BioSMART International Conference

Multi-Task CNN model for emotion recognition from EEG Brain maps

Evgenii Rudakov, Lou Laurent, Valentin Cousin, Ahmed Roshdi, Regis Fournier, Amine Nait-ali, Taha Beyrouthy, Samer Al Kork

UPEC UNIVERSITÉ PARIS-EST CRÉTEIL | FACULTÉ DES SCIENCES ET TECHNOLOGIE

Presentation plan

- Introduction
- Dataset description
- Data processing and feature extraction
- Model
- Discussion
- Conclusion

Introduction (1)

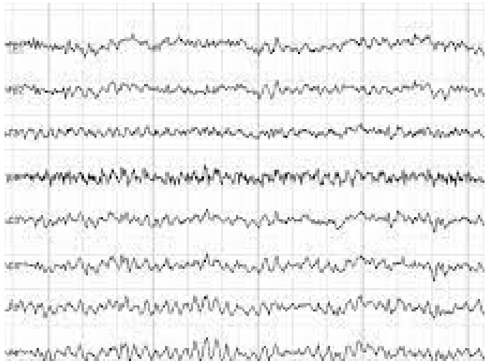
- Emotion is a mental and physiological state experienced as the reaction of a person to everyday external and internal events
- Emotions play a major role in human interactions
- Recognizing emotions will help researchers to understand better individuals
- Many domains can be considered for emotion recognition:

Non physiological	Physiological
Face	Heart rate
Speech	Functionnal MRI
Posture	EEG

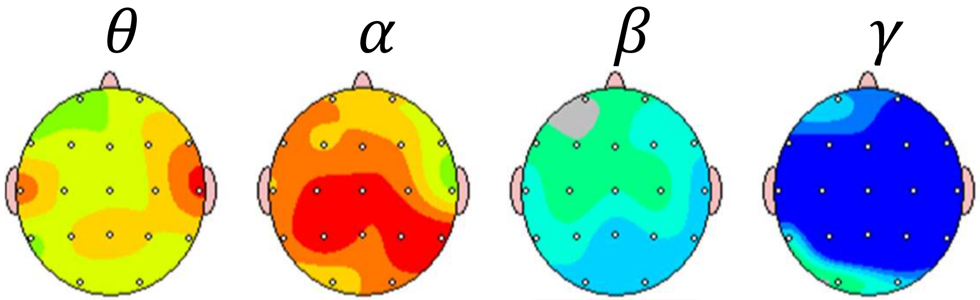
UPEC UNIVERSITÉ PARIS-EST CRÉTEIL FACULTÉ DES SCIENCES ET TECHNOLOGIE

Introduction (2)

- EEG signals can objectively reflect a variety of emotions and can be a reliable way to detect real emotions.
- Raw EEG signals are susceptible to noise and difficult to analyze, the problem arises of generation of highly-discriminative robust features
- EEG Brain maps are a spatial representation of features extracted from EEG signals



Raw EEG signals



Brain maps representation for visual analysis

DEAP dataset

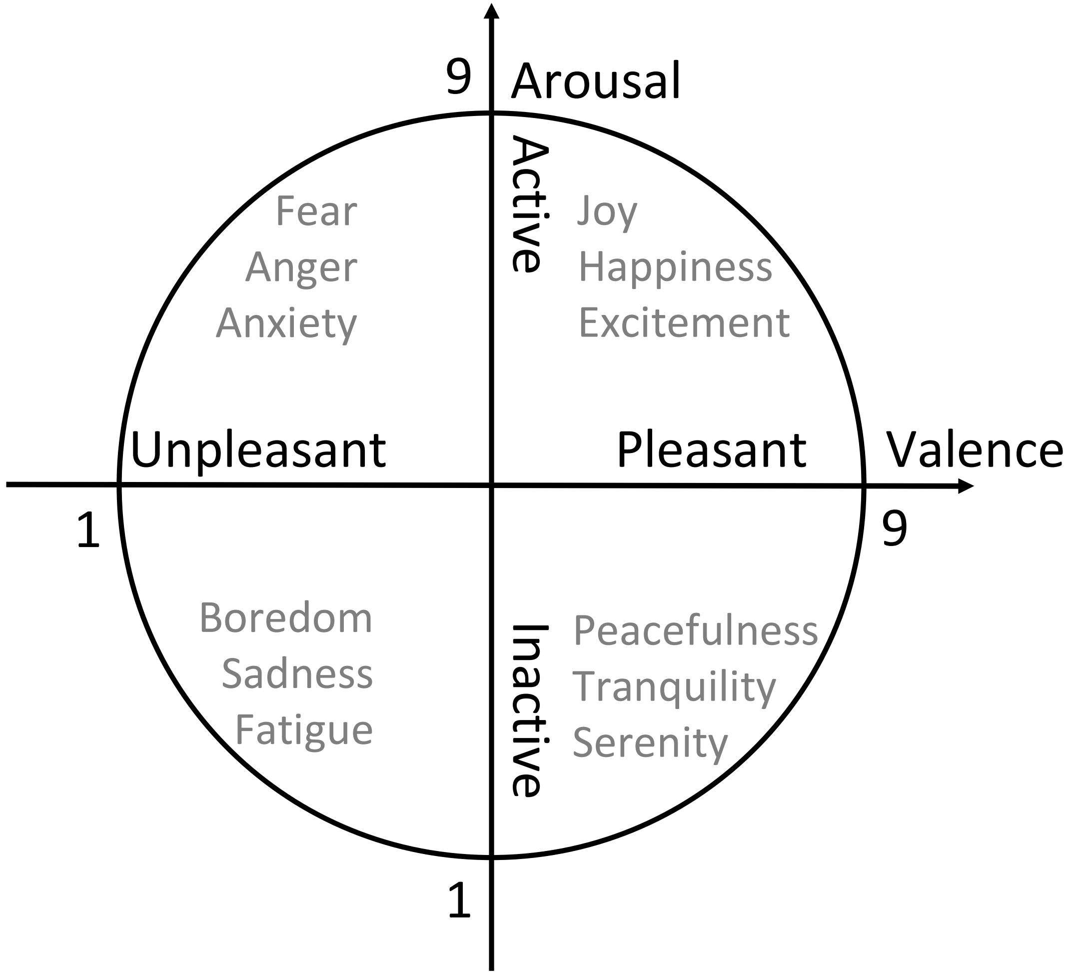
- 32 participants
- 40 stimuli (music video clips)
- Each trial is ranked in valence and arousal

DEAP dataset participant with connected sensors and in front of a PC with translated stimuli [1]

[1] Koelstra, S., C. Muhl, M. Soleymani, Jong-Seok Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, et I. Patras. DEAP: A Database for Emotion Analysis; Using Physiological Signals . IEEE Transactions on Affective Computing 3, no 1 (janvier 2012): 18-31. https://doi.org/10.1109/TAFFC.2011.15.

DEAP dataset

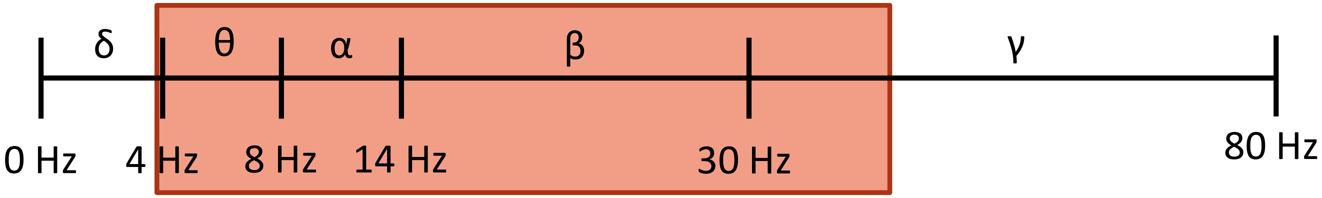
- Arousal spectrum refers to the activity level
- Valence spectrum refers to the pleasure level



Brain map creation (1)

4 brainwaves bands used:

- θ, 4 – 8 Hz
- α, 8 – 14 Hz
- β, 14 – 30 Hz
- γ, 30 – 45 Hz

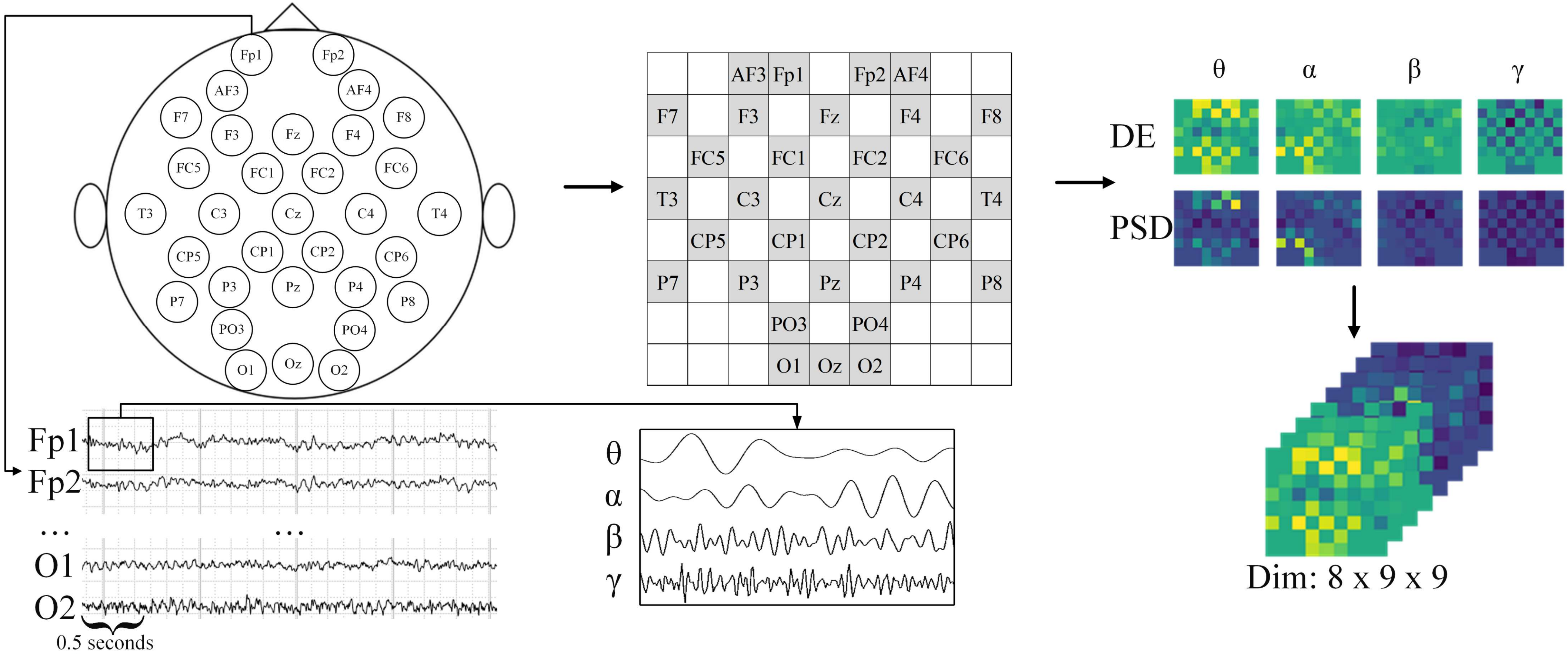


Frequency range for analysis

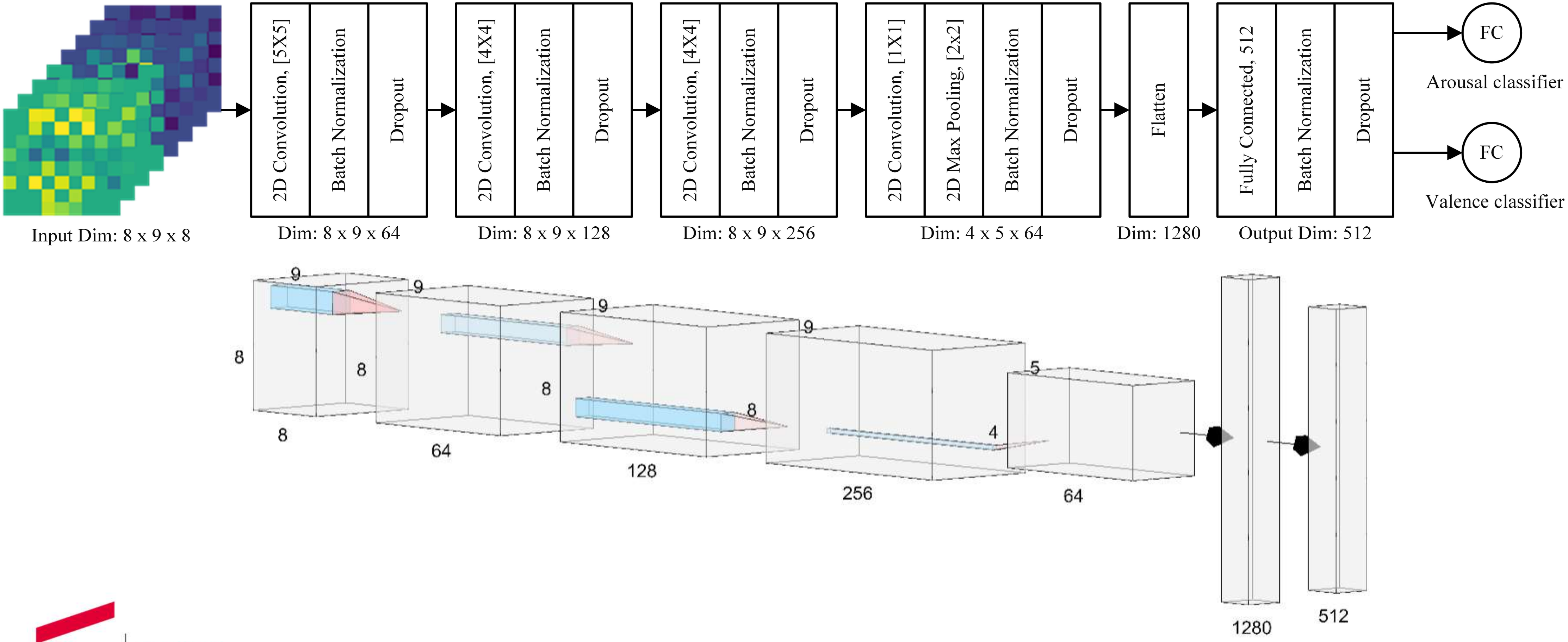
2 features extracted:

- Power Spectral Density (PSD)
- Differential Entropy (DE)

Brain map creation (2)



Model



UPEC
UNIVERSITÉ PARIS-EST CRÉTEIL
FACULTÉ DES SCIENCES ET TECHNOLOGIE

Training process (1)

- Due to the high inter-subject variability, models performance degrade when the whole subject set is used for training instead of subject-specific set.
- Common practice is to have unique models to all subject in the dataset

Model	Arousal, Accuracy	Valence, Accuracy
Reproduced SOTA, training on the specific subject only [2], **64 separate models**	94.21%	93.93%
Training on the whole subject dataset, **2 models**	90.15% (–4.06%)	88.92% (–5.01%)

[2] Shen, F., Dai, G., Lin, G. et al. EEG-based emotion recognition using 4D convolutional recurrent neural network. Cogn Neurodyn 14, 815–828 (2020). https://doi.org/10.1007/s11571-020-09634-1

Training process (2)

To handle high inter-subject variability we apply several neural network regularizations:

- Dropout
- Batch Normalizations
- Learning rate decay
- Balancing the training set

Training process (3)

In this study, we are dealing with two tasks:

- classification of the valence level
- classification of the arousal level

The final loss function is a weighted sum of the original loss functions of the tasks, which can then be used to solve the problem using the back-propagation method. It can be represented as:

$$\mathcal{L}(\theta) = w_v \mathcal{L}_v(\theta) + w_a \mathcal{L}_a(\theta)$$

Where $\mathcal{L}_v$ and $\mathcal{L}_a$ are the loss functions for valence and arousal tasks respectively and w_v and w_a are weighting coefficients for them. And $\theta \in R^n$ are training model parameters.

UPEC UNIVERSITÉ PARIS-EST CRÉTEIL | FACULTÉ DES SCIENCES ET TECHNOLOGIE

Training process (4)

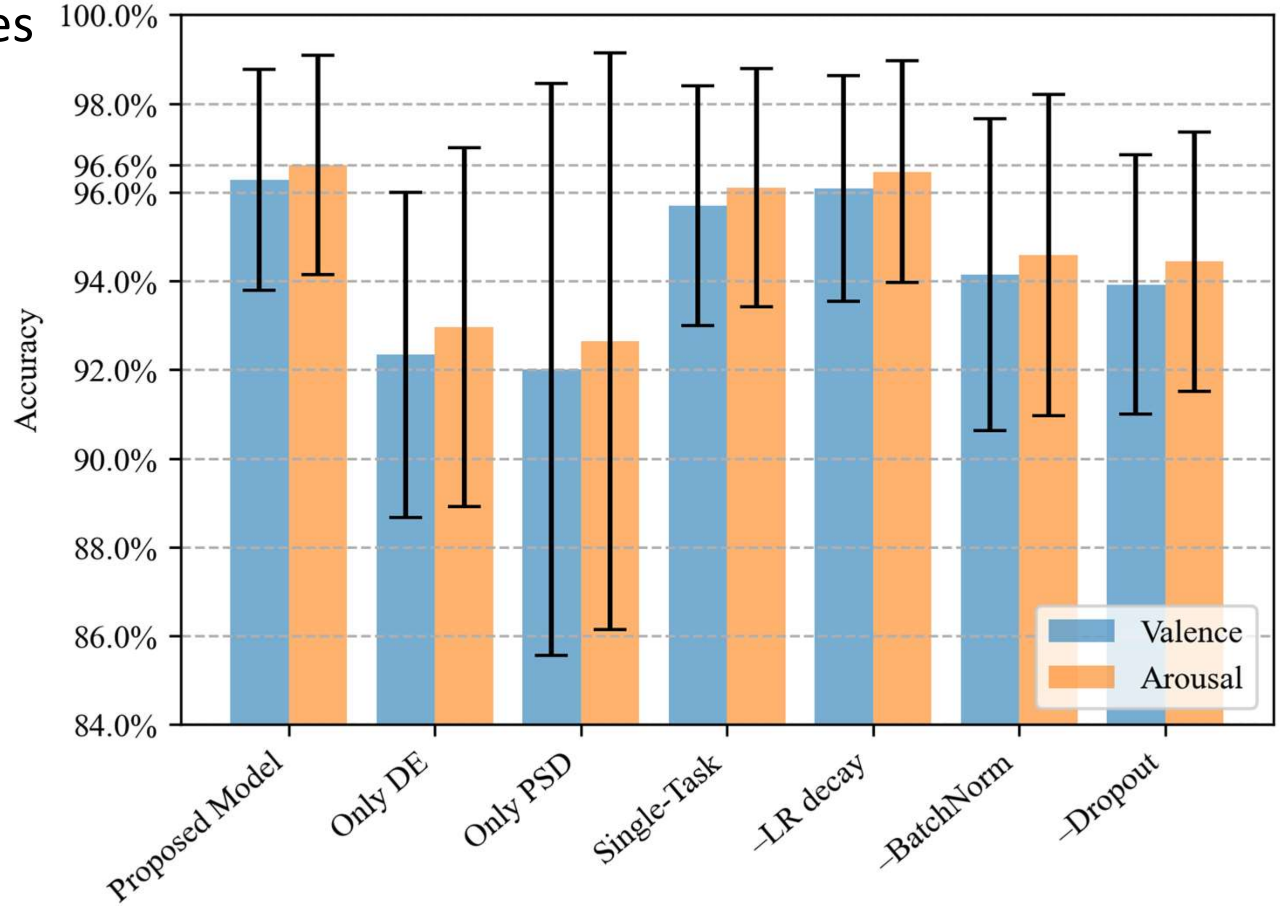
- The performance definition of the final model is made by a cross-validation method with five folds
- Model is trained to convergence with following training hyper-parameters:

Parameter	Value
Learning rate	0.001
Learning rate decay in plateau	1/2
Learning rate decay patience	5 epochs
Dropout rate	0.2
Batch size	64 samples
Valence loss weight coefficient	1
Arousal loss weight coefficient	1

Results (1)

We conducted the ablation studies to evaluate the impact of each modification on the performance of the model:

- PSD and DE as combined features for better representation
- Multi-task learning
- Dropout
- Batch Normalizations
- Learning rate decay



UPEC UNIVERSITÉ PARIS-EST CRÉTEIL | FACULTÉ DES SCIENCES ET TECHNOLOGIE

Results (2)

Nos.	Method	DEAP-valence	DEAP-arousal
1	Multi-column CNN [11]	90.01%	90.65%
2	SAE-LSTM [12]	81.10%	74.38%
3	4D-CRNN [8]	94.22%	94.58%
4	FBCCNN [13]	90.26%	88.90%
5	MT-CNN (Proposed method)	96.28%	96.62%

[8] Shen, Fangyao, Guojun Dai, Guang Lin, Jianhai Zhang, Wanzeng Kong, et Hong Zeng. EEG-Based Emotion Recognition Using 4D Convolutional Recurrent Neural Network . Cognitive Neurodynamics 14, no 6 (decembre 2020): 815-28. https://doi.org/10.1007/s11571-020-09634-1.

[11] Yang, Heekyung, Jongdae Han, et Kyungha Min. A Multi-Column CNN Model for Emotion Recognition from EEG Signals . Sensors 19, no 21 (31 octobre 2019): 4736. https://doi.org/10.3390/s19214736.

[12] Xing, Xiaofen, Zhenqi Li, Tianyuan Xu, Lin Shu, Bin Hu, et Xiangmin Xu. SAE+LSTM: A New Framework for Emotion Recognition From Multi-Channel EEG . Frontiers in Neurorobotics 13 (12 juin 2019): 37. https://doi.org/10.3389/fnbot.2019.00037.

[13] Pan, Bo, et Wei Zheng. Emotion Recognition Based on EEG Using Generative Adversarial Nets and Convolutional Neural Network. Computational and Mathematical Methods in Medicine. 2021 (11 octobre 2021): 1-11. https://doi.org/10.1155/2021/2520394.

Discussion

Area of future improving:

- Subject-independent and cross-dataset approaches
- Time-domain information capturing
- Multi-modal emotion recognition combining physiological and non-physiological data
- Generation and combination of new highly-discriminative features

Conclusion

Contribution:

- We propose a Multi-task CNN based on the Brain map EEG representation combining PSD and DE features
- We utilize the regularization methods which allow us to train the model on the broader subsets of subjects
- We utilize multi-task approach to come up with a single model for emotion recognition and increase the model domain knowledge
- We conduct extensive experiments on the benchmark dataset and the experimental results show that our model consistently outperforms all state-of-the-art models.

Thank you for your attention!

Evgenii Rudakov, Lou Laurent, Valentin Cousin, Ahmed Roshdi, Regis Fournier, Amine Nait-ali, Taha Beyrouthy, Samer Al Kork

UPEC UNIVERSITÉ PARIS-EST CRÉTEIL | FACULTÉ DES SCIENCES ET TECHNOLOGIE