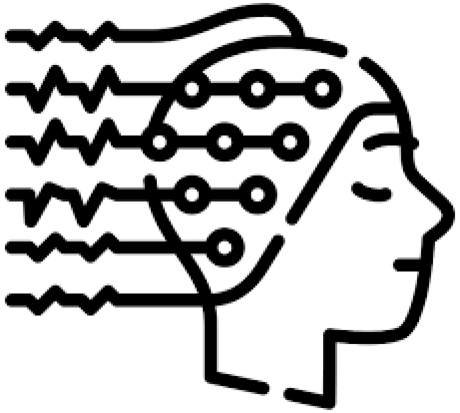
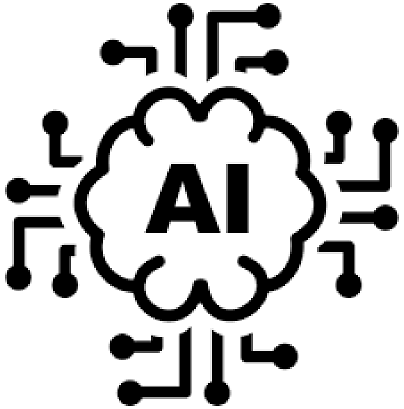




Machine Learning Project

EEG & Auditory Attention Detection

Foad Moslem

EEG & Auditory Attention Detection

Reference Paper

- (2021) Extracting the Auditory Attention in a Dual - Speaker Scenario from EEG using a Joint CNN-LSTM Model

  https://arxiv.org/pdf/2102.03957v2.pdf

- Aim:
  - Presents a novel neural network framework that makes use of the speech spectrogram of multiple speakers and the EEG signals as inputs to classify the auditory attention

---

- Datasets:
  - FAU Dataset (Kuruvila et al., 2021)
  - DTU Dataset (Fuglsang et al., 2018)
  - KUL Dataset (Das et al., 2019)

Our Work

- Final Project of Machine Learning Course
- EEG & Auditory Attention Detection

- Aim:
  - Preprocessing (Segmentation & feature extraction)
  - Rebuild model
  - The Train And Test ACC For Model

- Dataset:
  - KUL Dataset (Das et al., 2019)

  https://zenodo.org/record/3377911#.ZGytBOxBzEY

EEG & Auditory Attention Detection

KUL Dataset (Das et al., 2019)

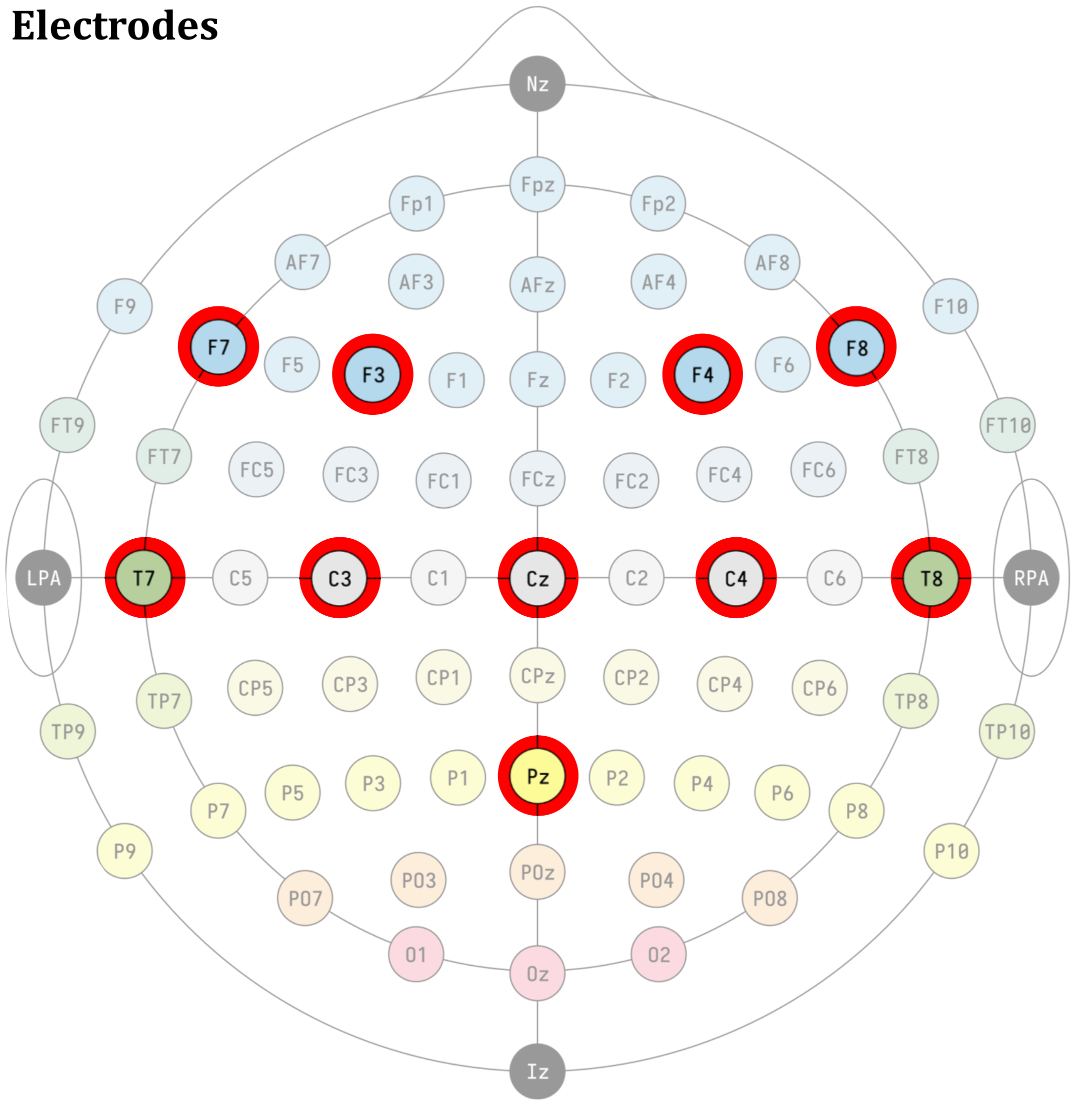
- 16 subjects
- Conditions: HRTF, dichotic and repeated
- 3 experiments for each subject
- 4 presentation in each experiment
- Each presentation contained 2 audio played simultaneously
- Each presentation is approximately 6 minutes
- EEG analyzed in 64 electrodes

KUL Dataset (Reference Paper & Our Work)

- 16 subjects
- Conditions: dichotic
- 2 experiments for each subject
- 2 presentation in each experiment
- Each presentation contained 2 audio played simultaneously
- Each presentation is approximately 6 minutes
- EEG analyzed in 10 electrodes

EEG & Auditory Attention Detection

Electrodes



Stimuli

4 Stories (Parts)

- Part 1
  - Track #1 – 6~7 min
  - Track #2 – 6~7 min
- Part 2
  - Track #1 – 6~7 min
  - Track #2 – 6~7 min
- Part 3
  - Track #1 – 6~7 min
  - Track #2 – 6~7 min
- Part 4
  - Track #1 – 6~7 min
  - Track #2 – 6~7 min

Subject # (within a single recording session)

Experiment #1
Experiment #2

- Presentation #1 – HRTF stimuli – Part ? – Track ? & Part ? – Track ? (played simultaneously)
- Presentation #2 – HRTF stimuli – Part ? – Track ? & Part ? – Track ? (played simultaneously)
- Presentation #3 – Dichotic stimuli – Part ? – Track ? & Part ? – Track ? (played simultaneously)
- Presentation #4 – Dichotic stimuli – Part ? – Track ? & Part ? – Track ? (played simultaneously)

Questions Asked To The Subject After Each Presentation

EEG & Auditory Attention Detection

Dataset

subject	TrialID	condition	repetition	experiment	part	attended track	attended ear	stimuli1	stimuli2
S1	3	dry	0	1	3	1	R	part3_track2_dry.wav	part3_track1_dry.wav
S1	4	dry	0	1	4	1	L	part4_track1_dry.wav	part4_track2_dry.wav
S1	7	dry	0	2	3	2	R	part3_track1_dry.wav	part3_track2_dry.wav
S1	8	dry	0	2	4	2	L	part4_track2_dry.wav	part4_track1_dry.wav
S2	2	dry	0	1	2	1	R	part2_track2_dry.wav	part2_track1_dry.wav
S2	3	dry	0	1	3	1	L	part3_track1_dry.wav	part3_track2_dry.wav
S2	6	dry	0	2	2	2	R	part2_track1_dry.wav	part2_track2_dry.wav
S2	7	dry	0	2	3	2	L	part3_track2_dry.wav	part3_track1_dry.wav
S3	1	dry	0	1	1	1	R	part1_track2_dry.wav	part1_track1_dry.wav
S3	2	dry	0	1	2	1	L	part2_track1_dry.wav	part2_track2_dry.wav
S3	5	dry	0	2	1	2	R	part1_track1_dry.wav	part1_track2_dry.wav
S3	6	dry	0	2	2	2	L	part2_track2_dry.wav	part2_track1_dry.wav
S4	1	dry	0	1	1	1	L	part1_track1_dry.wav	part1_track2_dry.wav
S4	4	dry	0	1	4	1	R	part4_track2_dry.wav	part4_track1_dry.wav
S4	5	dry	0	2	1	2	L	part1_track2_dry.wav	part1_track1_dry.wav
S4	8	dry	0	2	4	2	R	part4_track1_dry.wav	part4_track2_dry.wav
S5	3	dry	0	1	3	1	R	part3_track2_dry.wav	part3_track1_dry.wav
S5	4	dry	0	1	4	1	L	part4_track1_dry.wav	part4_track2_dry.wav
S5	7	dry	0	2	3	2	R	part3_track1_dry.wav	part3_track2_dry.wav
S5	8	dry	0	2	4	2	L	part4_track2_dry.wav	part4_track1_dry.wav
S6	2	dry	0	1	2	1	R	part2_track2_dry.wav	part2_track1_dry.wav
S6	3	dry	0	1	3	1	L	part3_track1_dry.wav	part3_track2_dry.wav
S6	6	dry	0	2	2	2	R	part2_track1_dry.wav	part2_track2_dry.wav
S6	7	dry	0	2	3	2	L	part3_track2_dry.wav	part3_track1_dry.wav
S7	1	dry	0	1	1	1	R	part1_track2_dry.wav	part1_track1_dry.wav
S7	2	dry	0	1	2	1	L	part2_track1_dry.wav	part2_track2_dry.wav
S7	5	dry	0	2	1	2	R	part1_track1_dry.wav	part1_track2_dry.wav
S7	6	dry	0	2	2	2	L	part2_track2_dry.wav	part2_track1_dry.wav
S8	1	dry	0	1	1	1	L	part1_track1_dry.wav	part1_track2_dry.wav
S8	4	dry	0	1	4	1	R	part4_track2_dry.wav	part4_track1_dry.wav
S8	5	dry	0	2	1	2	L	part1_track2_dry.wav	part1_track1_dry.wav
S8	8	dry	0	2	4	2	R	part4_track1_dry.wav	part4_track2_dry.wav

subject	TrialID	condition	repetition	experiment	part	attended track	attended ear	stimuli1	stimuli2
S9	3	dry	0	1	3	1	R	part3_track2_dry.wav	part3_track1_dry.wav
S9	4	dry	0	1	4	1	L	part4_track1_dry.wav	part4_track2_dry.wav
S9	7	dry	0	2	3	2	R	part3_track1_dry.wav	part3_track2_dry.wav
S9	8	dry	0	2	4	2	L	part4_track2_dry.wav	part4_track1_dry.wav
S10	2	dry	0	1	2	1	R	part2_track2_dry.wav	part2_track1_dry.wav
S10	3	dry	0	1	3	1	L	part3_track1_dry.wav	part3_track2_dry.wav
S10	6	dry	0	2	2	2	R	part2_track1_dry.wav	part2_track2_dry.wav
S10	7	dry	0	2	3	2	L	part3_track2_dry.wav	part3_track1_dry.wav
S11	1	dry	0	1	1	1	R	part1_track2_dry.wav	part1_track1_dry.wav
S11	2	dry	0	1	2	1	L	part2_track1_dry.wav	part2_track2_dry.wav
S11	5	dry	0	2	1	2	R	part1_track1_dry.wav	part1_track2_dry.wav
S11	6	dry	0	2	2	2	L	part2_track2_dry.wav	part2_track1_dry.wav
S12	1	dry	0	1	1	1	L	part1_track1_dry.wav	part1_track2_dry.wav
S12	4	dry	0	1	4	1	R	part4_track2_dry.wav	part4_track1_dry.wav
S12	5	dry	0	2	1	2	L	part1_track2_dry.wav	part1_track1_dry.wav
S12	8	dry	0	2	4	2	R	part4_track1_dry.wav	part4_track2_dry.wav
S13	3	dry	0	1	3	1	R	part3_track2_dry.wav	part3_track1_dry.wav
S13	4	dry	0	1	4	1	L	part4_track1_dry.wav	part4_track2_dry.wav
S13	7	dry	0	2	3	2	R	part3_track1_dry.wav	part3_track2_dry.wav
S13	8	dry	0	2	4	2	L	part4_track2_dry.wav	part4_track1_dry.wav
S14	2	dry	0	1	2	1	R	part2_track2_dry.wav	part2_track1_dry.wav
S14	3	dry	0	1	3	1	L	part3_track1_dry.wav	part3_track2_dry.wav
S14	6	dry	0	2	2	2	R	part2_track1_dry.wav	part2_track2_dry.wav
S14	7	dry	0	2	3	2	L	part3_track2_dry.wav	part3_track1_dry.wav
S15	1	dry	0	1	1	1	R	part1_track2_dry.wav	part1_track1_dry.wav
S15	2	dry	0	1	2	1	L	part2_track1_dry.wav	part2_track2_dry.wav
S15	5	dry	0	2	1	2	R	part1_track1_dry.wav	part1_track2_dry.wav
S15	6	dry	0	2	2	2	L	part2_track2_dry.wav	part2_track1_dry.wav
S16	1	dry	0	1	1	1	L	part1_track1_dry.wav	part1_track2_dry.wav
S16	4	dry	0	1	4	1	R	part4_track2_dry.wav	part4_track1_dry.wav
S16	5	dry	0	2	1	2	L	part1_track2_dry.wav	part1_track1_dry.wav
S16	8	dry	0	2	4	2	R	part4_track1_dry.wav	part4_track2_dry.wav

EEG & Auditory Attention Detection

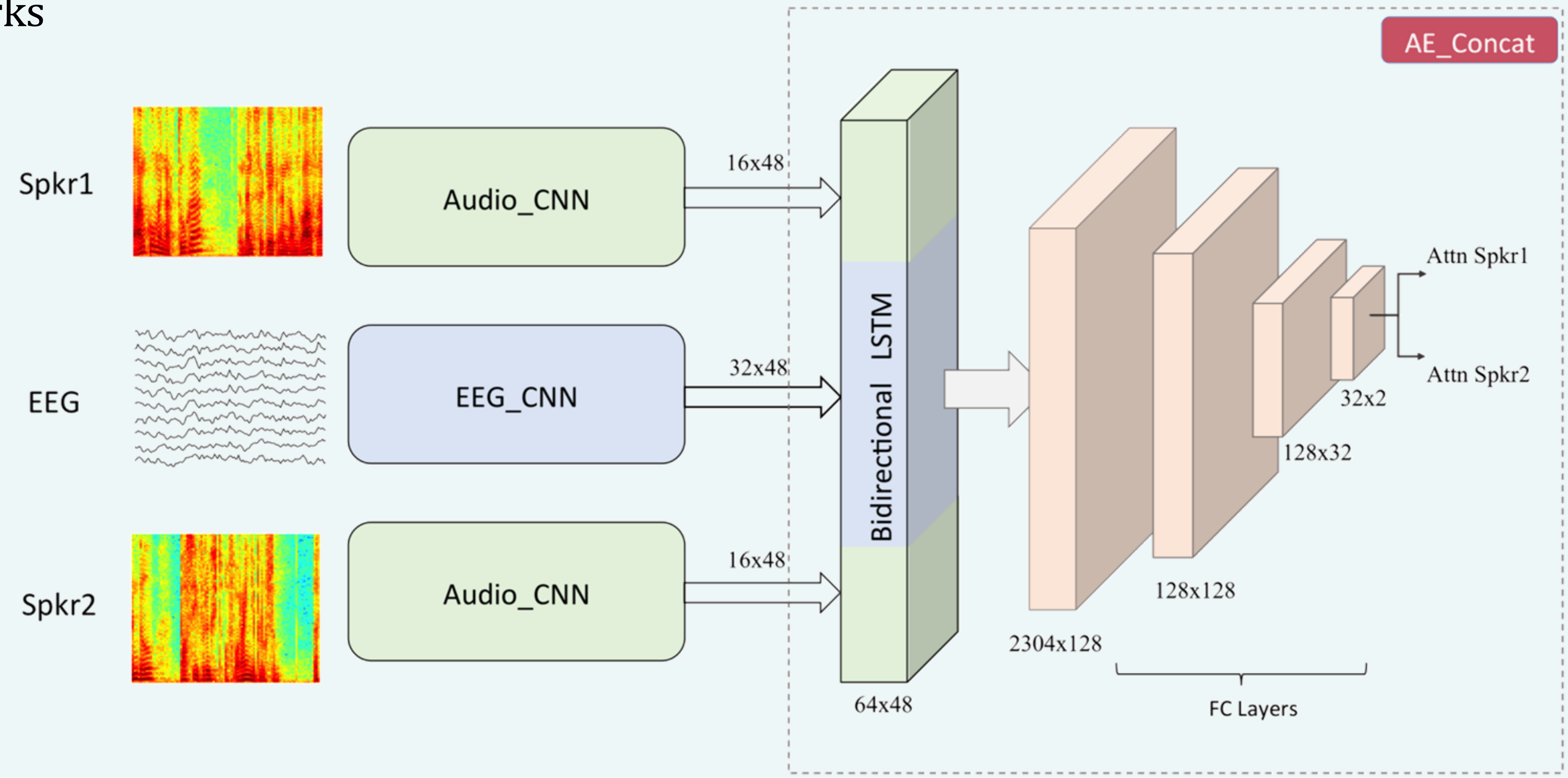
Dataset – All epochs

index	EEG	Audio1 (Left Ear)	Audio2 (Right Ear)	stimuli (Target)
0_0	S1_trials3_10Ch_epoch0	part3_track2_dry.wav_epoch0	part3_track1_dry.wav_epoch0	2
0_1	S1_trials3_10Ch_epoch1	part3_track2_dry.wav_epoch1	part3_track1_dry.wav_epoch1	2
0_2	S1_trials3_10Ch_epoch2	part3_track2_dry.wav_epoch2	part3_track1_dry.wav_epoch2	2
0_3	S1_trials3_10Ch_epoch3	part3_track2_dry.wav_epoch3	part3_track1_dry.wav_epoch3	2
0_4	S1_trials3_10Ch_epoch4	part3_track2_dry.wav_epoch4	part3_track1_dry.wav_epoch4	2
⋮	⋮	⋮	⋮	⋮
39_383	S10_trials7_10Ch_epoch383	part3_track2_dry.wav_epoch383	part3_track1_dry.wav_epoch383	1
39_384	S10_trials7_10Ch_epoch384	part3_track2_dry.wav_epoch384	part3_track1_dry.wav_epoch384	1
39_385	S10_trials7_10Ch_epoch385	part3_track2_dry.wav_epoch385	part3_track1_dry.wav_epoch385	1
40_0	S11_trials1_10Ch_epoch0	part1_track2_dry.wav_epoch0	part1_track1_dry.wav_epoch0	2
40_1	S11_trials1_10Ch_epoch1	part1_track2_dry.wav_epoch1	part1_track1_dry.wav_epoch1	2
40_2	S11_trials1_10Ch_epoch2	part1_track2_dry.wav_epoch2	part1_track1_dry.wav_epoch2	2
⋮	⋮	⋮	⋮	⋮
63_393	S16_trials8_10Ch_epoch393	part4_track1_dry.wav_epoch393	part4_track2_dry.wav_epoch393	2
63_394	S16_trials8_10Ch_epoch394	part4_track1_dry.wav_epoch394	part4_track2_dry.wav_epoch394	2
63_395	S16_trials8_10Ch_epoch395	part4_track1_dry.wav_epoch395	part4_track2_dry.wav_epoch395	2

EEG & Auditory Attention Detection

Subnetworks

- Three subnetworks
  - EEG_CNN
  - Audio_CNN
  - AE_Concat

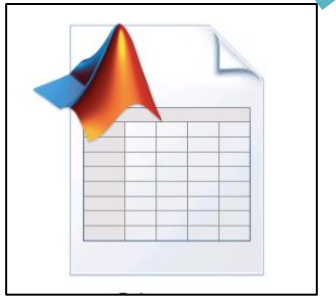


EEG & Auditory Attention Detection

EEG CNN

	Number of Kernels	Kernel Size	Dilation	Padding	Maxpool
Layer 1	32	24x1	1,1	12,0	2,1
Layer 2	32	7x1	2,1	6,0	1,2
Layer 3	32	7x5	1,1	3,2	2,5
Layer 4	32	7x1	1,1	3,0	1,1

- MNE Library
- Downsample EEG Data From 128 Hz To 64 Hz
- Filter Band-pass Frequencies Between 1 Hz And 32 Hz
- Normalize filtered EEG Signals
- Segment filtered EEG File Into Trials With A Duration Of 3 Seconds With 2 Seconds Overlap
- A Convolutional Neural Network With Four Layers Based On Table 3
- Input Is The Epochs Of Filtered EEG Signals
- Apply Max Pooling, Batch Normalization, Dropout And Relu Activation To The Convolution Output
- Output With A Fixed Dimension At 48x32



EEG_Extractor_From_mat_files.m

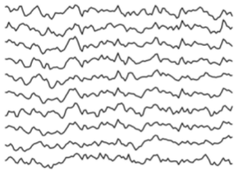
1_2_Specified_EEG_Dataset_Extracter.py

1_3_EEG_Processing.py

1_4_EEG_CNN_Function.py

1_5_EEG_CNN_FeatureMaps.py

EEG

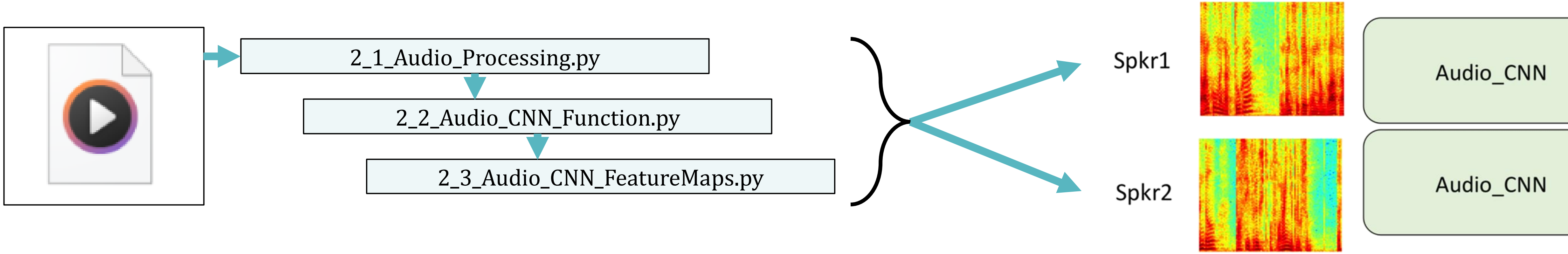


EEG_CNN

Audio CNN

- LIBROSA Library
- Cut-Off Frequency Of 8 KHz
- Downsample Audios From 48 KHz To 16 KHz
- Segment Audio Files Into Trials With A Duration Of 3 Seconds With 2 Seconds Overlap
- Obtain The Spectrogram Of Each Trial Of Audio Files by STFT
- Using A Hann Window Of 32 ms Duration With A 12 ms Overlap
- A Convolutional Neural Network With Five Layers Based On Table 4
- Apply Max Pooling, Batch Normalization, Dropout And Relu Activation To The Convolution Output
- Output With A Fixed Dimension At 48x16

	Number of Kernels	Kernel Size	Dilation	Padding	Maxpool
Layer 1	32	1x7	1,1	0,3	1,1
Layer 2	32	7x1	1,1	0,0	1,4
Layer 3	32	3x5	8,8	0,16	1,2
Layer 4	32	3,3	16,16	0,16	1,1
Layer 5	1	1x1	1,1	0,0	2,2

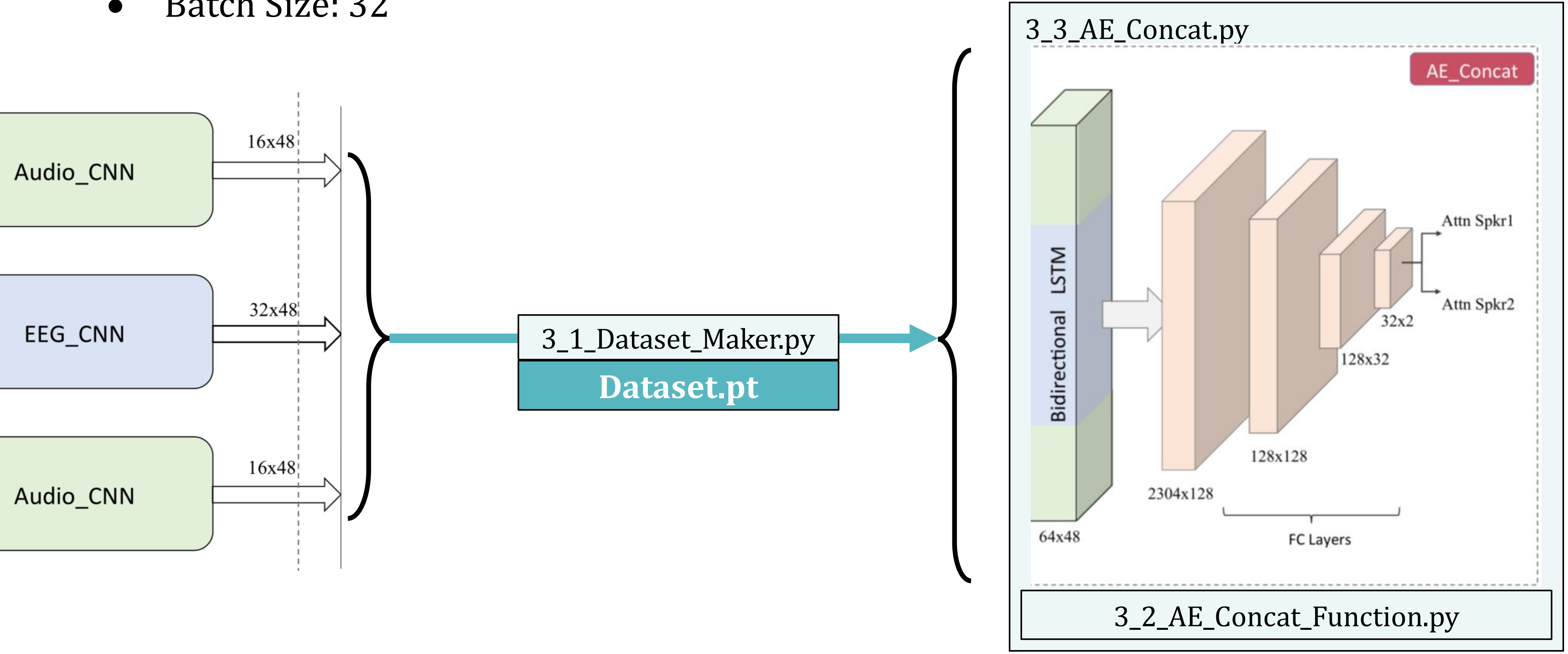


EEG & Auditory Attention Detection

AE Concat

- Concatenated Feature Map With Dimensions Of 48x64
- Bidirectional Long Short-Term Memory (blstm) Layer
- Four Fully Connected (FC)
- 80 Epochs For Trained
- Batch Size: 32

- Learning Rate: 5∗E−4
- Drop Out Probability: 0.25
- Optimize Using Adam Optimizer
- Loss Function Of Binary Cross Entropy



Epoch 10, Loss: 0.6930

Epoch 20, Loss: 0.6928

Epoch 30, Loss: 0.6925

Epoch 40, Loss: 0.6914

Epoch 50, Loss: 0.6895

Epoch 60, Loss: 0.6848

Epoch 70, Loss: 0.6793

Epoch 80, Loss: 0.6726

Accuracy: 97.28%

EEG & Auditory Attention Detection

Codes & Directories

EEG_Extractor_From_mat_files.m	1_1_EEG_Original_mat_Files
1_2_Specified_EEG_Dataset_Extracter.py	1_2_EEG_Original_csv_Files
1_3_EEG_Processing.py	1_3_EEG_10Ch_csv_Files
1_4_EEG_CNN_Function.py	1_4_EEG_10Ch_RawSignal
1_5_EEG_CNN_FeatureMaps.py	1_5_EEG_Dataset_Feature_Maps
2_1_Audio_Processing.py	2_1_Audio_Original_wav_Files
2_2_Audio_CNN_Function.py	2_2_Audio_dry_wav_Files
2_3_Audio_CNN_FeatureMaps.py	2_3_Audio_dry_Spectrogram
3_1_Dataset_Maker.py	2_4_Audio_Dataset_Feature_Maps
3_2_AE_Concat_Function.py	2_5_All_epochs_Dataset
3_3_AE_Concat.py	3_1_AE_Concats

Thank You

Foad Moslem

foad.moslem@gmail.com