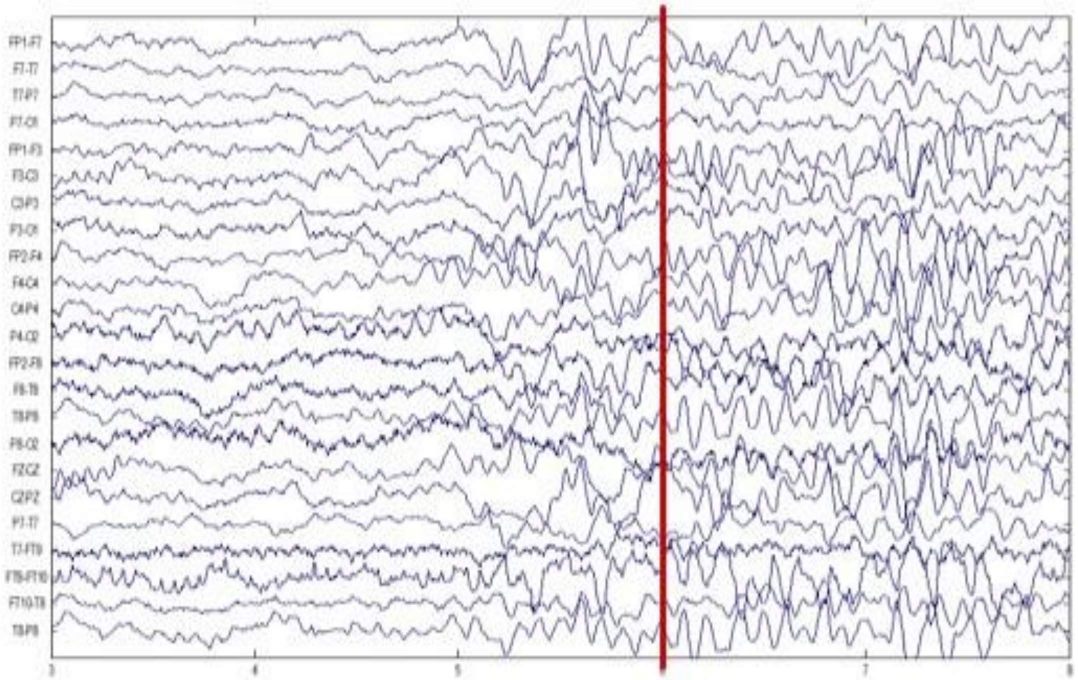
PADERBORN UNIVERSITY

Motivation

[1] Examples of EEG patterns across different patients



(a) Scalp EEG signals of patient A



(b) Scalp EEG signals of patient B

[1] Yuan et al. (2014)

Motivation

PADERBORN UNIVERSITY

Challenge

- Manual seizure detection is time-consuming
- Requires expert analysis due to high variability in rhythmic pattern
- Can potentially delay critical care to the patient

Approach

- An automated pipeline capable of processing EEG signals quickly and accurately
- Ensuring timely seizure detection and intervention

Literature Review - Papers

Category	Key Techniques/Methods	Highlights	Referenced Papers
Feature extraction	1. Time-domain features: statistical measures, entropy etc	Requires significant domain knowledge	Tong & Thakor (2009) ; Acharya et al. (2015)
	2. Spectral-domain features: Fourier transform (FFT), wavelet transform	Captures frequency-related changes in EEG signals	Samiee et al. (2015) ; Faust et al. (2015)
Deep Learning	1. CNNs: Detect spatial patterns in EEG data	Improved accuracy over time and spectral domain features	Johansen et al. (2016) ; Antoniades et al. (2016)
	2. Stacked Autoencoders: Unsupervised learning from EEG data	Handles high-dimensional data efficiently	Supratak et al. (2014) ; Gogna et al. (2017)

Literature Review – Potential gaps

Challenges/Limitations	Proposed Solutions/Improvements
Channel redundancy	Introduce **channel-aware modules** to focus on critical EEG channels and reduce irrelevant data.
Lack of multi-view learning	Develop **multi-view frameworks** to exploit inter- and intra-channel correlations in EEG signals.
Limited end-to-end frameworks	Design **unified training pipelines** combining unsupervised and supervised objectives.

Outline

PADERBORN UNIVERSITY

1. **Overview of the research paper**
    - 1.1 The overall pipeline
    - 1.2 Role of autoencoders
    - 1.3 Role of channel aware module
    - 1.4 Combine training and loss

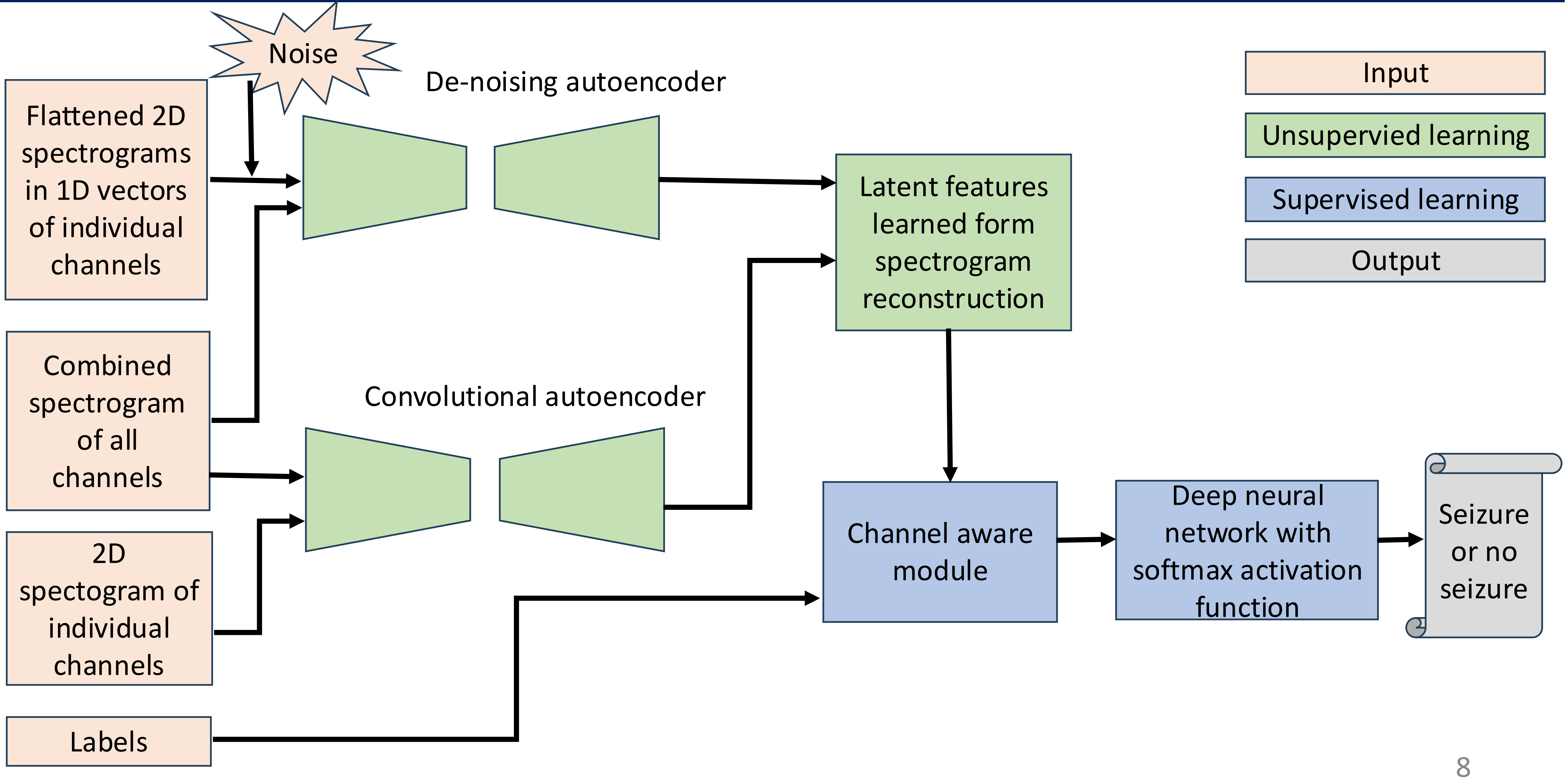
2. **Evaluation and results**
    - 2.1 Dataset and Input processing
    - 2.2 Evaluation - Research paper
    - 2.3 Evaluation - My Implementation

3. **Conclusion and key takeaways**

Outline

PADERBORN UNIVERSITY

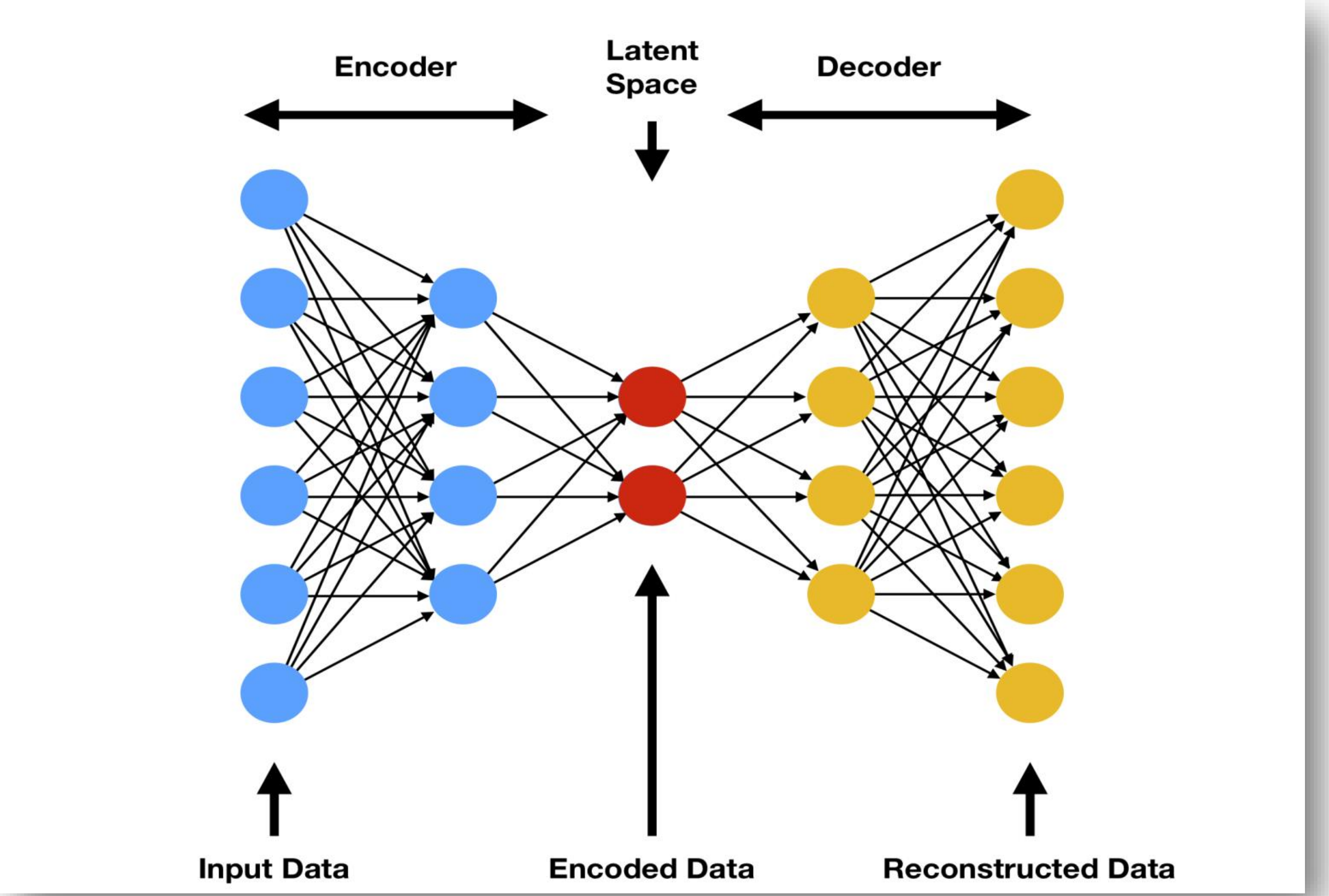
1.1 The overall pipeline



[2] Autoencoders :

- A type of neural network designed to learn efficient, compressed representations of input data

- Consists of an encoder (compresses input into a latent space) and a decoder (reconstructs data from the latent space)



[2] Supratak et al. (2014)

PADERBORN UNIVERSITY

[3]**Denoising Autoencoders (DA)**:

- Learns representations (latent features) by reconstructing corrupted inputs

- DA cross channel reconstruction can be represented as:

$$\hat{x}_r = Decoder_{cross}(Encoder_{cross}(x_{1:C}))$$

$$\hat{x}_r = f(W_r' f(W_r \bar{x}_{1:C} + b_r) + b_r')$$

- DA intra channel reconstruction can be represented as:

$$\hat{x}_c = Decoder_{intra}(Encoder_{intra}(x_c))$$

$$\hat{x}_c = f(W_c' f(W_c \bar{x}_c + b_c) + b_c'),$$

[3]Gogna et al. (2017)

[4]**Convolutional Autoencoder (ConvA)**:

- Retains spatial locality in spectrograms better by using convolution layers to share weights across input locations
- ConvA cross channel reconstruction can be represented as

$$\hat{x}_r^c = f\left(\sum_{k\in H_r} h_r^k * \tilde{W}_{rc}^k + c_r^c\right),$$

$$h_r^k = f\left(\sum_{c=1}^{C} x_c * W_{rc}^k + b_r^k\right),$$

- ConvA intra channel reconstruction can be represented as

$$\hat{x}_c = f\left(\sum_{k\in H_c} h_c^k * \tilde{W}_{cc}^k + c_{cc}\right),$$

$$h_c^k = f(x_c * W_{cc}^k + b_{cc}^k),$$

[4]Gogna et al. (2017)

Combine objective / loss function:

- Combines cross-entropy reconstruction losses for both cross-channel and intra-channel autoencoders for unsupervised learning

$$J_{\mathrm{MAE}} = \frac{1}{m} \sum_{i=1}^{m} \mathrm{LIH}(x^{(i)}, \hat{x}_r^{(i)}) + \sum_{c=1}^{C} \mathrm{LIH}(x_c^{(i)}, \hat{x}_c^{(i)})$$

Cross Entropy:

- Cross Entropy can be defined as:

$$\mathrm{LIH}(x, \hat{x}) = -\sum_{j=1}^{n} [x_j \log(\hat{x}_j) + (1 - x_j) \log(1 - \hat{x}_j)]$$

PADERBORN UNIVERSITY

Channel aware module for seizure detection:

- First input to calculate the response energy is intra channel latent features ($h_1, h_2, \ldots, h_c$)

- Measure channel importance using response energy. Retain the top channels based on R_c and some threshold. Response energy is calculated as:

$$R_c = ||h_c - \hat{\rho}_c||^2,$$

Fused hidden representation:

- Second input is the cross channel latent features. Through the combination of selected intra-channel and cross-channel latent features, a fused unified representation is created

$$h_u = f(W_u[h_r; \bar{h}_c] + b_u),$$

- This fused hidden representation (h_u) is now ready for classification. It is combined with the labels for supervised learning

$$\hat{y} = softmax(W_s h_u + b_s),$$

Combined training and optimization objective:

- Combines unsupervised feature learning with supervised seizure detection.
- Unsupervised component:

$$J_{\mathrm{MAE}} = \frac{1}{m}\sum_{i=1}^{m} \mathrm{LIH}(x^{(i)}, \hat{x}_r^{(i)}) + \sum_{c=1}^{C} \mathrm{LIH}(x_c^{(i)}, \hat{x}_c^{(i)})$$

- Supervised component:

$$J_{MSD} = \frac{1}{m}\sum_{i=1}^{m} \mathcal{L}_{\mathrm{IH}}(y^{(i)}, \hat{y}^{(i)}),$$

Outline

2.1 Dataset and input processing

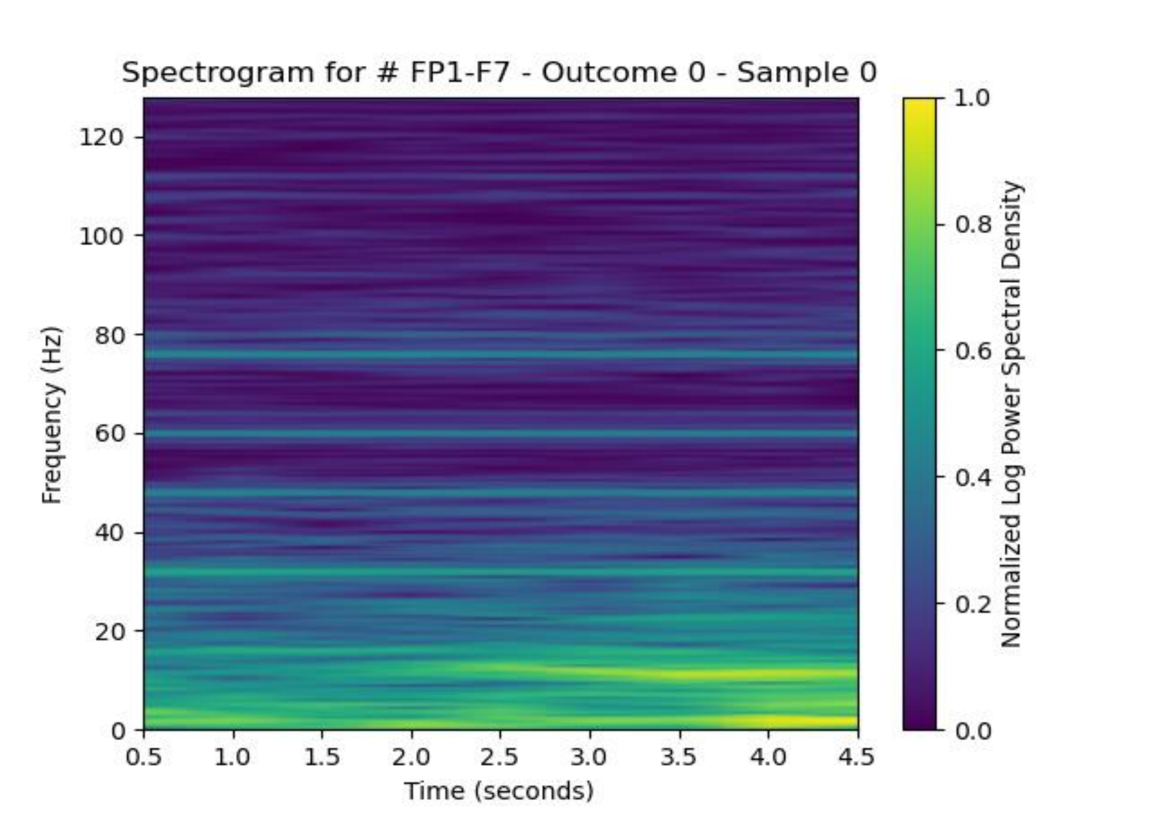
PADERBORN UNIVERSITY

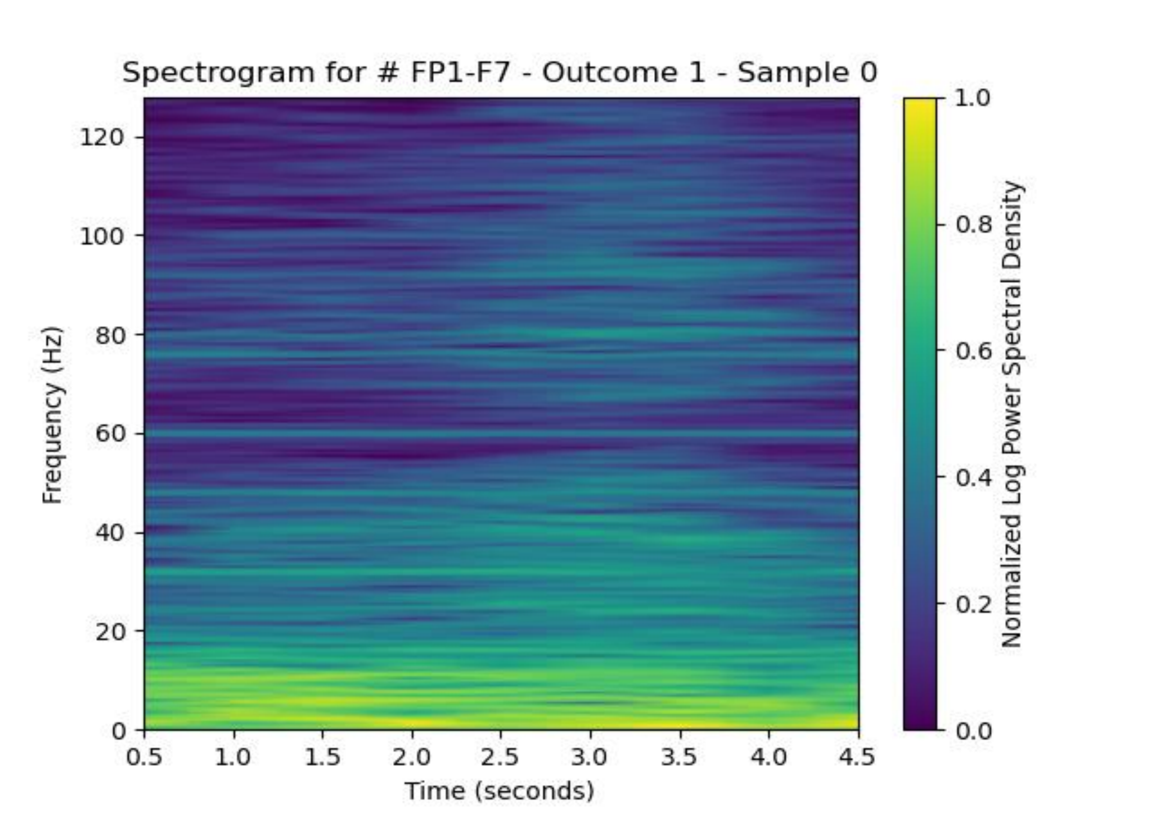
Features	Explanations
Dataset	CHB-MIT Scalp EEG Database that has around 2 million datapoints for 23 patients (5M, 18F) & (ages between 2 – 22 years) in csv format
Sampling rate	The data is recorded at a sampling rate of 256 Hz
No. of channels	The data is recorded with 23 bipolar channels
Class distribution	Ictal (seizure) and Non-ictal (no seizure): 50% of the data each for both the classes

2.1 Dataset and input processing (cont.)

PADERBORN UNIVERSITY

- The EEG signals are converted into 2D spectrograms using a window size of 4.5 secs and an overlap of 1 sec





2.2 Evaluation – Research paper

PADERBORN UNIVERSITY

Performance metrics

- F1 Score
- Test accuracy
- Precision – Recall (PR) curve
- Receiver Operator Characteristics (ROC) curve

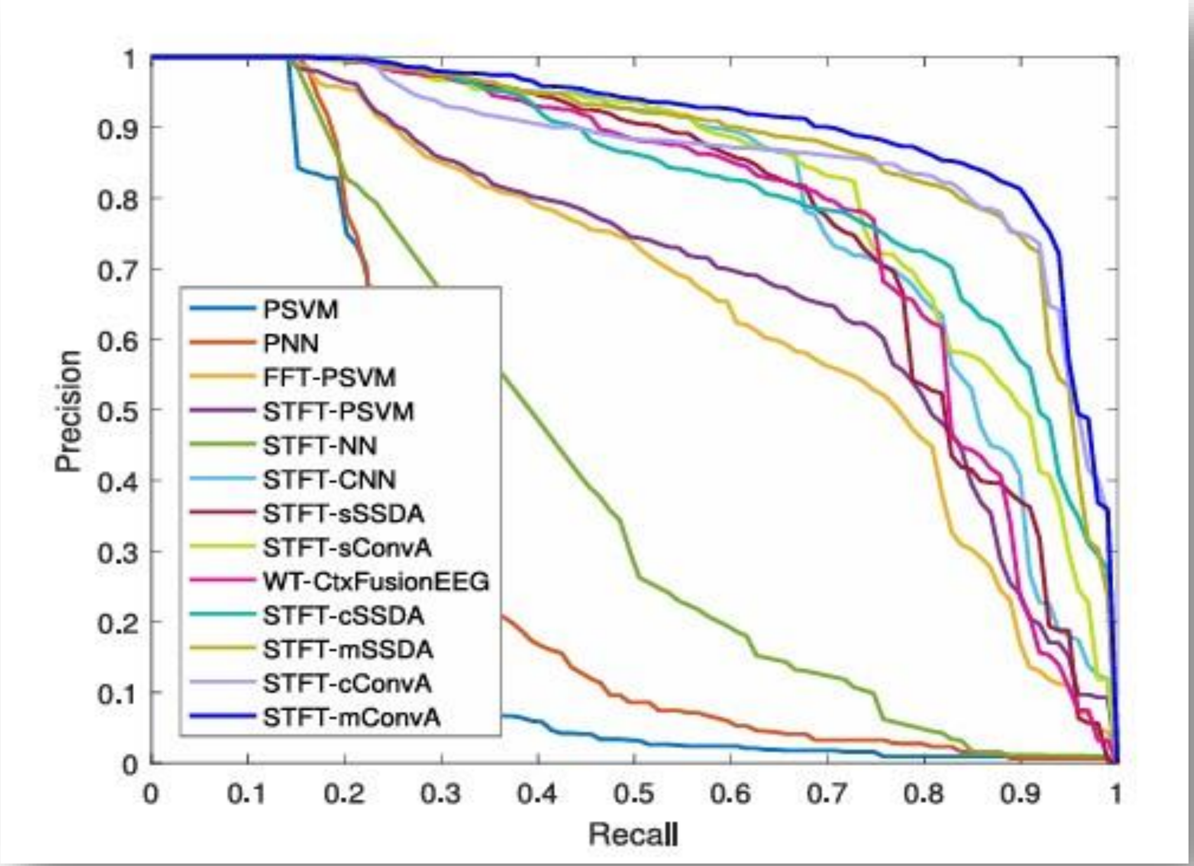
2.2 Evaluation - Research Paper (cont.)

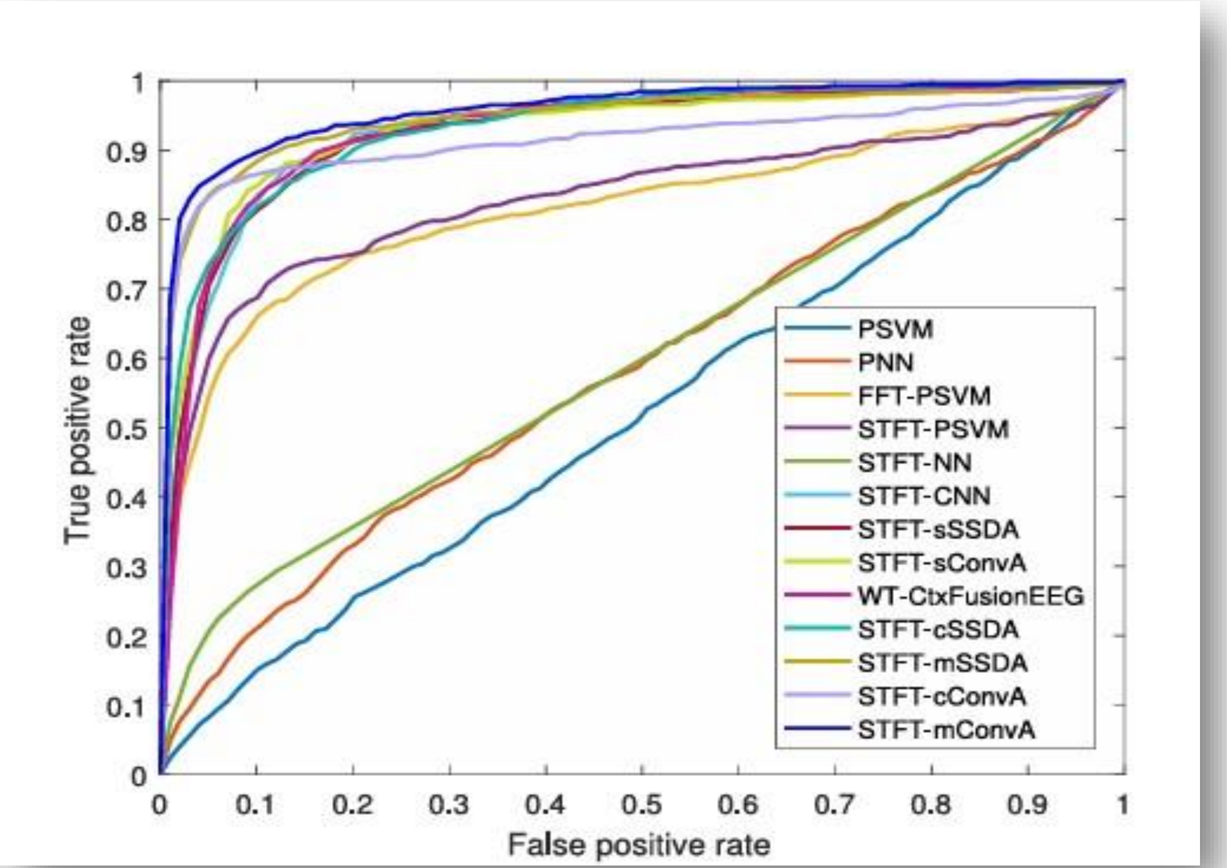
- The methodology used in this paper, referred to as STFT ConvA(m), achieved the highest F1 score and accuracy compared to other baseline classifiers

Models	F1 score	Accuracy
FFT PSVM	0.61	86%
STFT PSVM	0.64	88%
STFT CNN	0.75	88%
STFT SSDAs	0.74	89%
STFT ConvAs	0.76	89%
WT-CTX Fusion EEG	0.72	90%
STFT ConvAm	**0.85**	**94%**

PADERBORN UNIVERSITY

- Additionally, the PR (Precision-Recall) and ROC (Receiver Operating Characteristic) curves of the mentioned methodology surpass those of the other baseline models

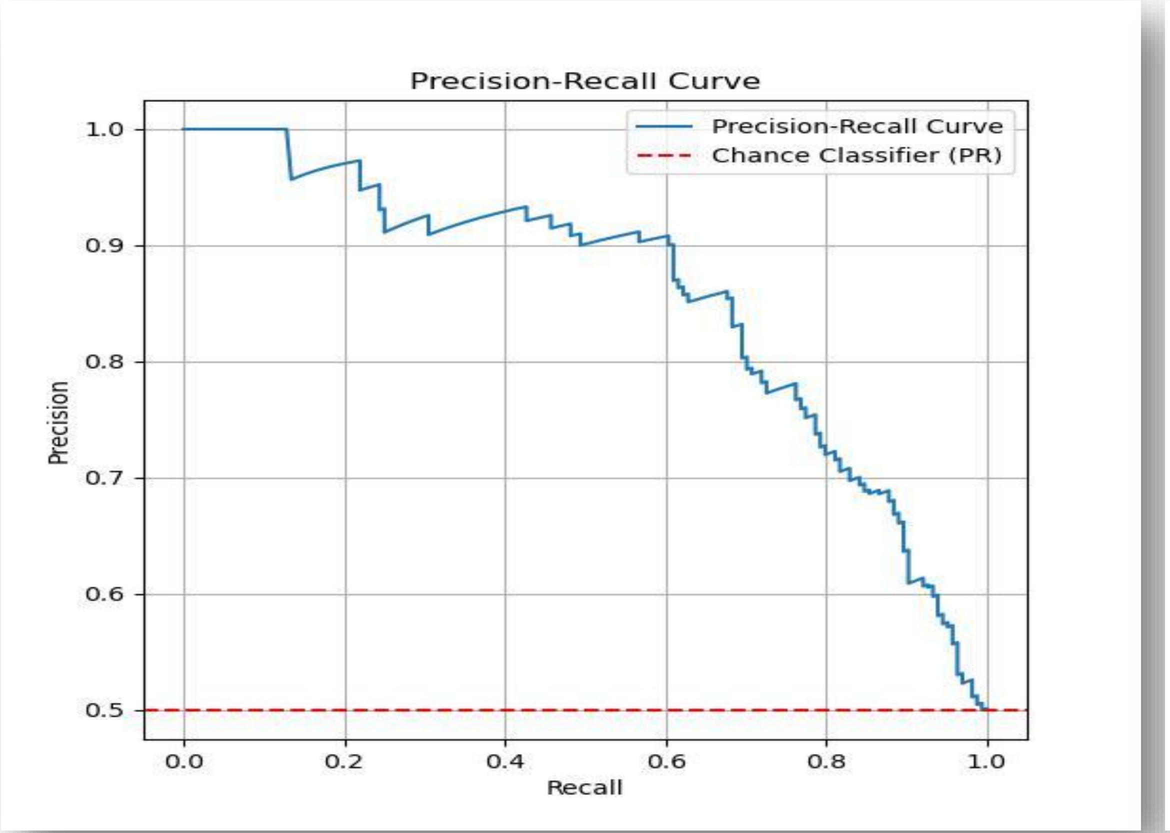


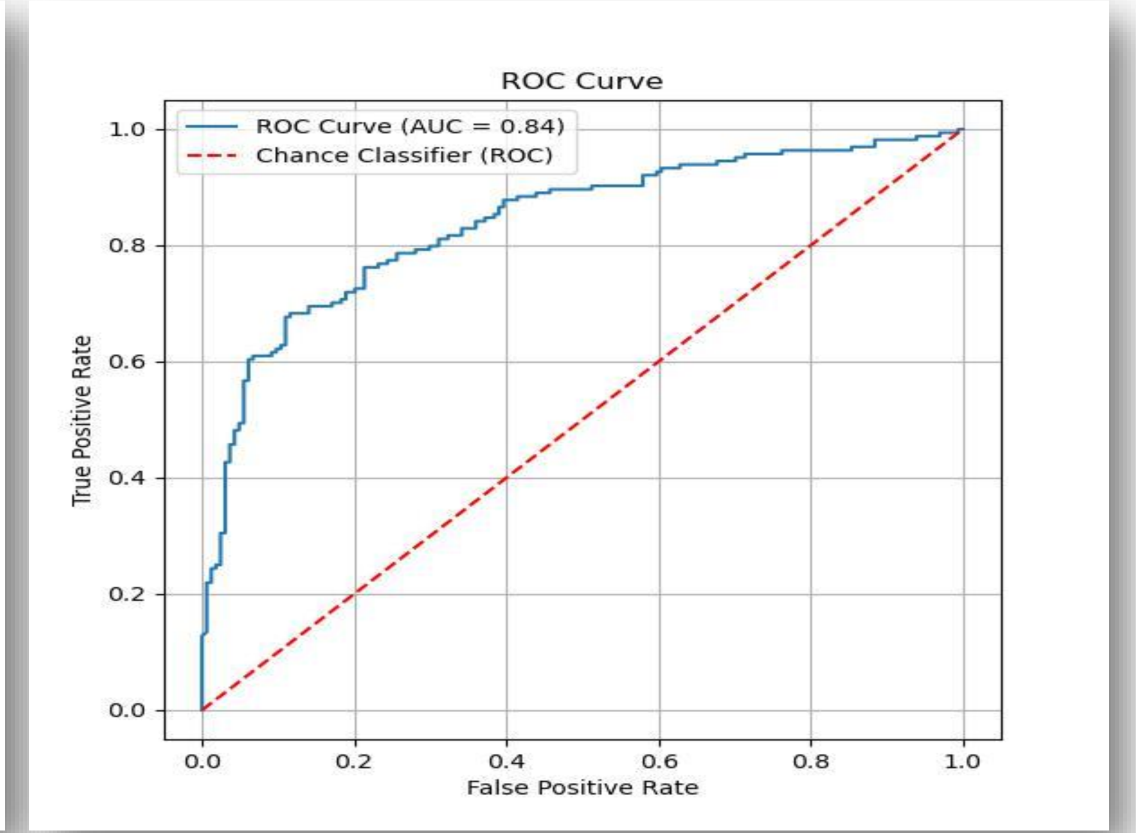


2.3 Evaluation - My implementation

PADERBORN UNIVERSITY

- The implementation of the methodology in this paper resulted in a test accuracy of 81%, compared to the benchmark of 94%
- The average F1 score for both classes is 0.78, compared to benchmark of 0.85





2.3 Evaluation - My implementation (cont.)

PADERBORN UNIVERSITY

There is a difference between the results of my implementation and those achieved by the authors of the paper. The primary reason for this disparity lies in the data used:

- **Authors' data**: Included 96 channels with timestamps recorded in real-time, from which they generated labels.
- **My data**: Used a pre-processed version with only 23 channels, and the labels were already provided in a CSV format.

- This difference in data likely caused the inferior performance of my model.

Outline

PADERBORN UNIVERSITY

1. **Overview of the research paper**

    1.1 The overall pipeline

    1.2 Role of autoencoders

    1.3 Role of channel aware module

    1.4 Combine training and loss

2. **Evaluation and results**

    2.1 Dataset and Input processing

    2.2 Evaluation - Research paper

    2.3 Evaluation - My Implementation

3. **Conclusion and key takeaways**

3 Summary

PADERBORN UNIVERSITY

- **Unified Framework:** End-to-end multi-view deep learning for EEG seizure detection using cross and intra channel correlations.

- **Multi-View Autoencoders:** Utilized DA and ConvA autoencoders for multi-channel EEG analysis.

- **Channel-Aware Detection:** Focused on key EEG channels with a channel-aware seizure detection module.

- **Broad Applicability:** Validated on CHB-MIT dataset; adaptable to similar medical tasks.

Thank you!

PADERBORN UNIVERSITY

Questions?