Feature Reconstruction Based Channel Selection For Emotion Recognition Using EEG

James Ronald Msonda
Zhimin He, Ph. D.
Chuan Lu, Ph. D.

The 2021 IEEE Signal Processing in Medicine and Biology Symposium

Introduction

- Emotions are natural responses to circumstances (internal and/or external events).
- Human emotions have been predicted by observing parameters like facial expression.
- There is interest in the use of Electroencephalogram (EEG)

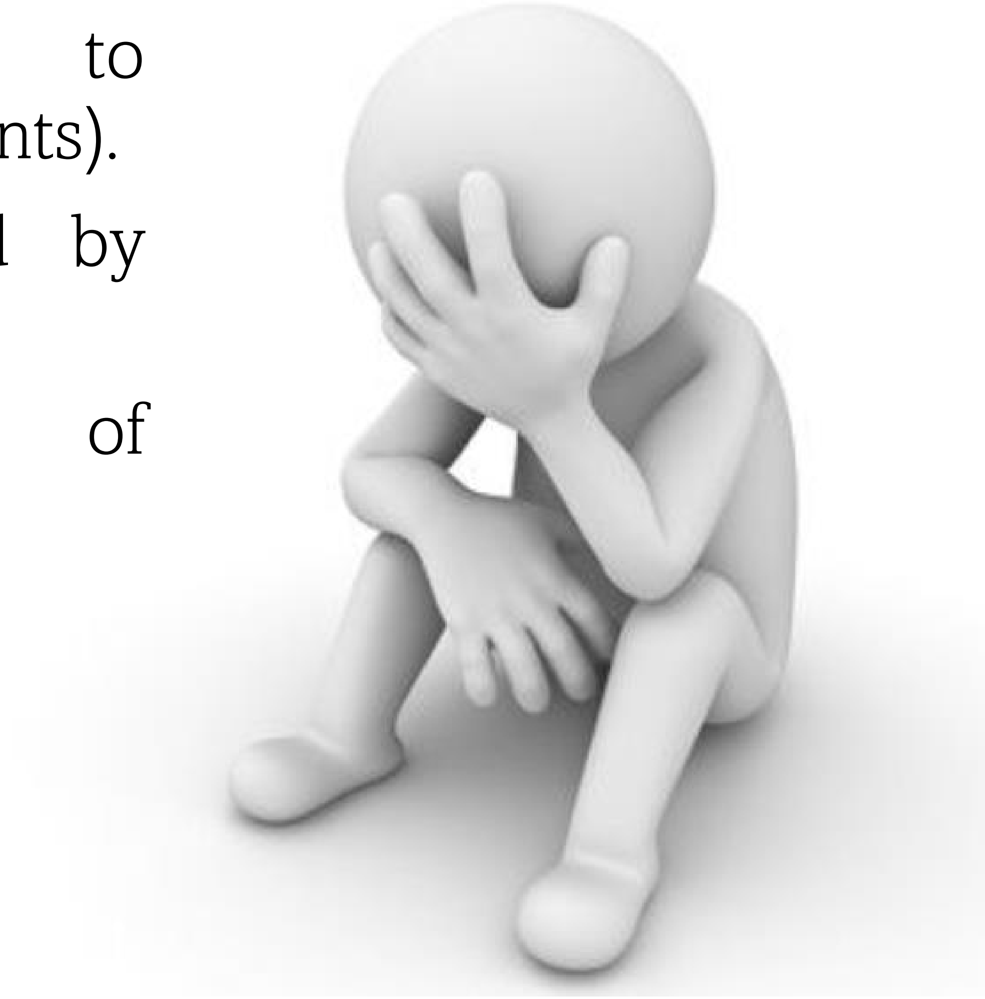


Image Credit: https://www.pinterest.co.uk/

The Problem

- Established EEG systems have up to 256 electrodes but are very expensive.
- Wearable EEG devices come in different configurations
- The question is: What is the minimum number of electrodes and where do we place them on the scalp for emotion recognition?

International 10/20 System

- Standardised method to describe EEG electrode location.
- Based on percentage of left to right or front to back distance

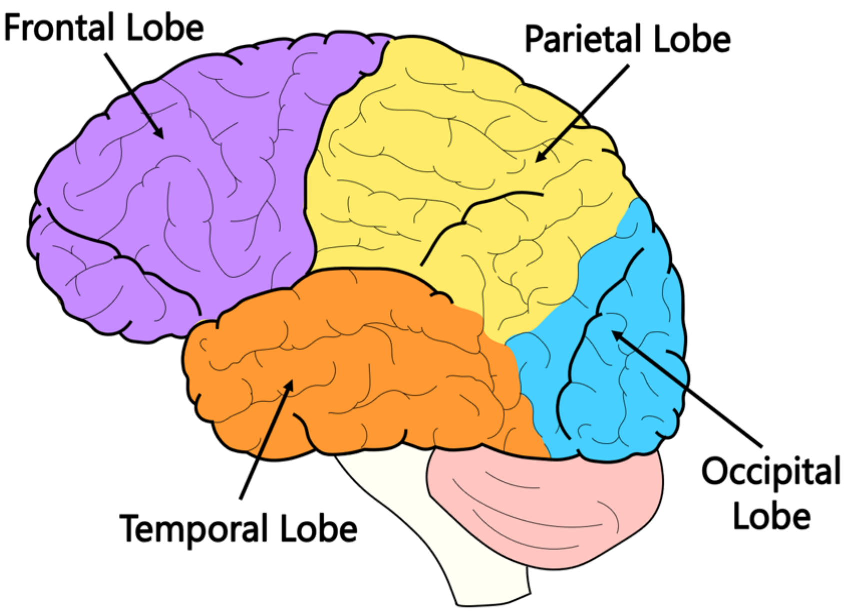
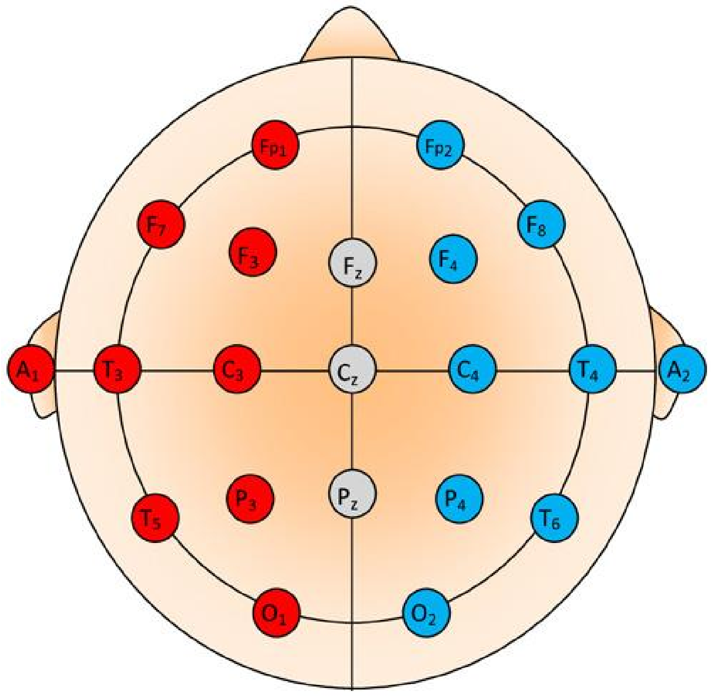


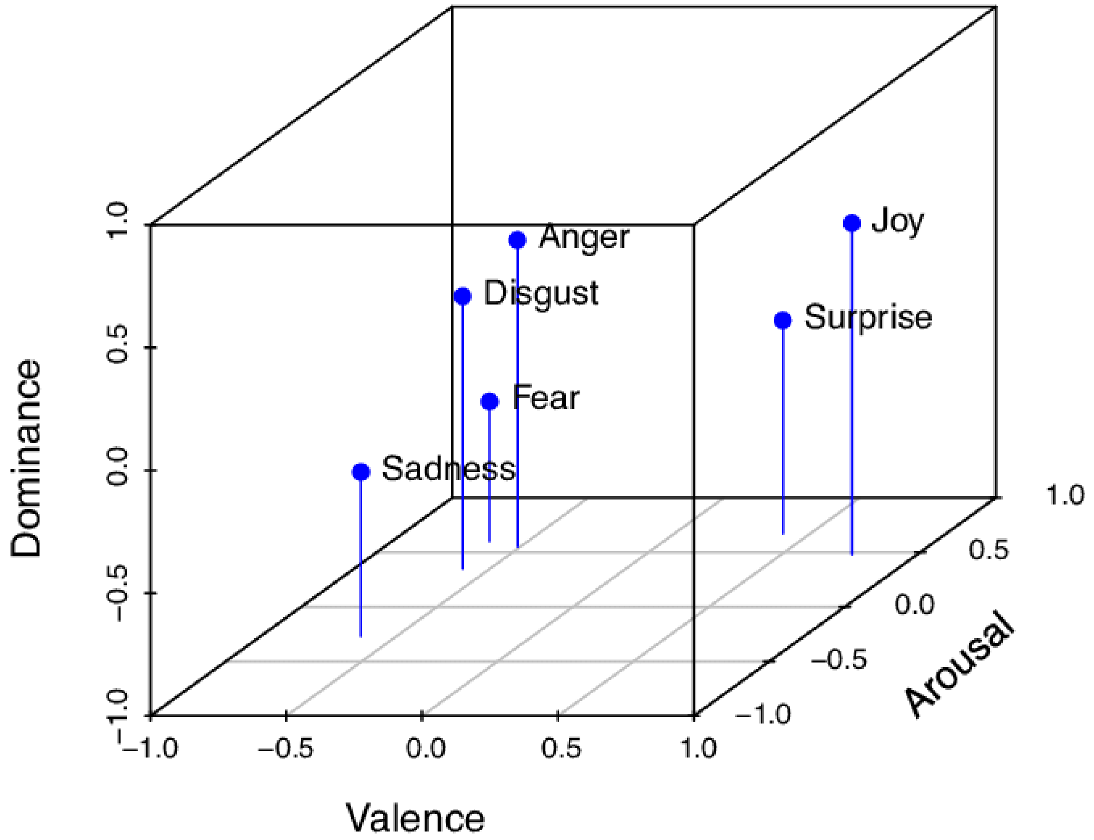
Image Credit: https://brainframe-kids.com/brain/facts-lobes.htm

- F - Frontal lobe
- T - Temporal lobe
- C - Central region
- P - Parietal lobe
- O - Occipital lobe



Data Description

- DEAP: 32 EEG channels
- DREAMER: 14 channels
- AMIGOS: 14 channels
- Self-Assessment Manikin (SAM) for self rating of emotions.

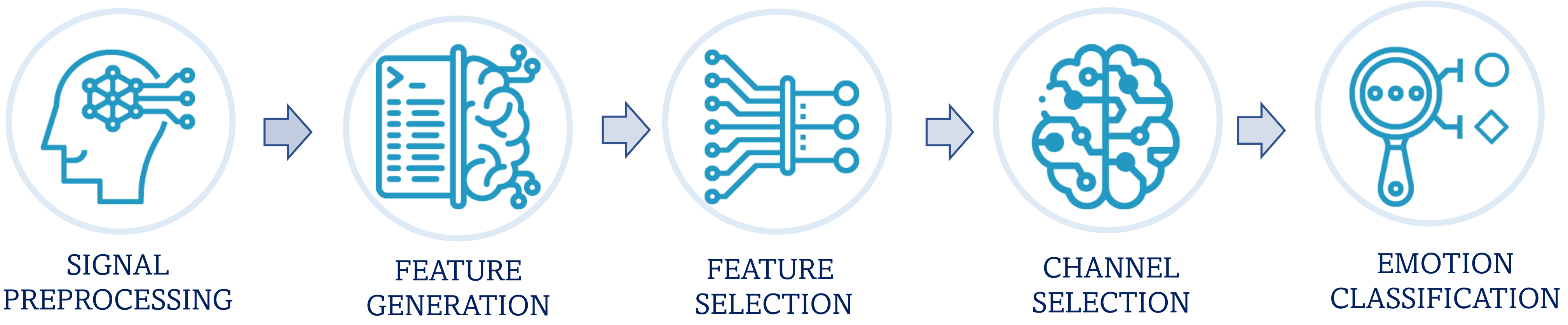


3D Representation of the SAM Scale

Image Credit: https://www.mdpi.com/2073-8994/12/1/21

Our Approach

- Goal is to identify suitable features for emotion recognition and thereafter, reconstruct the selected features using fewer channels/electrodes.
- It is a five-stage pipeline.



Preliminary Steps

Step 1

- Label binarization
- Extract final 60 secs of signal.
- Segment into 5 sec epochs.
- Channel mean subtraction

Step 2

- Full Band Signal
- EEG Band Decomposition
- Empirical Mode Decomposition
- Wavelet Decomposition

Step 3

- Fractal dimension measures
- Entropy based features
- Statistical features
- Energy… etc..

- Tested feature selection methods like ANOVA, BorutaShap, Tree Ensemble feature importance, recursive feature elimination etc.

Channel Selection Algorithm Summary

- Given output of feature selection as set of B positional features.
- To get best k out of N channels:
  1. Collect all possible channel combinations, $\mathbf{C}_{\mathbf{k}}^{\mathbf{N}}$;
  2. Calculate minimum sum of square error between each combination and corresponding features in B;
  3. REPEAT (2) for all combinations; and
  4. RETURN the combination with the least reconstruction error

Channel Selection Algorithm [1]

- Assuming we had :
  - five electrode positions - F3, T8, T7, P7, AF4
  - four features extracted from each position: power (POW), entropy (ENT), fractal dimension (FD) and zero crossing rate (ZR).
  - 2 class labels – 0, 1
- We will obtain 20 positional features and a label vector as below :

AF4_FD	AF4_ZR	LABEL
0.709	0.518	0
0.562	0.153	1
0.538	0.988	1
0.271	0.549	0

F3_POW	F3_ENT	F3_FD	F3_ZR	T8_POW	T8_ENT	T8_FD	T8_ZR	T7_POW	T7_ENT	T7_FD	T7_ZR	P7_POW	P7_ENT	P7_FD	P7_ZR	AF4_POW	AF4_ENT	AF4_FD	AF4_ZR	LABEL
0.676	0.285	0.836	0.398	0.529	0.986	0.684	0.096	0.989	0.373	0.412	0.465	0.478	0.669	0.445	0.732	0.303	0.241	0.709	0.518	0
0.385	0.467	0.833	0.858	0.287	0.157	0.146	0.351	0.737	0.986	0.898	0.491	0.156	0.552	0.459	0.589	0.818	0.685	0.563	0.153	1
0.086	0.736	0.093	0.602	0.918	0.895	0.049	0.463	0.133	0.935	0.469	0.854	0.207	0.377	0.664	0.471	0.781	0.988	0.538	0.988	1
0.759	0.732	0.478	0.053	0.937	0.547	0.851	0.332	0.541	0.131	0.555	0.037	0.763	0.908	0.601	0.347	0.179	0.085	0.271	0.549	0
0.869	0.281	0.634	0.772	0.385	0.504	0.201	0.293	0.361	0.856	0.528	0.244	0.205	0.565	0.221	0.793	0.589	0.922	0.468	0.148	0
0.893	0.253	0.403	0.727	0.698	0.006	0.408	0.501	0.301	0.009	0.779	0.979	0.349	0.482	0.573	0.801	0.929	0.908	0.606	0.168	1
0.692	0.012	0.924	0.371	0.163	0.883	0.524	0.735	0.244	0.319	0.225	0.693	0.387	0.282	0.192	0.893	0.253	0.119	0.205	0.342	0

Channel Selection Algorithm [2]

- Again, assume feature selection algorithm output as F3_POW, T8_ENT, AF4_ZR, P7_FD.
- To select the best two channels:

STEP 1: Combinations of two from five electrodes (F3, T8, T7, P7, AF4). Thus,

$$C_2^5 = \frac{5!}{2!\,(5-2)!} = 10$$

[F3, T8],	[F3, T7],	[F3, P7],	[F3, AF4],	[T8,T7],
[T8,P7],	[T8, AF4],	[T7,P7],	[T7, AF4] ,	[P7, AF4]

Channel Selection Algorithm [3]

STEP 2: For each of the combinations in step 1, we find the best substitute for each of the target features obtained after feature selection.

E.g., for candidate combination [T7, AF4] and target feature F3_POW, we compute:

$$A = \|T7\_POW - F3\_POW\|_2$$

$$B = \|AF4\_POW - F3\_POW\|_2$$

The one with a lower error is a better substitute for F3_POW

STEP 3: REPEAT step 2 to get substitutes for T8_ENT, AF4_ZR and P7_FD

Channel Selection Algorithm [4]

STEP 4: REPEAT steps 2 and 3 for all combinations to get substitutes such as:

COMBINATION	F3_POW	T8_ENT	AF4_ZR	P7_FD
[F3, T8]	F3_POW	T8_ENT	T8_ZR	F3_FD
[F3, T7]	-	-	-	-
[F3, P7]	-	-	-	-
[F3, AF4]	F3_POW	AF4_ENT	AF4_ZR	AF4_FD
[T8,T7]	-	-	-	-
[T8,P7]	-	-	-	-
[T8, AF4]	-	-	-	-
[T7,P7]	-	-	-	-
[T7, AF4]	T7_POW	AF4_ENT	T7_ZR	AF4_FD
[P7, AF4]	AF4_POW	P7_ENT	AF4_ZR	AF4_FD

Channel Selection Algorithm [5]

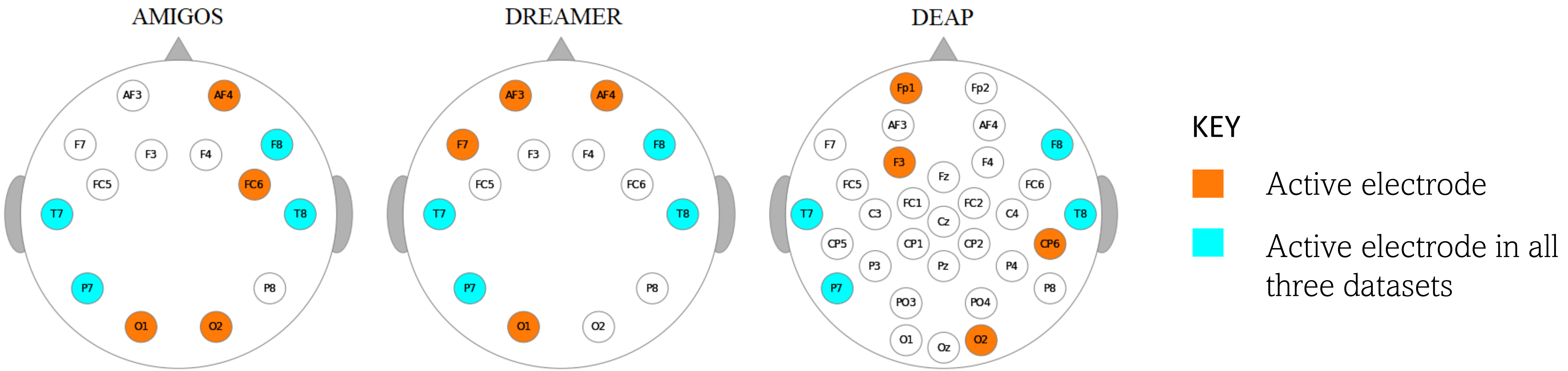
STEP 5: Calculate reconstruction error between channel combinations and the target features as the total mean squared errors from all substitute features to the targets.

COMBINATION	F3_POW	T8_ENT	AF4_ZR	P7_FD	SUM
[F3, T8]	0.000	0.009	0.002	0.009	0.0246
[F3, T7]	0.000	0.011	0.008	0.006	0.0251
[F3, P7]	0.000	0.002	0.003	0.007	0.0132
[F3, AF4]	0.002	0.001	0.005	0.002	0.0099
[T8,T7]	0.006	0.004	0.001	0.002	0.0127
[T8,P7]	0.001	0.009	0.003	0.002	0.0116
[T8, AF4]	0.011	0.009	0.008	0.005	0.0334
[T7,P7]	0.006	0.001	0.003	0.004	0.0138
[T7, AF4]	0.002	0.009	0.007	0.004	0.0217
[P7, AF4]	0.008	0.004	0.003	0.006	0.0216

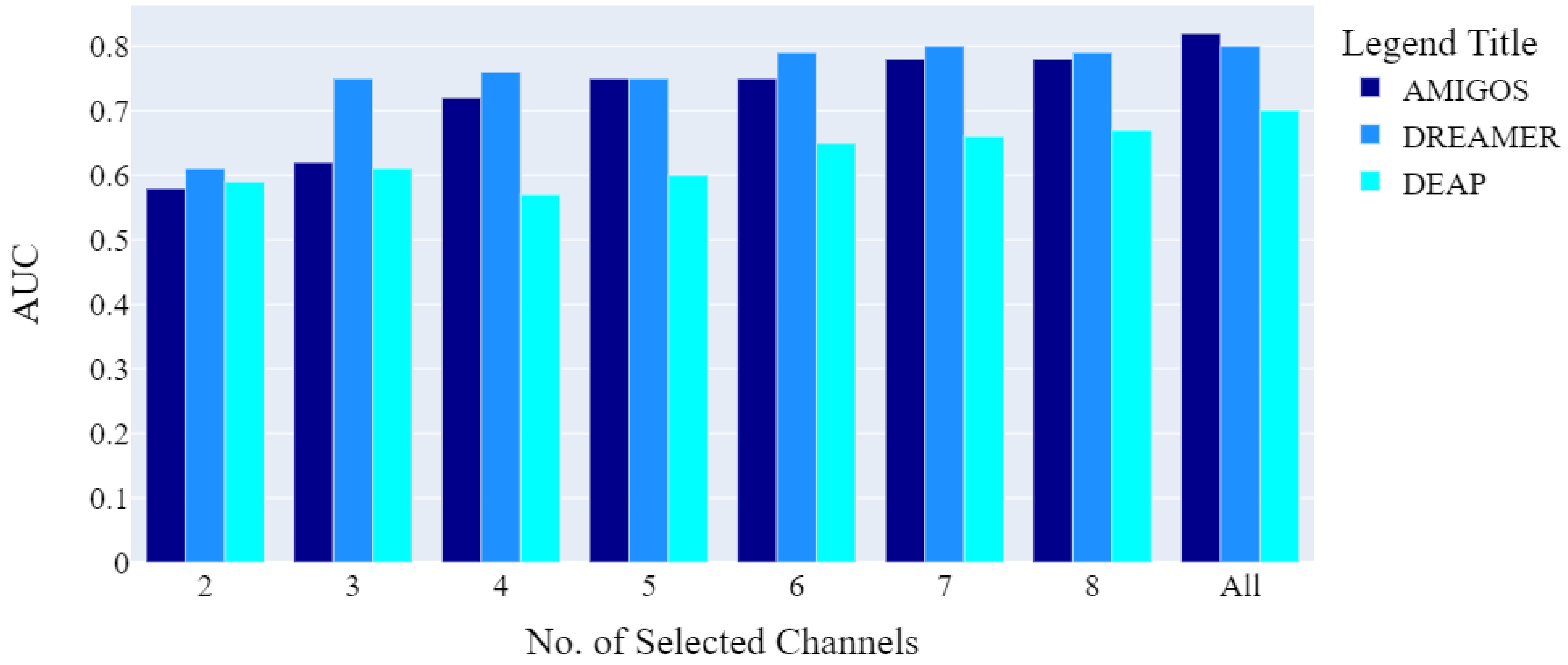
Best substitute features F3_POW, AF4_ENT, F3_ZR and AF4_FD

Results [1]

NUMBER OF CHANNELS	AMIGOS	DREAMER	DEAP
2	['FC6', 'O1']	['AF4', 'P7']	['CP6', 'T8']
3	['FC6', 'O1', 'P7']	['AF4', 'P7', 'T8']	['CP6', 'F3', 'T8']
4	['FC6', 'O1', 'P7', 'T8']	['AF4', 'F8', 'P7', 'T8']	['F3', 'Fp1', 'P7', 'T8']
5	['AF4', 'FC6', 'O1', 'P7', 'T8']	['AF4', 'F7', 'F8', 'P7', 'T8']	['F3', 'Fp1', 'O2', 'P7', 'T8']
6	['AF4', 'FC6', 'O1', 'P7', 'T7', 'T8']	['AF3', 'AF4', 'F7', 'F8', 'P7', 'T8']	['F3', 'Fp1', 'O2', 'P7', 'T7', 'T8']
7	['AF4', 'F8', 'FC6', 'O1', 'P7', 'T7', 'T8']	['AF4', 'F7', 'F8', 'FC6', 'O1', 'P7', 'T8']	['F3', 'F8', 'Fp1', 'O2', 'P7', 'T7', 'T8']
8	['AF4', 'F8', 'FC6', 'O1', 'O2', 'P7', 'T7', 'T8']	['AF3', 'AF4', 'F7', 'F8', 'O1', 'P7', 'T7', 'T8']	['CP6', 'F3', 'F8', 'Fp1', 'O2', 'P7', 'T7', 'T8']

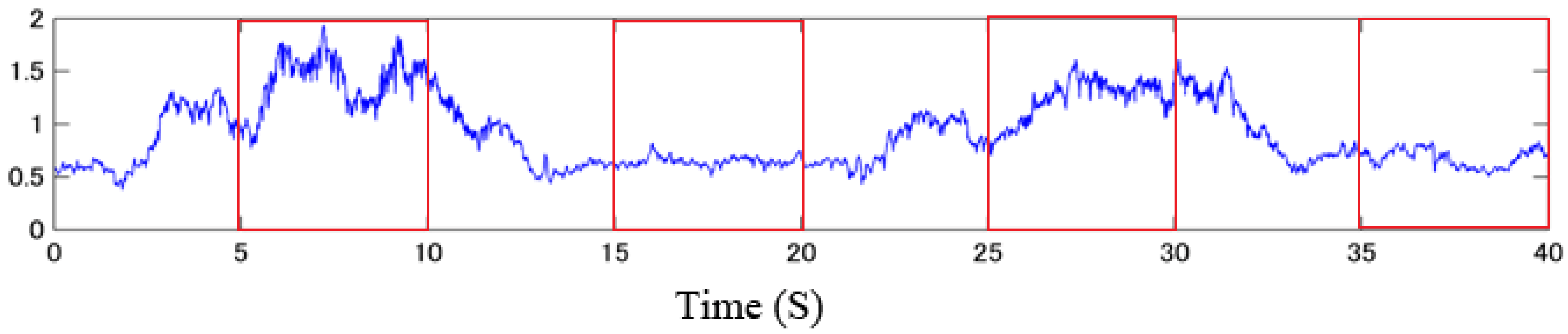


Results [2]



AUC
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
2
3
4
5
6
7
8
All
No. of Selected Channels
Legend Title
AMIGOS
DREAMER
DEAP

Weaknesses and Future Work



- Time segmentation of the signal may not be ideal
- L2 norm tends to penalize outliers
- Next step : Graph Convolution Network models

Thank You!

Contact: jam139@aber.ac.uk / jmsonda@mubas.ac.mw