Centre for Human Body Data Science
South China University of Technology

A Feature Learning Approach Based on Multimodal Human Body Data for Emotion Recognition

The 2021 IEEE Signal Processing in Medicine and Biology Symposium

Science Education and Research Center, Temple University, 1925 North 12th Street, Philadelphia, Pennsylvania, 19122, USA

Presented by: Perry Fordson

South China University of Technology, GZ, CH

December 4, 2021

Who Am I?

Ghanaian

Affective Computing

Emotion Recognition

Deep Learning

- PhD. Information and Communication Engineering.

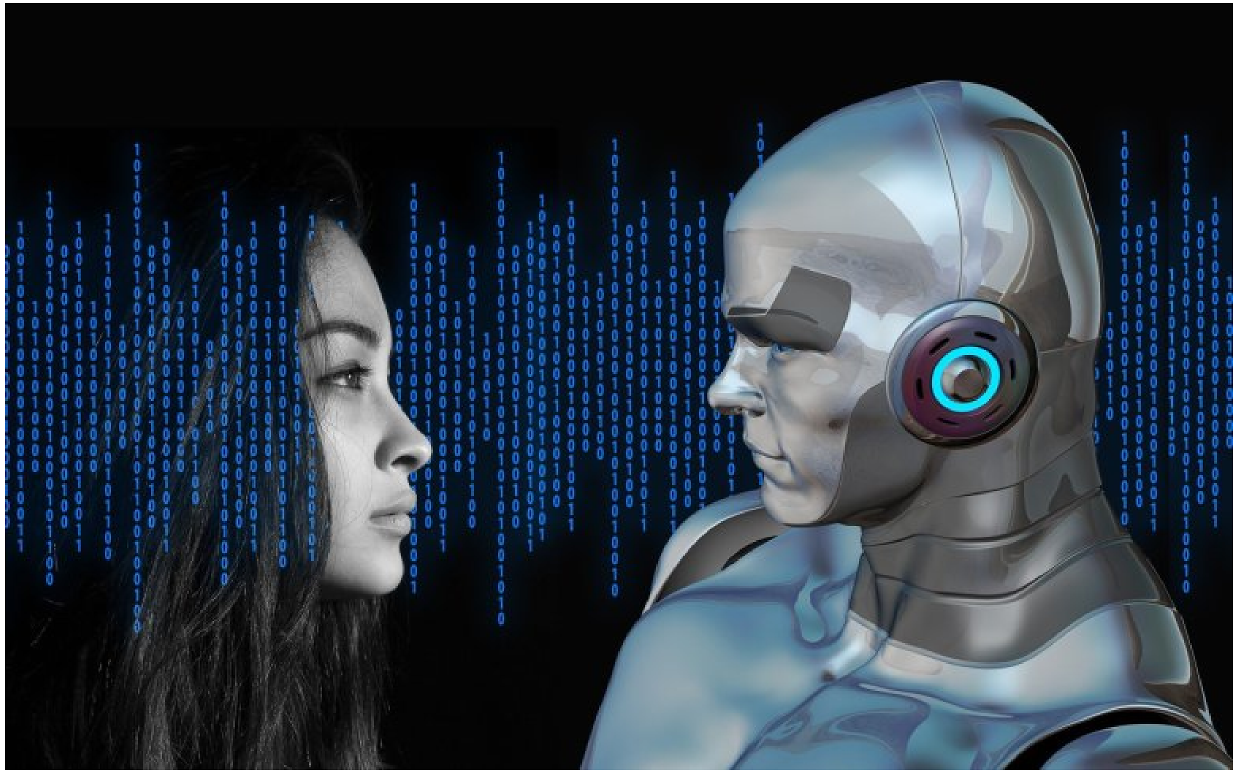
M.Eng. Information and Communication Engineering.

B.Eng. Computer Science and Technology.

Introduction

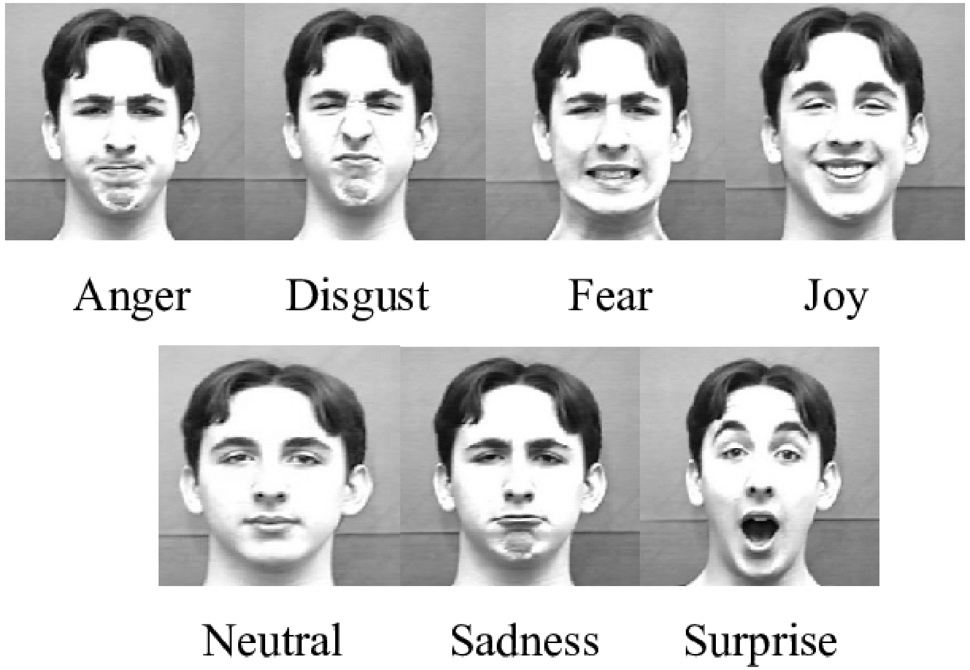
Affective Computing

- Human aspect of AI.
- Systems and devices that can recognize, interpret, process, and simulate human affect.

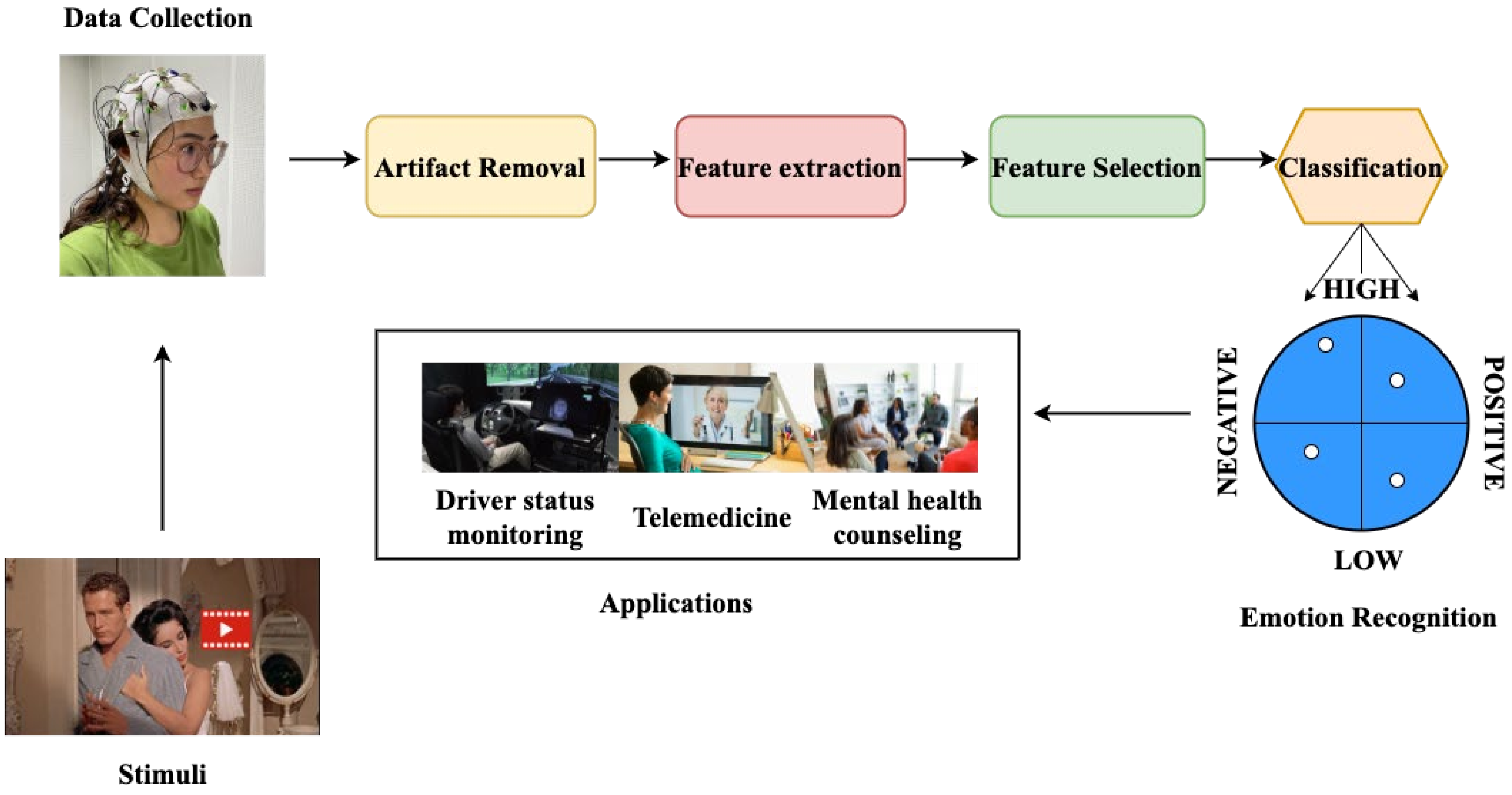


What are emotions?

- State of feeling;
- Physical and psychological changes;
- Thought and behavior.



Emotion Recognition



Fundamental modules of emotion recognition and its applications

Problem Statement

Problem Statement

- Which physiological signals are useful to distinguish between valence and arousal states?
- Can we build a model with strong representation ability to learn differences and similarities in feature nodes?
- Can we design an ensemble fusion paradigm that captures these relationships to improve multimodal performance?

Methodology

DEAP datasets

- Collected for analysis of human affective states.

Attributes

Part	**32 (16 M /16 F)**
Signals	EEG, GSR, RES, EMG (128 HZ)
Videos	40
Self-report	Valence, arousal
Rating scale	1-9

Two-Defined classes

Valence	**Arousal**	**Rating (r)**
Negative	High	$r \leq 4.5$
Positive	Low	$4.5 \leq r$

Three-Defined classes

Valence	**Arousal**	**Rating (r)**
Unpleasant	Calm	$1 \leq r \leq 3$
Neutral	Average	$4 \leq r \leq 6$
Pleasant	Excited	$7 \leq r \leq 9$

Emotional Keywords

Dimension	**Affective classes**	**Tagging**
Valence	Unpleasant Neutral Pleasant	Angry, sad, neutral, surprise, happy, amuse
Arousal	Calm Average Activated	Sad, Neutral, happy, amuse, surprise, angry

Features

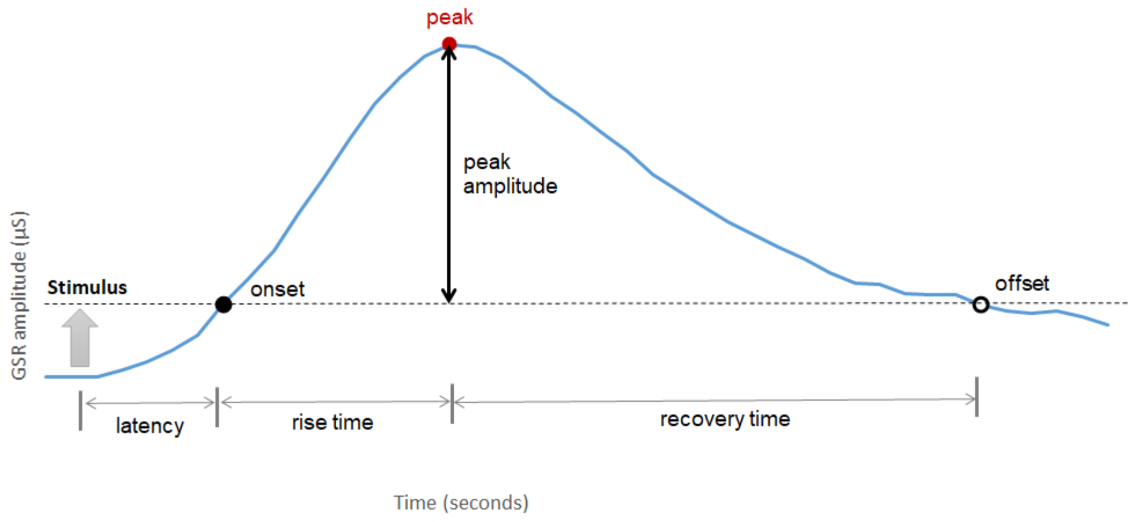
- Distinctive traits of signals.
- Require training on various models to perform classification.

Power spectral density;

Statistical moments;

Frequency;

Peaks.

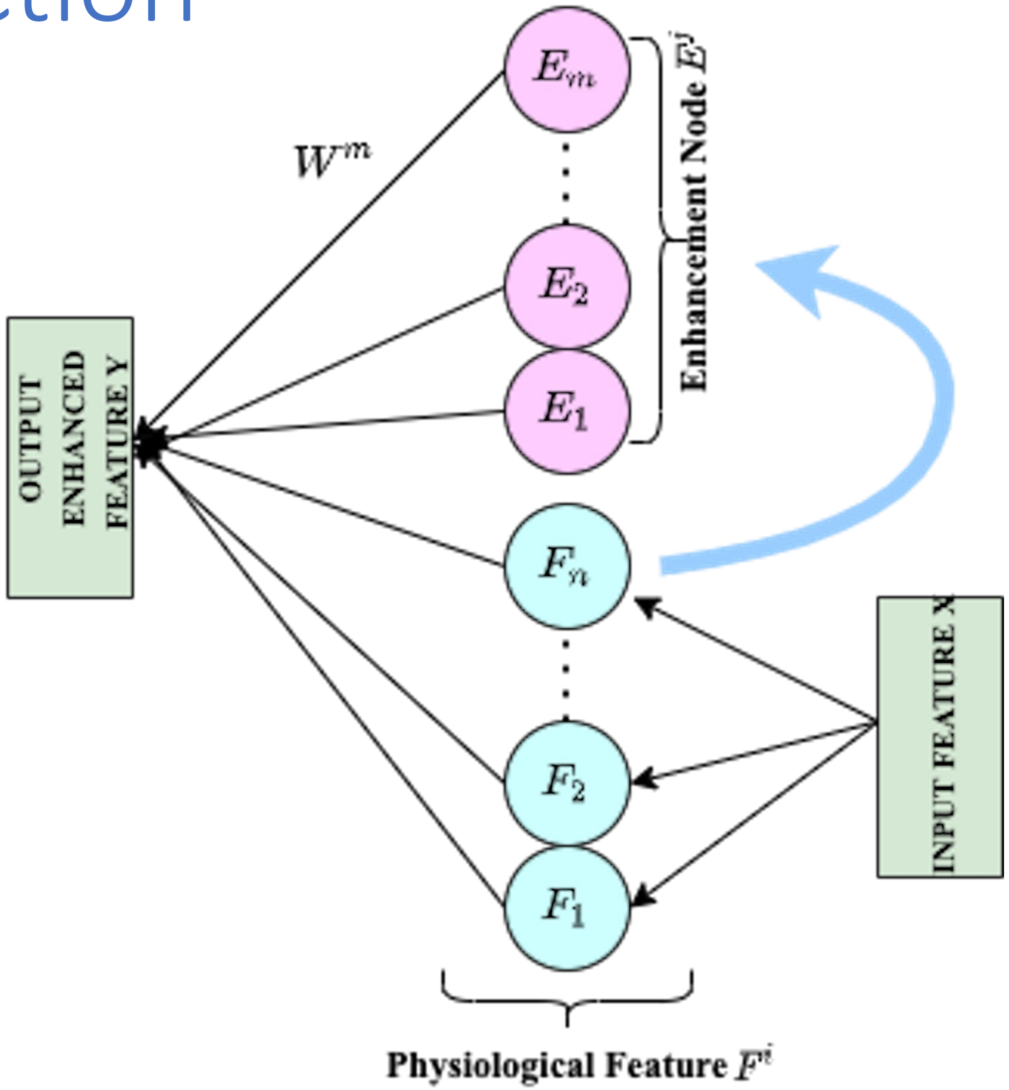


Ideal GSR signal with typically computed features

Hyper Enhanced Learning System (HELS)

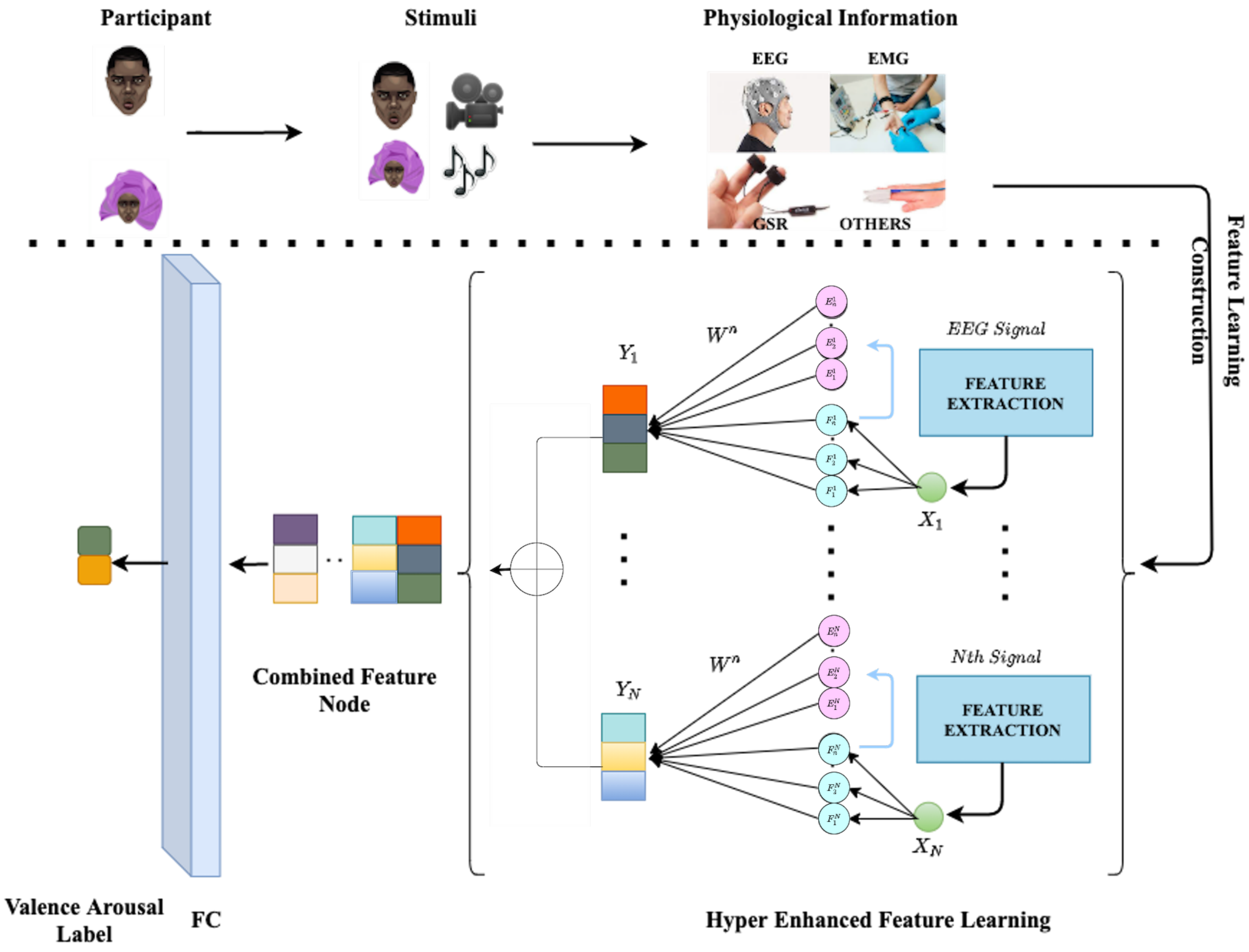
- Let's assume input feature X, projected using $\phi_i(XW_{ei} + \beta_{ei})$, as the ith mapped physiological feature;
- First group, $F^i \equiv [F_1, F_2, \dots, F_i]$;
- Similarly for enhancement node, $\zeta_j\left(F^i W_{hj} + \beta_{hj}\right)$ is E_j;
- First group, $E^j \equiv \left[E_1, E_2, \dots, E_j\right]$;
- Obtain richer features.

HELS Construction



Centre for Human Body Data Science
South China University of Technology

Overall framework



Experimental Results

Centre for Human Body Data Science
South China University of Technology

(Valence = V, Arousal = A)

(a) Two-Defined classes

Signals	V (%)	A (%)
EEG	70.1	72.2
GSR	73.3	75.6
RES	72.8	74.5
EMG	69.8	72.4
EMG+EEG	69.2	71.1
GSR+EEG	71.3	72.2
RES+EEG	68.7	70.2
GSR+EMG	68.8	70.3
RES+EMG	65.9	70.2
RES+GSR	69.4	70.9
GSR+EMG+EEG	71.3	71.4
RES+EMG+EEG	70.8	69.9
RES+GSR+EEG	71.0	71.2
RES+GSR+EMG	71.1	68.9
EEG+GSR+RES+EMG	**78.6**	**79.9**

(b) Three-Defined classes

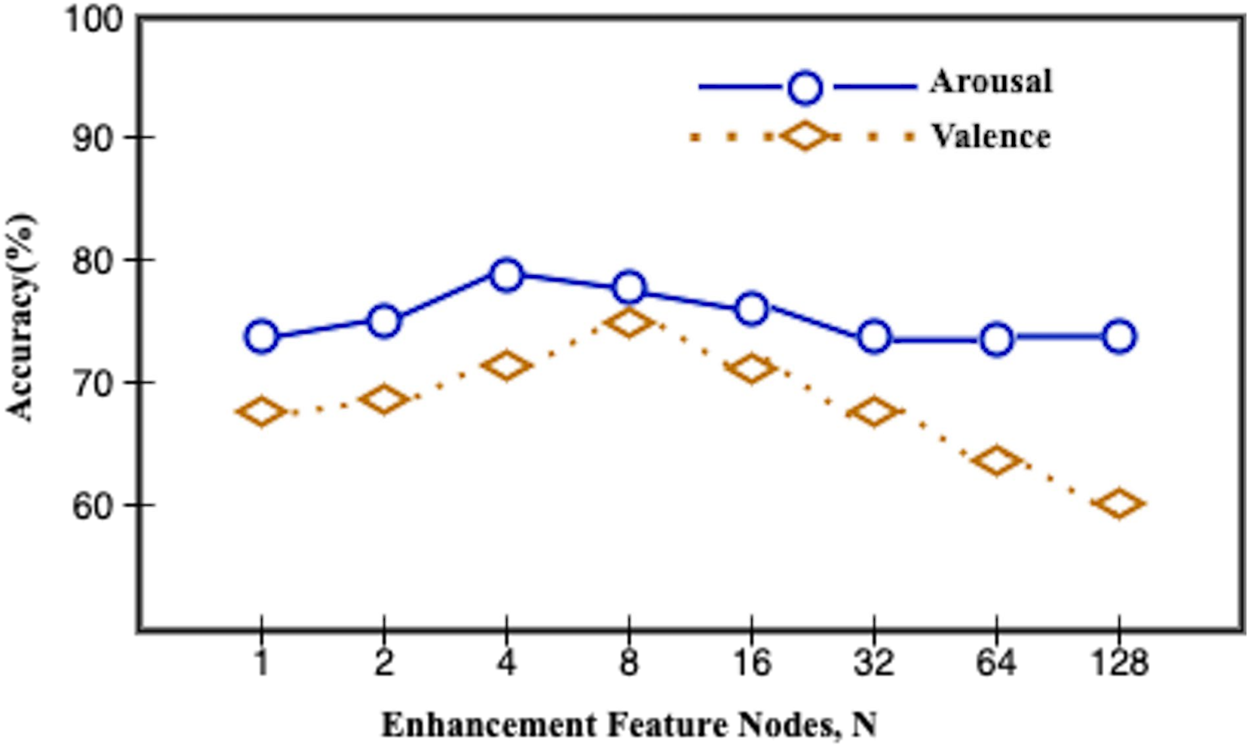
Signal	V (%)	A (%)
EEG	67.4	69.6
GSR	67.3	70.2
RES	68.4	69.7
EMG	66.6	68.0
EMG+EEG	66.2	68.5
GSR+EEG	67.4	70.1
RES+EEG	65.2	67.4
GSR+EMG	67.5	66.4
RES+EMG	67.9	65.3
RES+GSR	68.4	69.1
GSR+EMG+EEG	68.4	70.1
RES+EMG+EEG	67.1	69.8
RES+GSR+EEG	68.2	70.0
RES+GSR+EMG	69.1	69.9
EEG+GSR+RES+EMG	**69.7**	**71.9**

(c) Coded emotional keywords

Signal	V (%)	A (%)
EEG	68.1	70.6
GSR	69.2	72.9
RES	69.7	71.1
EMG	68.5	70.8
EMG+EEG	67.3	69.1
GSR+EEG	68.6	70.6
RES+EEG	66.2	68.1
GSR+EMG	69.7	68.4
RES+EMG	68.4	67.8
RES+GSR	69.2	66.6
GSR+EMG+EEG	69.2	71.3
RES+EMG+EEG	68.9	70.5
RES+GSR+EEG	69.0	71.1
RES+GSR+EMG	70.2	71.6
EEG+GSR+RES+EMG	**72.2**	**75.4**

Results

Influence of enhancement nodes on performance



Comparison with related works

Works	Valence %	Arousal %
Koelstra et al. [18]	62.7	57.0
Soleymani et al. [20]	57.0	52.3
Zhang et al. [11]	69.6	70.1
Luo et al. [3]	78.0	74.0
Ours	**78.6**	**79.9**

Summary

Summary

- Multimodal framework using physiological signals for accurate valence - arousal emotion recognition;
- Hyper-enhanced learning system that takes input as mapped features and generates random weight;
- Combining signals to validate emotional states of subjects;
- Obtained results are promising compared to related works.

Thank you and I may take your questions