

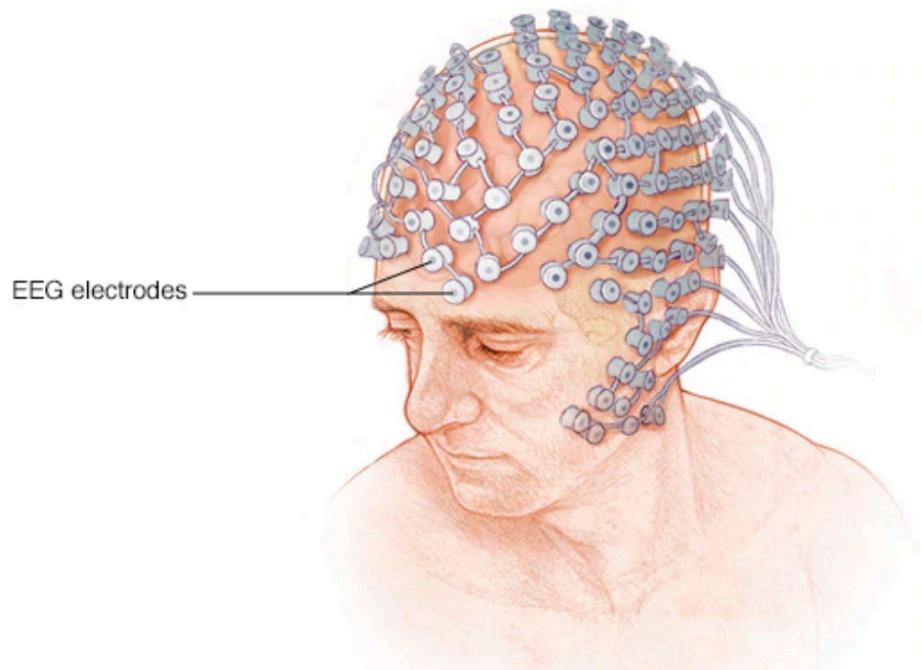
Deep Learning for EEG Classification

Ziying Feng
December 13, 2019

Outline

- **Introduction**
- **Deep Learning Model Design**
- **Preprocessing and Optimization**
- **Ablation Study and Test**
- **Failure Case Analysis**
- **Summary**

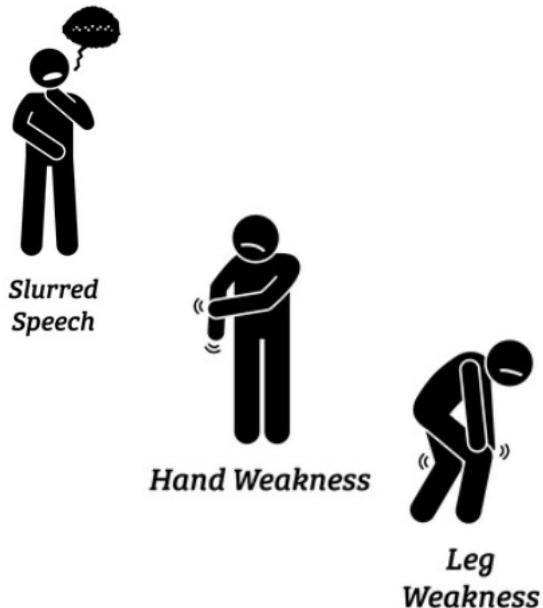
Introduction – EEG (Electroencephalogram)



- **EEG (Electroencephalogram)**
- **Record electrical activity of the brain**
- **Non-invasive, placed along the scalp**

Introduction – Decode Movement Info From EEG

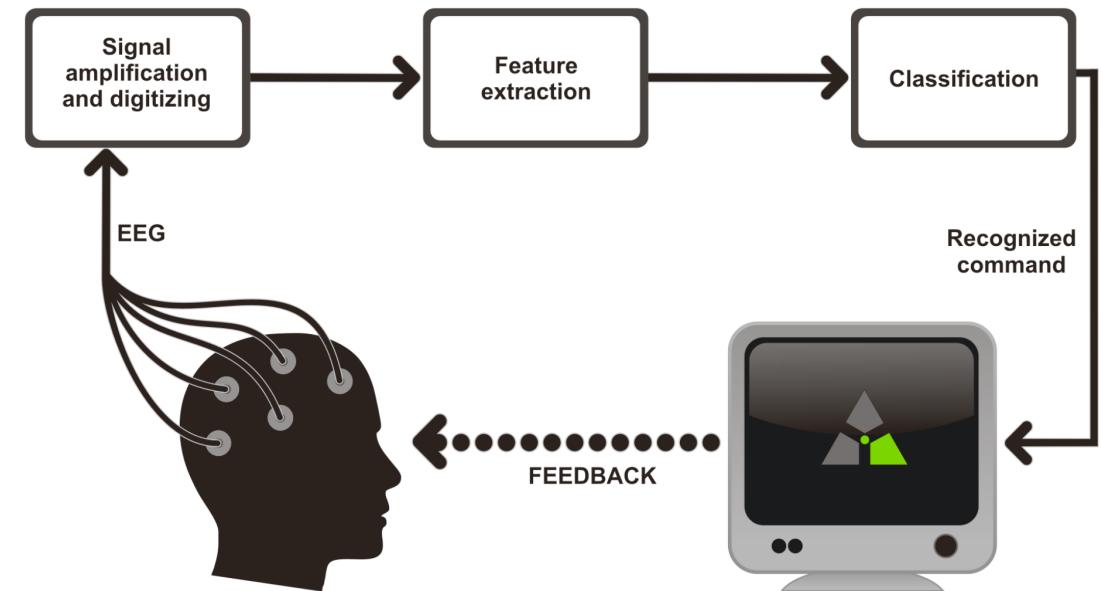
Why is it important ?



ALS patients

<http://www.hawking.org.uk/>

PLoS ONE 2011 6(6): e20674



Left



Right



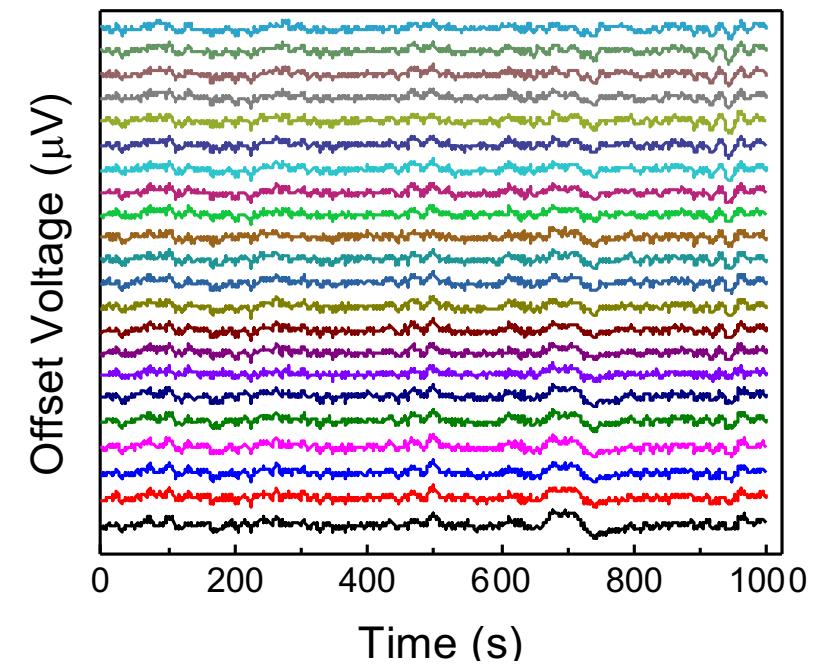
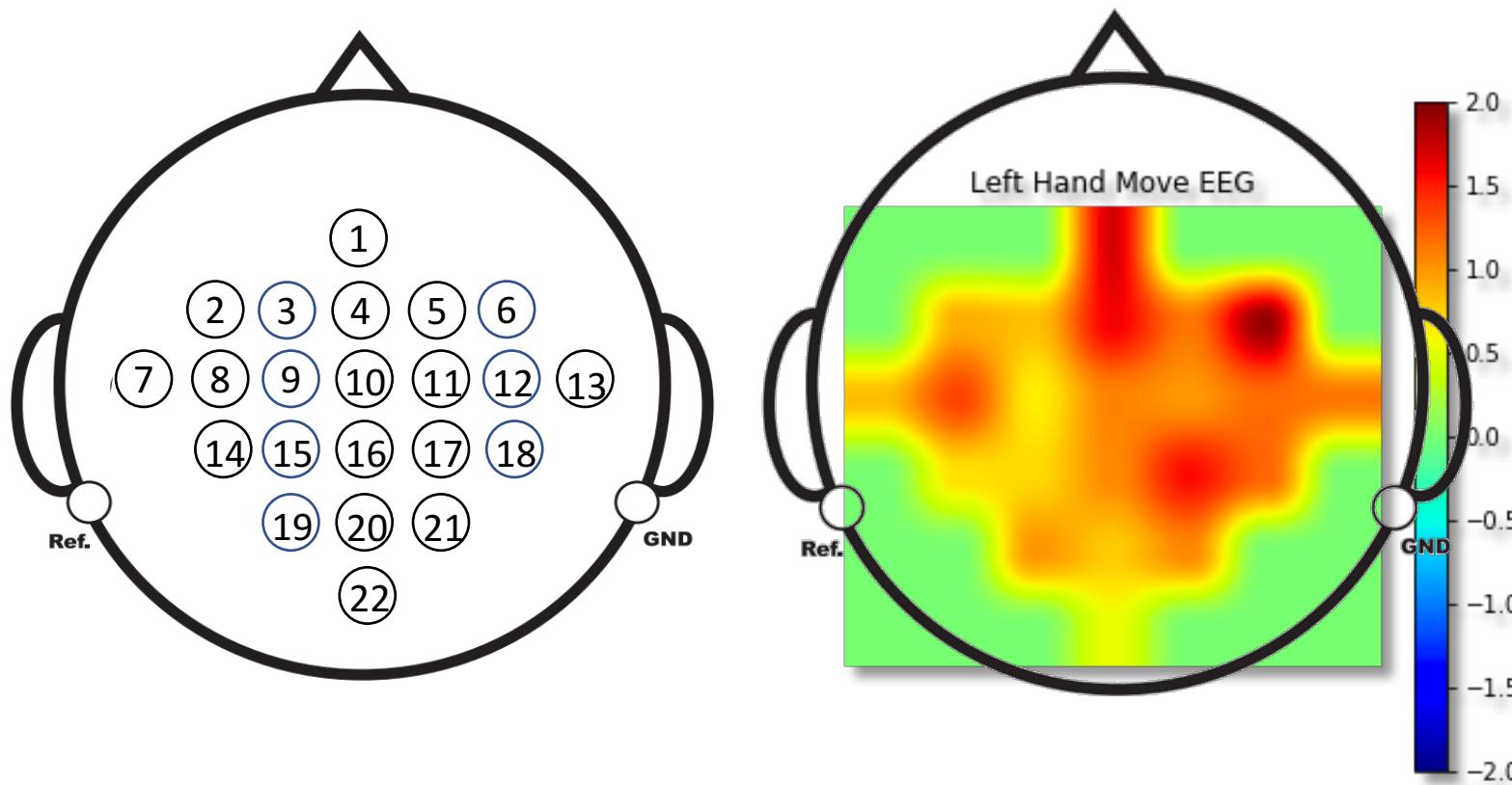
Feet



Tongue

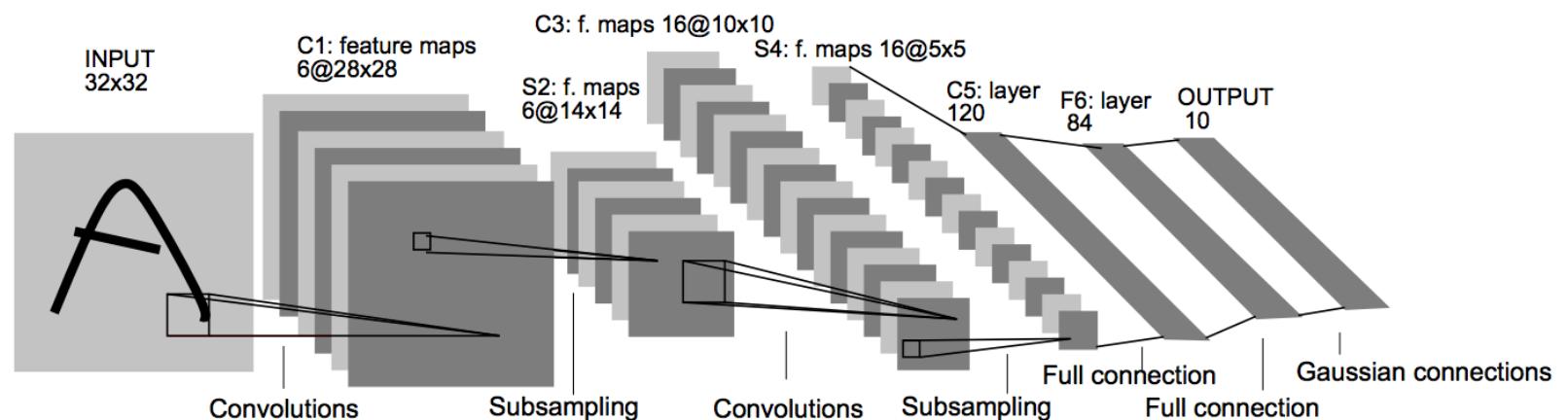
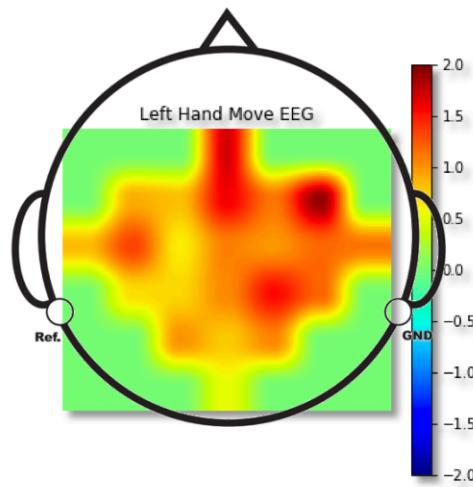
Introduction – EEG Training Data Example

EEG data set for this project



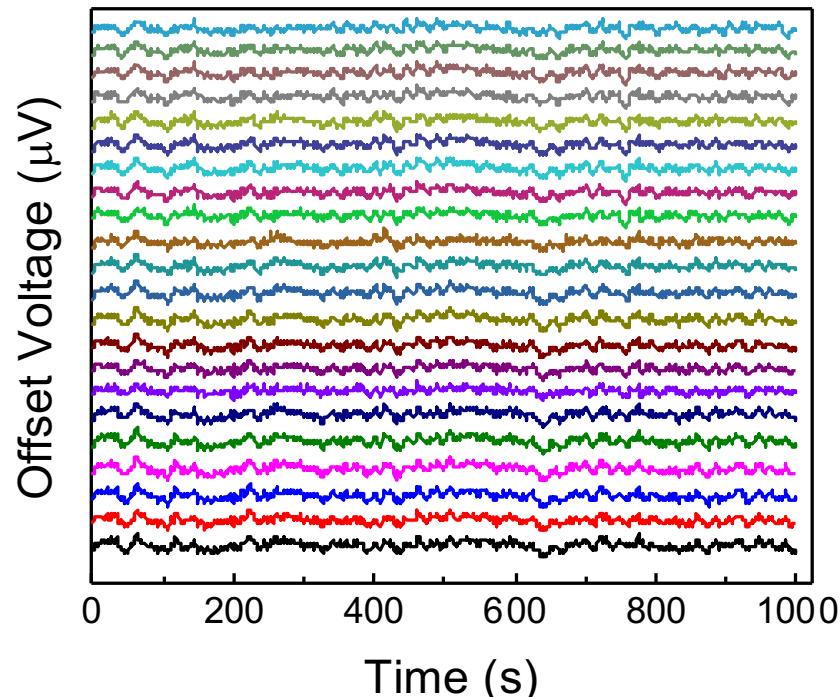
Why Convolutional Neural Network ?

- Too many features: $22 \text{ electrodes} \times 1000 \text{ time bins} = 22000$
- CNN (Convolutional Neural Network) reduces the parameter numbers by applying filter

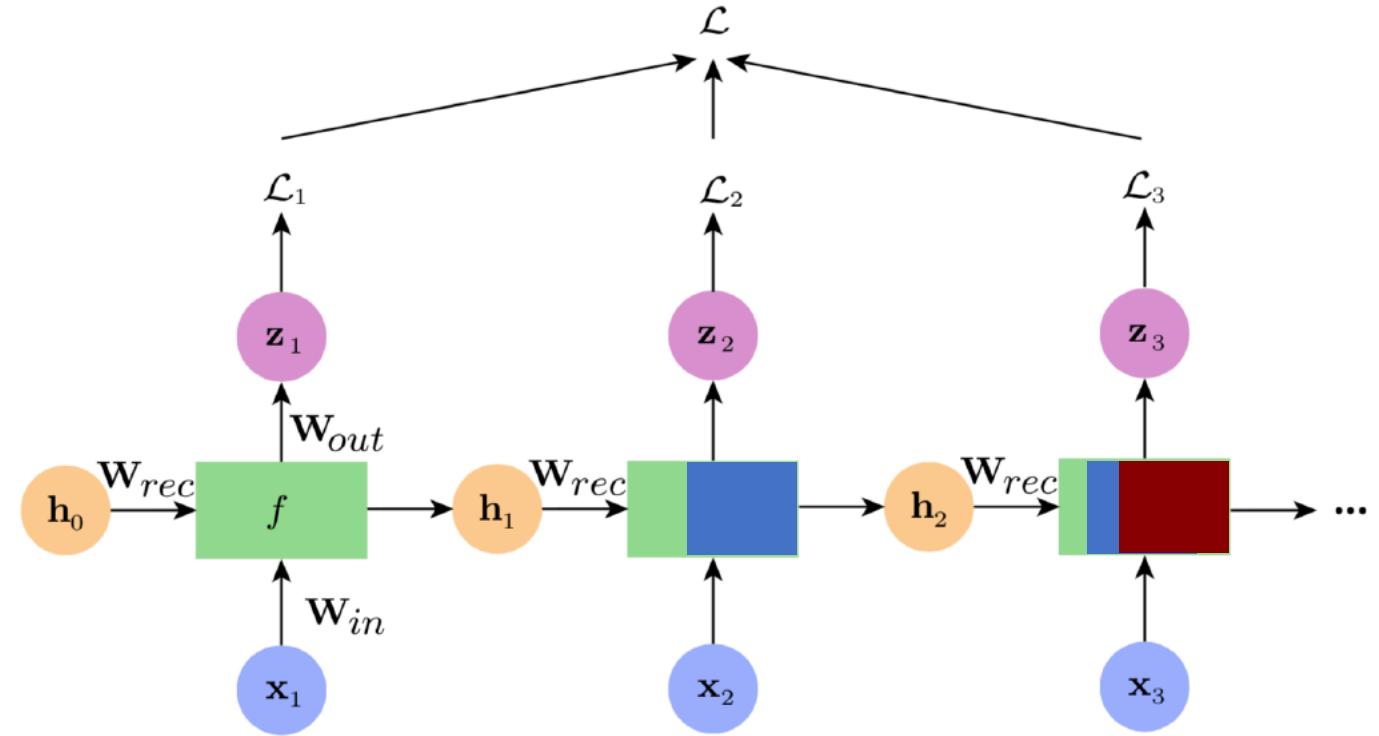


How to Deal With Temporal Feature ?

- **RNN (Recurrent Neural Network)**
- **Has hidden state from the previous step – Memory**



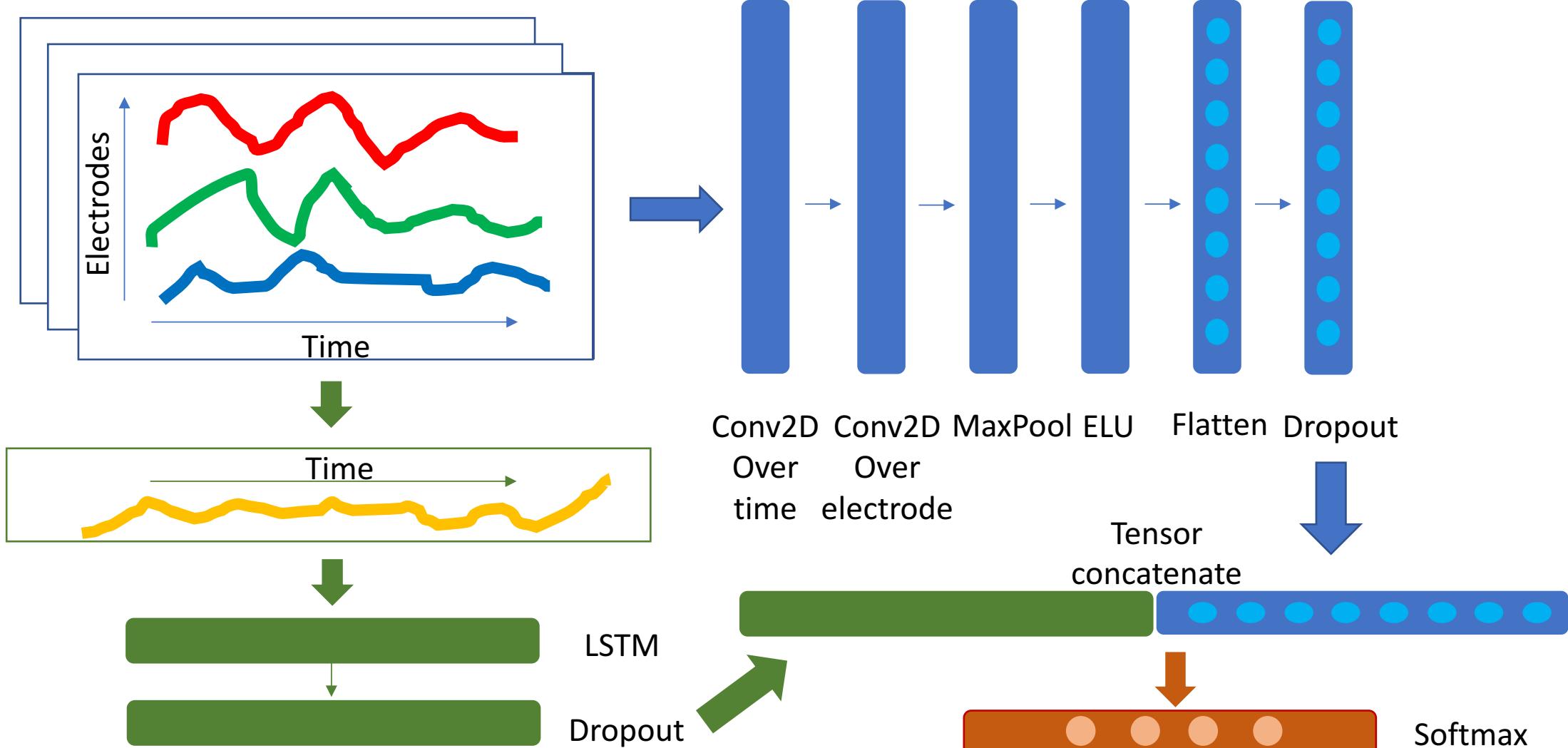
Sample Phrases	I am happy	I am not happy
Without RNN	Positive	Positive
With RNN	Positive	Negative



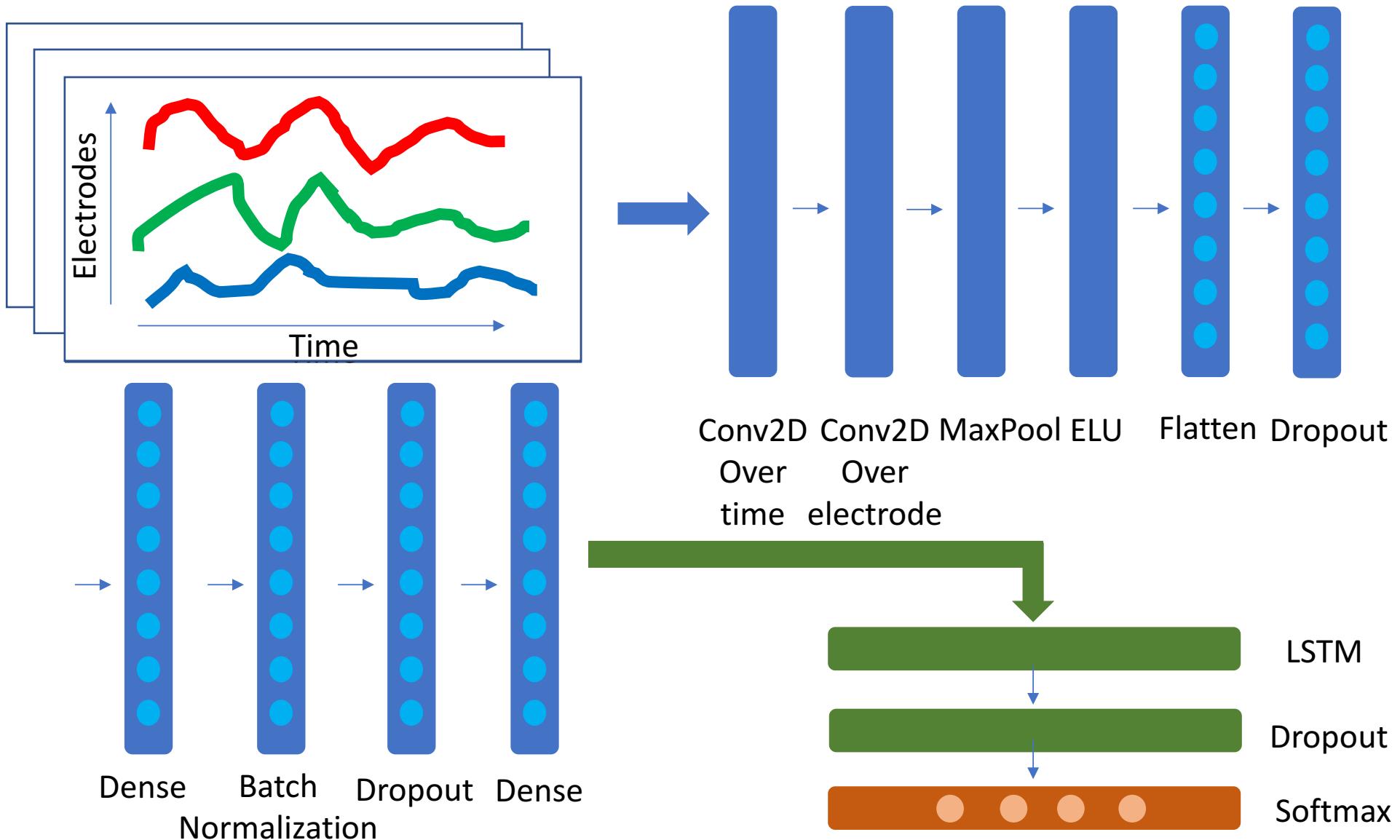
Outline

- Introduction
- Deep Learning Model Design
- Preprocessing and Optimization
- Ablation Study and Test
- Failure Case Analysis
- Summary

Model – CRNN Parallel



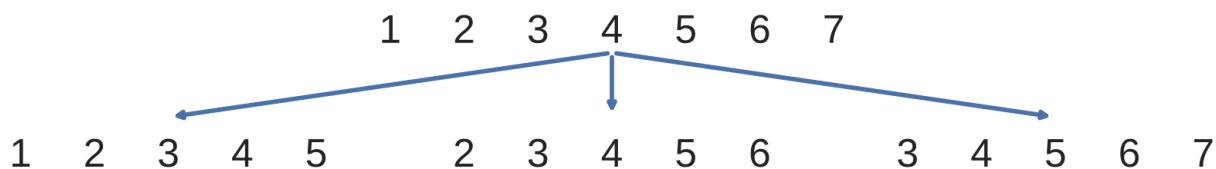
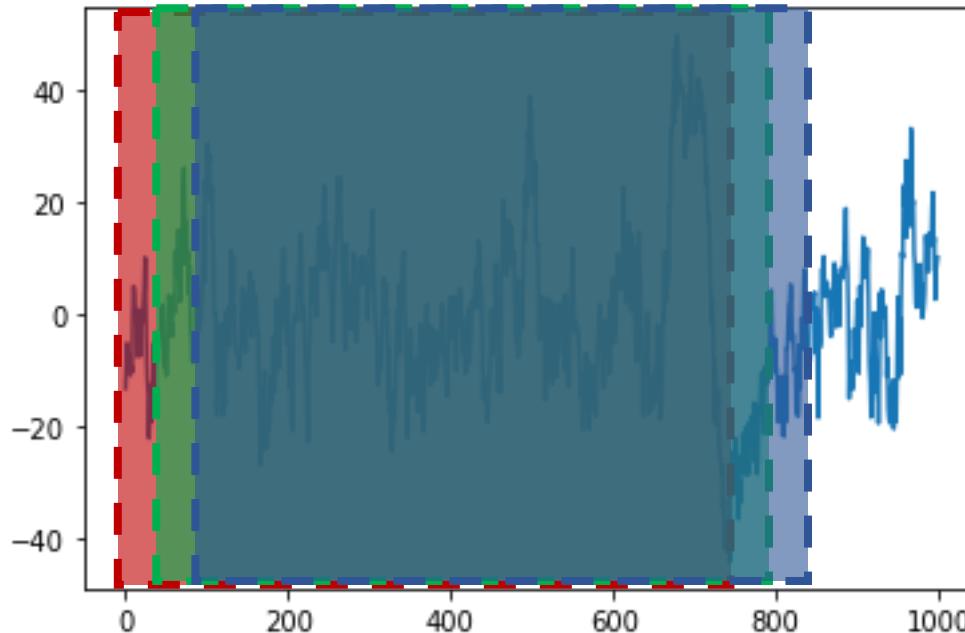
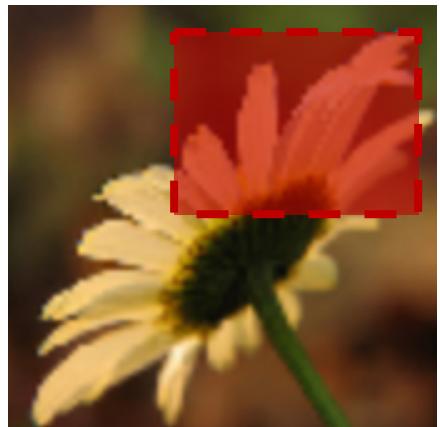
Model – CRNN Cascade



Outline

- Introduction
- Deep Learning Model Design
- **Preprocessing and Optimization**
- Ablation Study and Test
- Failure Case Analysis
- Summary

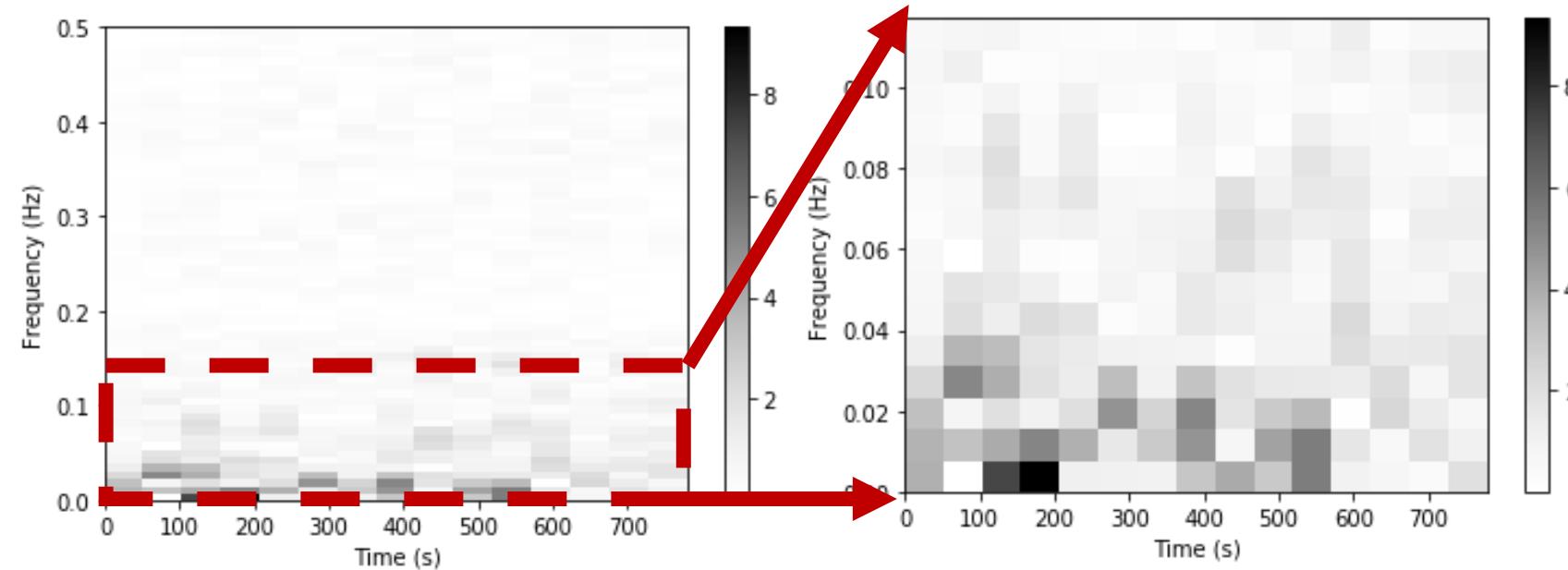
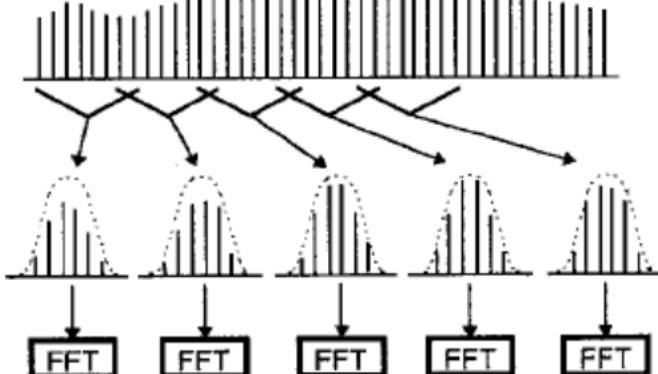
Data Augmentation – Cropping



- Cropping, enlarges training data
- Prevent overfitting
- Better performance

Feature Extraction – STFT

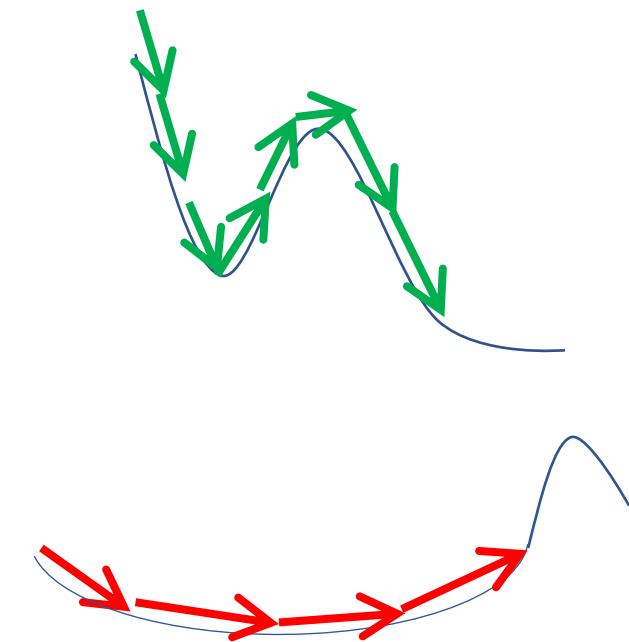
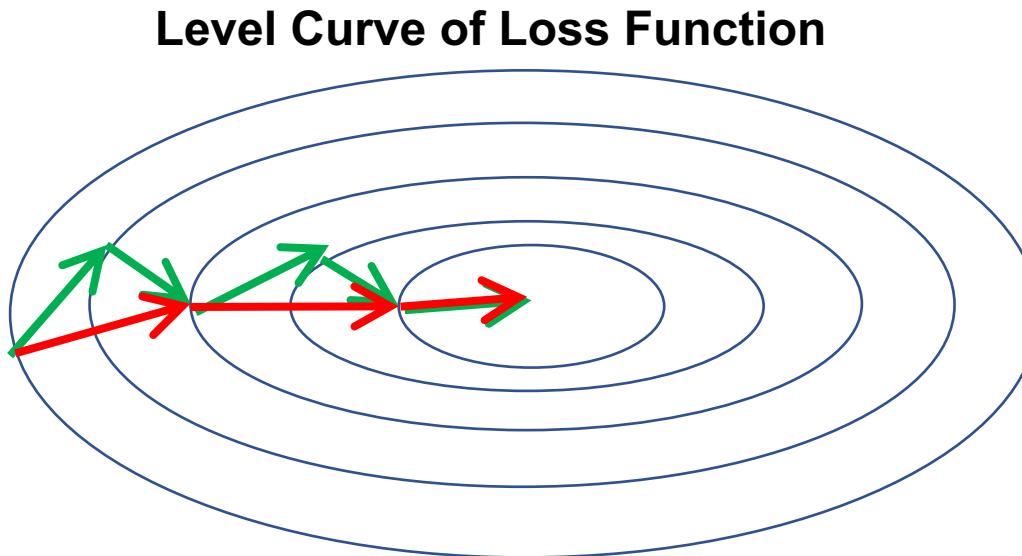
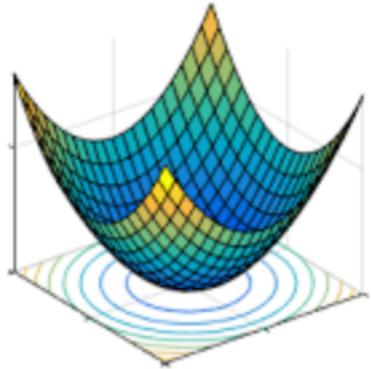
- STFT (Short Time Fourier Transform)
- Extract the frequency feature for small time frame
- Frequency evolution with time



Low Pass Filter

Adam Optimizer

- Adam optimizer = Moving average + Learning rate reduction
- Faster learning
- Avoid deep local minima



Training Details

- GPU 1xTesla K80
- Adam optimizer
- Batch size: 128
- Data size:
 - Initial input shape: 2115
 - After data processing: 19035
- Learning rate: 1e-02
- Iterations: 1k
- Data augmentation: Cropping (crop size 750s on time scale)
- Feature extraction: STFT

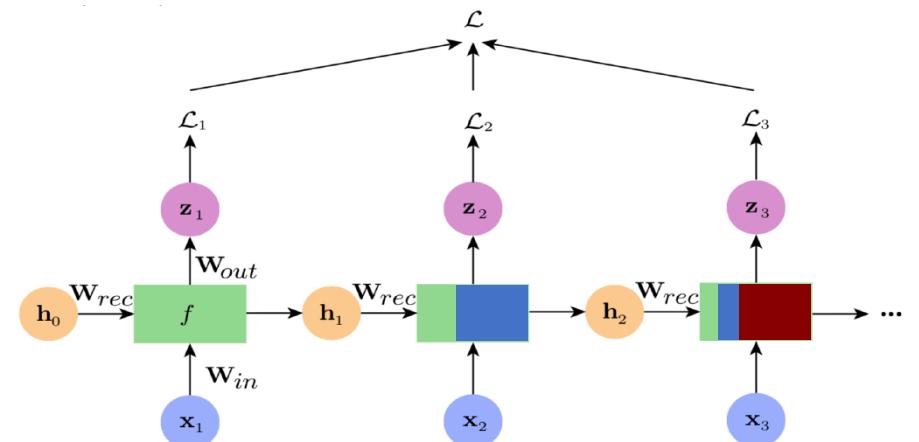
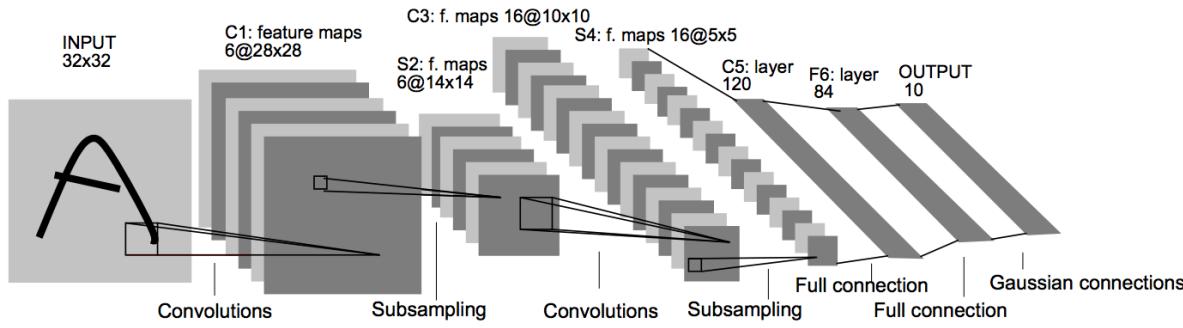


Outline

- Introduction
- Deep Learning Model Design
- Preprocessing and Optimization
- **Ablation Study and Test**
- Failure Case Analysis
- Summary

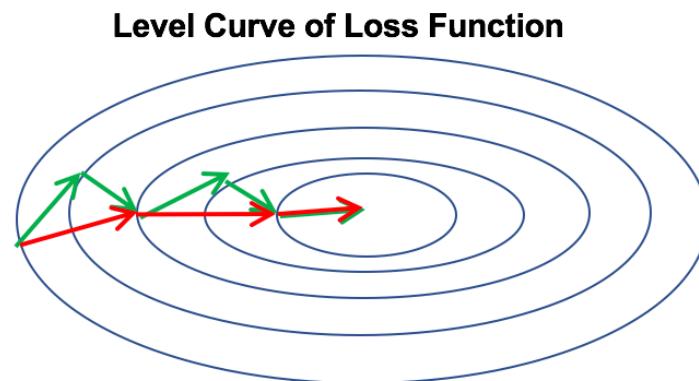
Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279



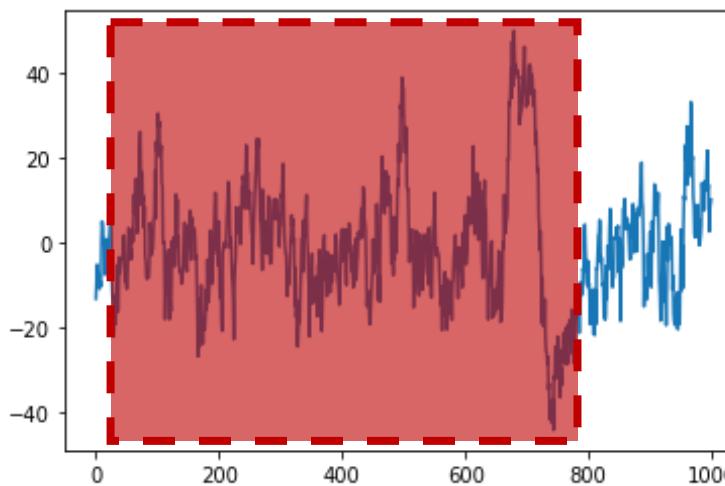
Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279
✓	✓				43.1%	1,657,287



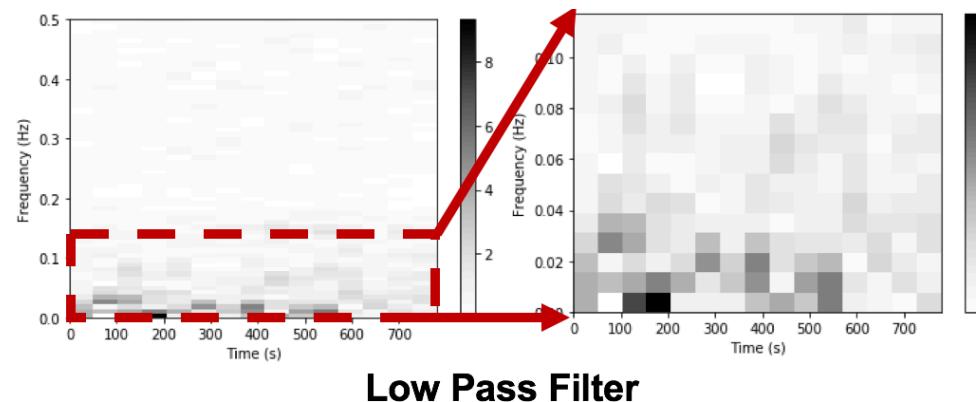
Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279
✓	✓				43.1%	1,657,287
✓	✓	✓			44.3%	295,580,889



Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279
✓	✓				43.1%	1,657,287
✓	✓	✓			44.3%	295,580,889
✓	✓	✓	✓	✓	49.6%	349,353



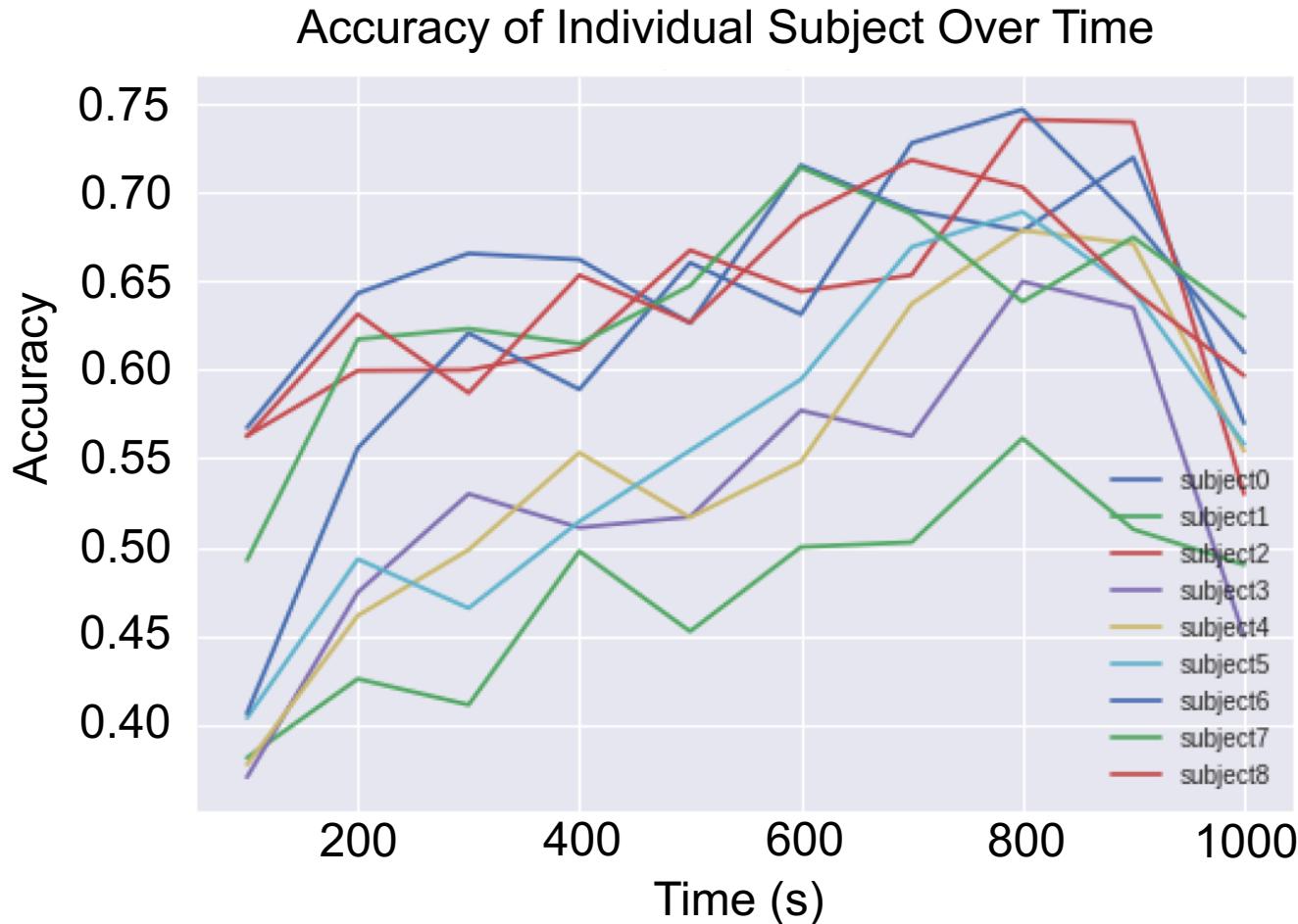
Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279
✓	✓				43.1%	1,657,287
✓	✓	✓			44.3%	295,580,889
✓	✓	✓	✓		49.6%	349,353
✓	✓	✓	✓	✓ Parallel	55.6%	1,200,356
✓	✓	✓	✓	✓ Cascade	53.3%	9,937,802

Ablation Study

CNN	Adam	Cropping	STFT	RNN	Accuracy of All Subjects	Flops
				✓	21.7%	201,642
✓					25.1%	1,657,279
✓	✓				43.1%	1,657,287
✓	✓	✓			44.3%	295,580,889
✓	✓	✓	✓		49.6%	349,353
✓	✓	✓	✓	✓ Parallel	55.6%	1,200,356
✓	✓	✓	✓	✓ Cascade	53.3%	9,937,802

Test Over Time

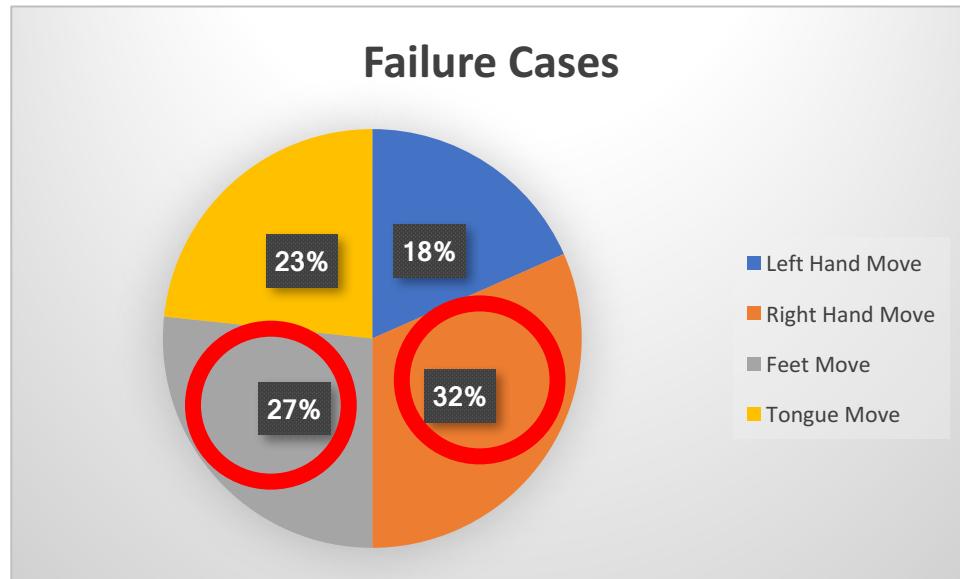


- **Performance consistent for different subjects**
- **Accuracy increases until 800s**
- **Drop data after 800s to improve performance**

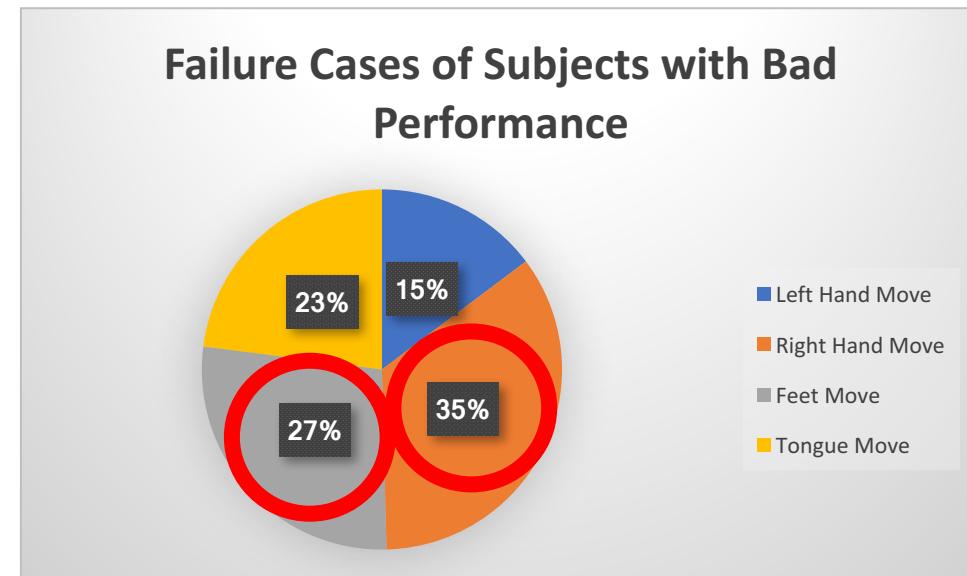
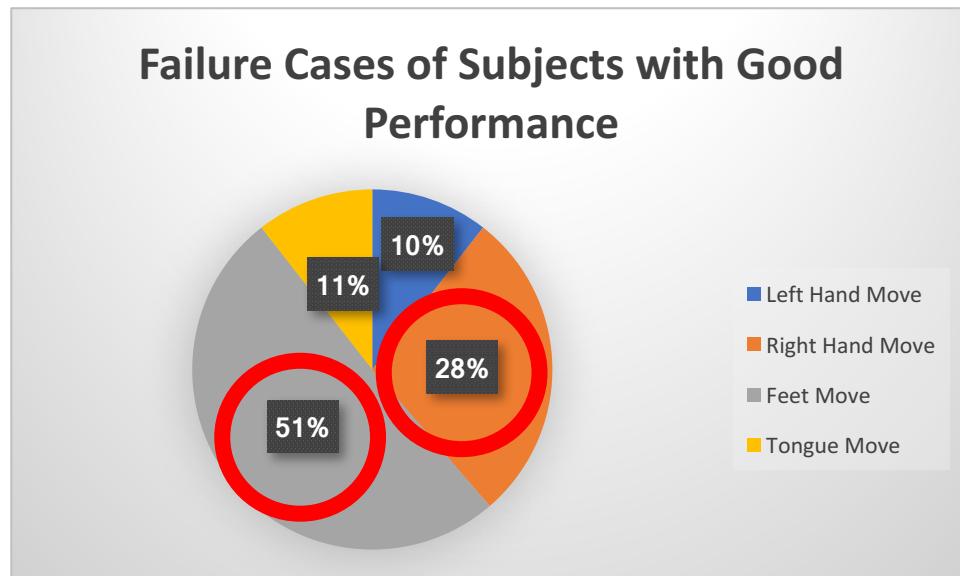
Outline

- Introduction
- Deep Learning Model Design
- Preprocessing and Optimization
- Ablation Study and Test
- **Failure Case Analysis**
- Summary

Failure Case Analysis

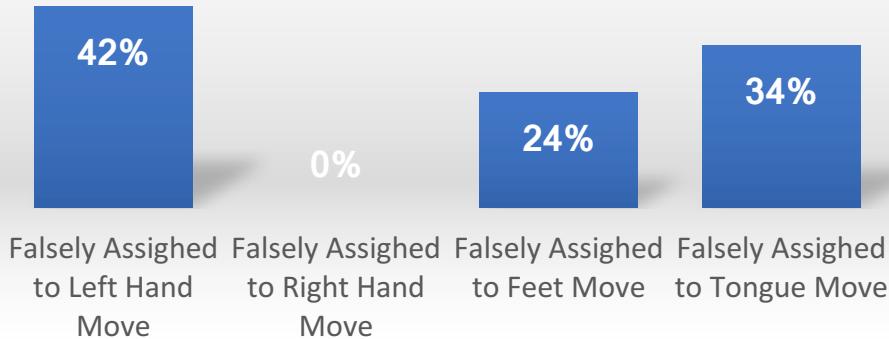


- Overall result consistent with individual subjects
- We could keep data of subject with bad performance

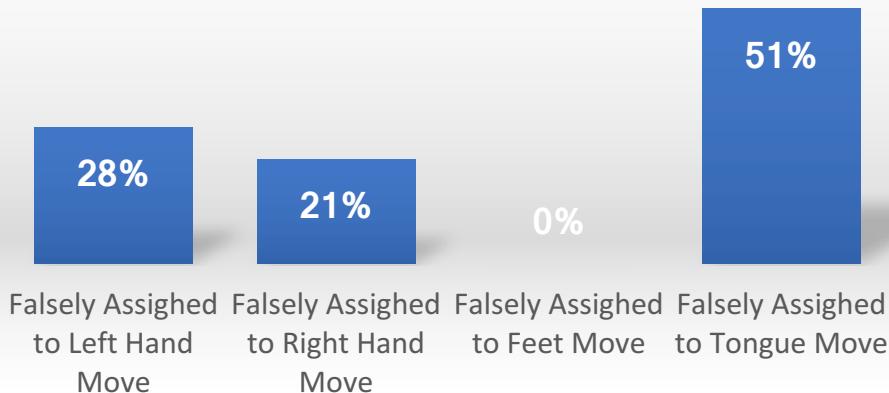


Failure Case Analysis

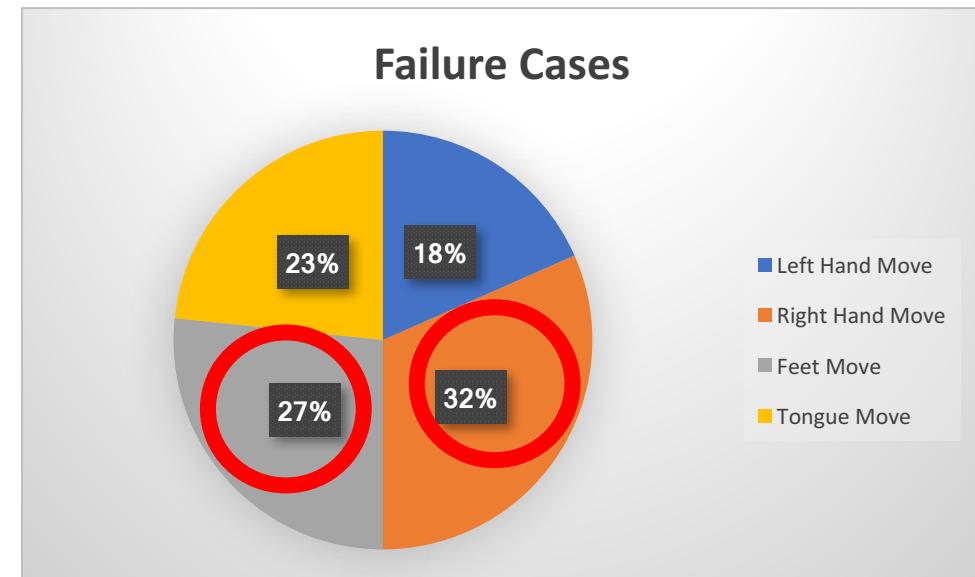
Right Hand Move Failure Cases Assignment



Feet Move Failure Cases Assignment

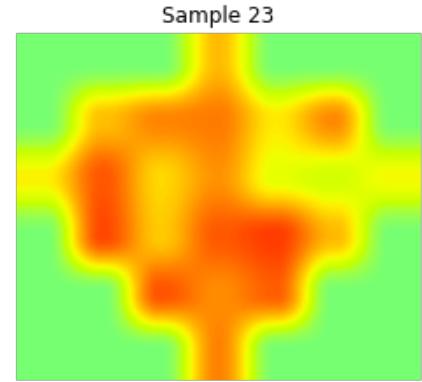
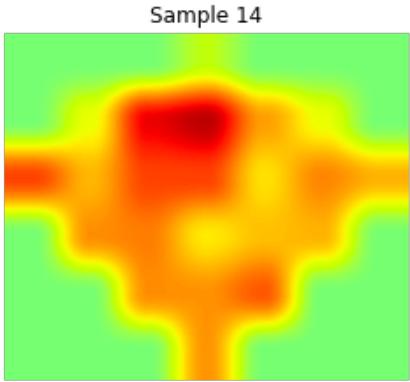


- **Right Hand Move falsely assigned to Left Hand Move or Tongue Move**
- **Feet Move falsely assigned to Tongue Move**



Right Hand Move Failure Cases

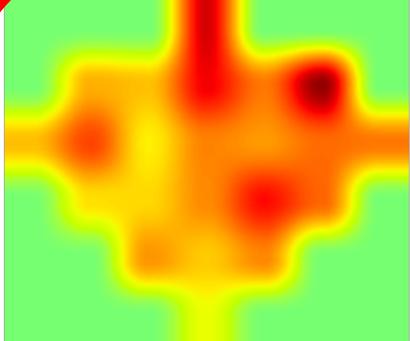
Right Hand Move Failure Cases



Failure reason:

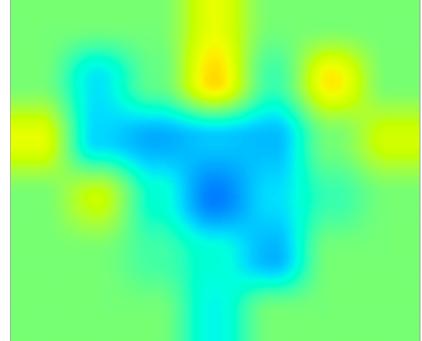
- Overall high response voltage

Left Hand Move EEG



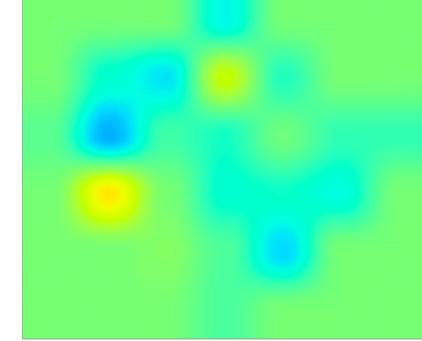
Left

Right Hand Move EEG



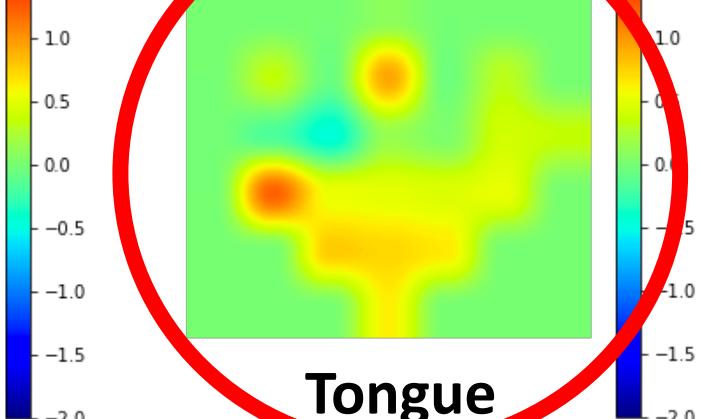
Right

Feet Move EEG



Feet

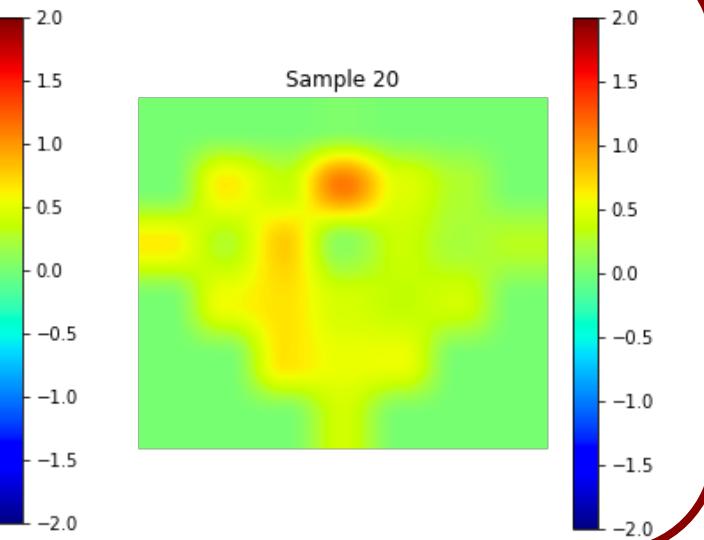
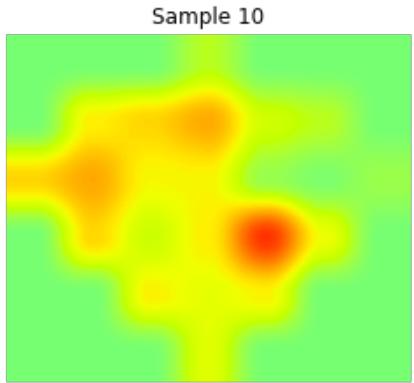
Tongue Move EEG



Tongue

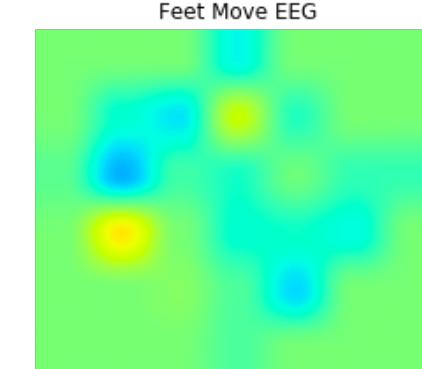
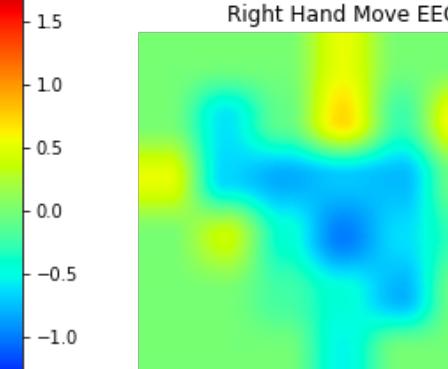
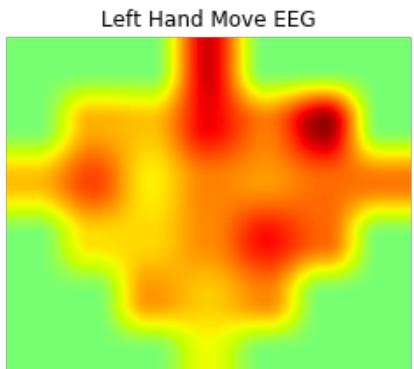
Feet Move Failure Cases

Feet Move Failure Cases



Failure reason:

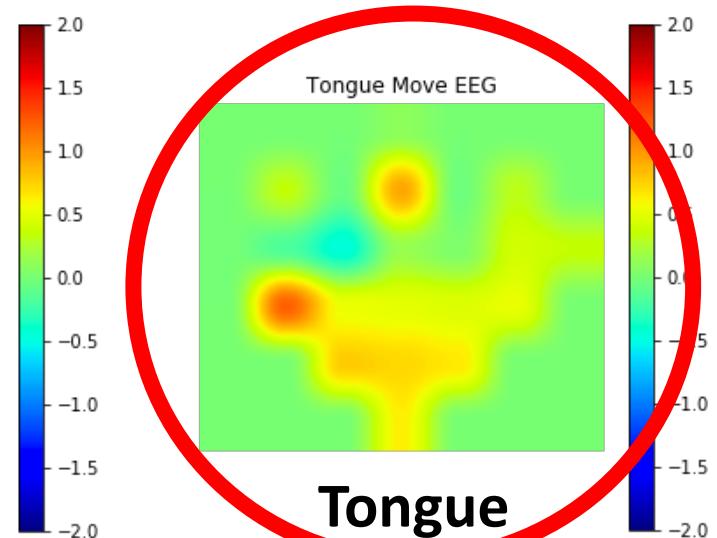
- **High response voltages distribute sparsely**
- **Overall positive voltage**



Left

Right

Feet



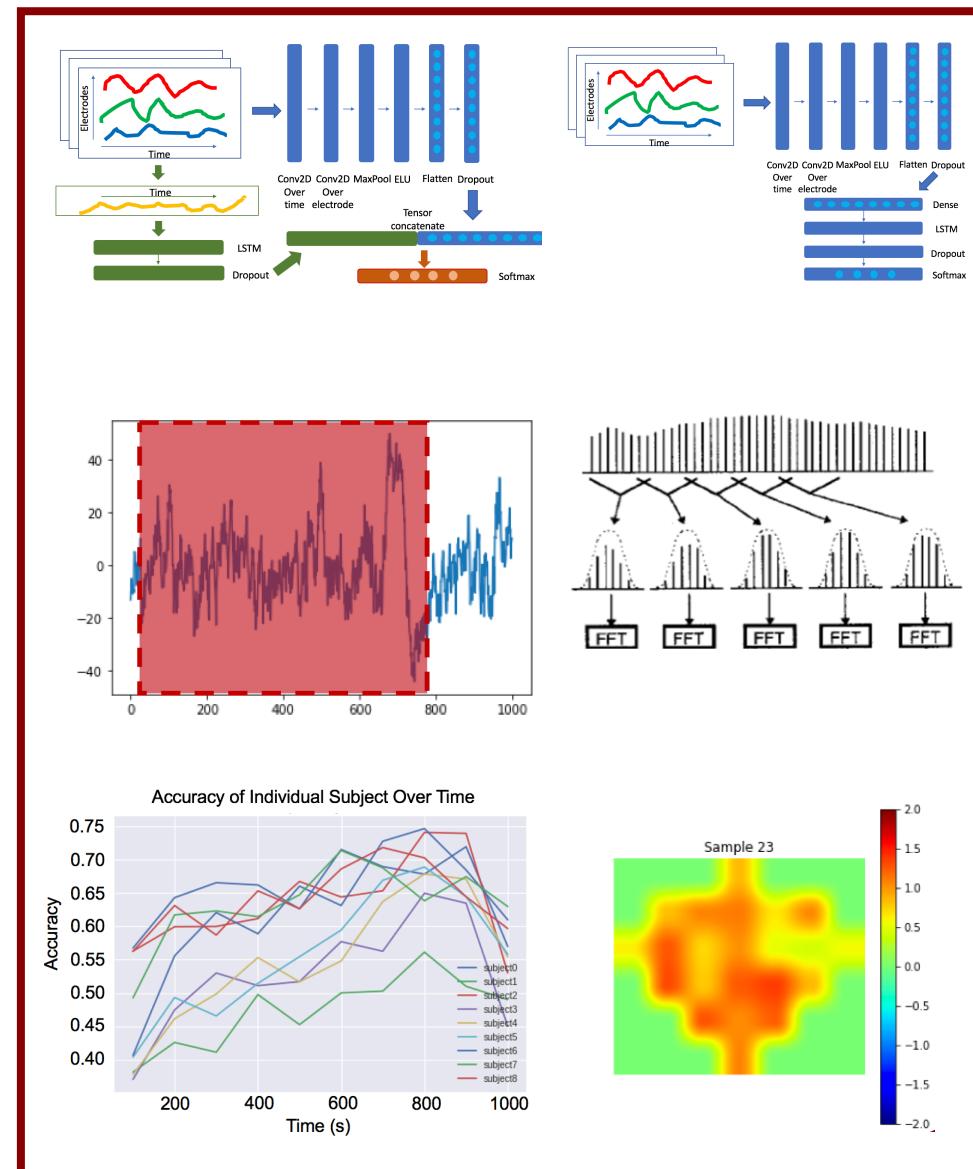
Tongue

Outline

- **Introduction**
- Deep Learning Model Design
- Preprocessing and Optimization
- Ablation Study and Test
- Failure Case Analysis
- **Summary**

Summary

- **Innovation:**
 - Developed CRNN model
 - Preprocessed data: cropping, STFT
- **Did ablation test for model validation**
 - Accuracy for all subject 56%, for single subject 74%
 - Small number of flops, 1M
- **Did failure analysis:**
 - Most failure cases are Right Hand Move (32%) and Feet Move (27%)
 - Failure reason is high voltage value
- **Model weakness and Improvements:**
 - Consider dropping noisy data after 800s
 - Train data at different frequency bands



Thank you !